CHEMICAL PROCESS DYNAMICS AND CONTROLS

Peter Woolf et al.
University of Michigan

LibreTexts™

University of Michigan

Chemical Process Dynamics and Controls

Peter Woolf et al.

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.

LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

The LibreTexts name, logo, and book covers are trademarked by LibreTexts, Inc. (a 501(c)(3) not-for-profit organization) and protected against unauthorized use.

This text was compiled on 08/26/2026

TABLE OF CONTENTS

Licensing

1: Overview

1.1: Introduction to Controls- Background and design methodology

1.2: Introduction to DCS- Control system hardware

1.3: Current Significance- Process controls and you

1.4: Failures in Process Control- Bhopal, Three Mile Island

1.5: Process Controls in Everyday Life- Applying process control thinking to everyday situations

2: Modeling Basics

2.1: Verbal Modeling- process description, control specifications, and connections

2.2: Degrees of Freedom- importance, calculation procedure, and examples

2.3: Incidence Graphs -Interpretations, Consistency, and Inconsistency

2.4: Excel Modeling - logical models, optimization with solver for nonlinear regression, sampling random numbers

2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles

2.6: Numerical ODE solving in Excel- Euler's method, Runge Kutta, Dead time in ODE solving

2.7: Solving ODEs with Mathematica- How to find numerical and analytical solutions to ODEs with Mathematica

2.8: Fitting ODE parameters to data using Excel- Using regression to fit complex models in Excel

2.9: Helpful Mathematica Syntax- Hints on how to use Mathematica to model chemical processes

3: Sensors and Actuators

3.1: Control Systems - Measurement Devices

3.2: Temperature Sensors

3.3: Pressure Sensors

3.4: Level Sensors

3.5: Flow Sensors

3.6: Composition Sensors

3.7: pH and Viscosity Sensors

3.8: Miscellaneous Sensors

3.9: Valves - Types, Kinds, and Selection

3.10: Valves - Modeling Dynamics

4: Piping and Instrumentation Diagrams

4.1: PandID General Information

4.2: Piping and Instrumentation Diagram Standard Notation

4.3: Piping and Instrumentation Diagrams - Location of Controls and Standard Control Structures

4.4: Piping and Instrumentation Diagrams - Standard Pitfalls

4.5: Safety Features in Piping and Instrumentation Diagrams

4.6: Regulatory Agencies and Compliance

5: Logical Modeling

5.1: Boolean Models - Truth Tables and State Transition Diagrams

5.2: Logical Control Programs - IF... THEN... WHILE...

6: Modeling Case Studies

6.1: Surge Tank Model

6.2: Heated Surge Tank

6.3: Bacterial Chemostat

6.4: ODE and Excel CSTR model with heat exchange

6.5: ODE and Excel model of a Simple Distillation Column

6.6: ODE and Excel model of a Heat Exchanger

6.7: ODE and Excel Model of an Adiabatic PFR

6.8: Modeling and PID Controller Example - Cruise Control for an Electric Vehicle

6.9: Blood Glucose Control in Diabetic Patients

7: Mathematics for Control Systems

7.1: Dirac delta (impulse) function

7.2: First-order Differential Equations

7.3: Second-order Differential Equations

7.4: Taylor Series

7.5: Laplace Transforms

8: Optimization

8.1: Introduction to Optimization

8.2: Linear Optimization

8.3: Non-linear Optimization

9: Proportional-Integral-Derivative (PID) Control

9.1: Constructing Block Diagrams- Visualizing control measurements

9.2: P, I, D, PI, PD, and PID control

9.3: PID Tuning via Classical Methods

9.4: PID tuning via Frequency Responses with Bode Plots

9.5: PID tuning via Optimization

9.6: PID Downsides and Solutions

10: Dynamical Systems Analysis

10.1: Finding fixed points in ODEs and Boolean models

10.2: Linearizing ODEs

10.3: Eigenvalues and Eigenvectors

10.4: Using Eigenvalues and Eigenvectors to Find Stability and Solve ODEs

10.5: Phase Plane Analysis - Attractors, Spirals, and Limit cycles

10.6: Root Locus Plots - Effect of Tuning

10.7: Routh Stability - Ranges of Parameter Values that are Stable

11: Control Architectures

11.1: Feedback Control

11.2: Feed Forward Control

11.3: Cascade Control

11.4: Ratio Control

11.5: Common control loops and model for liquid pressure and liquid level

11.6: Common Control Loops and Model for Temperature Control

11.7: Common control architectures and model for reactors

12: Multiple Input, Multiple Output (MIMO) Control

12.1: Determining if a system can be decoupled

12.2: MIMO control using RGA

12.3: MIMO using Model Predictive Control

12.4: Neural Networks for automatic model construction

12.5: Understanding MIMO Control Through Two Tanks Interaction

13: Statistics and Probability Background

13.1: Basic statistics- mean, median, average, standard deviation, z-scores, and p-value

13.2: SPC- Basic Control Charts- Theory and Construction, Sample Size, X-Bar, R charts, S charts

13.3: Six Sigma- What is it and what does it mean?

13.4: Bayes Rule, Conditional Probability and Independence

13.5: Bayesian Network Theory

13.6: Learning and analyzing Bayesian networks with Genie

13.7: Occasionally Dishonest Casino? Markov Chains and Hidden Markov Models

13.8: Continuous Distributions- normal and exponential

13.9: Discrete Distributions - Hypergeometric, Binomial, and Poisson

13.10: Multinomial Distributions

13.11: Comparisons of two means

13.12: Factor analysis and ANOVA

13.13: Correlation and Mutual Information

13.14: Random sampling from a stationary Gaussian process

14: Design of Experiments

14.1: Design of Experiments via Taguchi Methods - Orthogonal Arrays

14.2: Design of experiments via factorial designs

14.3: Design of Experiments via Random Design

14.4: Summary- Summary on Control Architectures' philosophies, advantages, and disadvantages.

Index

Glossary

Detailed Licensing

This page titled Chemical Process Dynamics and Controls (Woolf) is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

Licensing

A detailed breakdown of this resource's licensing can be found in ***Back Matter/Detailed Licensing***.

This page titled Licensing is shared under a license and was authored, remixed, and/or curated by LibreTexts.

1: OVERVIEW

1.1: Introduction to Controls- Background and design methodology
1.2: Introduction to DCS- Control system hardware
1.3: Current Significance- Process controls and you
1.4: Failures in Process Control- Bhopal, Three Mile Island
1.5: Process Controls in Everyday Life- Applying process control thinking to everyday situations

This page titled 1: Overview is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

1.1: Introduction to Controls- Background and design methodology

Introduction

Process controls is a mixture between the statistics and engineering discipline that deals with the mechanism, architectures, and algorithms for controlling a process. Some examples of controlled processes are:

- Controlling the temperature of a water stream by controlling the amount of steam added to the shell of a heat exchanger.
- Operating a jacketed reactor isothermally by controlling the mixture of cold water and steam that flows through the jacket of a jacketed reactor.
- Maintaining a set ratio of reactants to be added to a reactor by controlling their flow rates.
- Controlling the height of fluid in a tank to ensure that it does not overflow.

To truly understand or solve a design problem it is necessary to understand the key concepts and general terminology. The paragraphs below provide a brief introduction to process controls as well as some terminology that will be useful in studying controls. As you begin to look at specific examples contained here, as well as elsewhere on the wiki, you will begin to gain a better grasp on how controls operate and function as well as their uses in industry.

Process Control Background

The role of process control has changed throughout the years and is continuously shaped by technology. The traditional role of process control in industrial operations was to contribute to safety, minimized environmental impact, and optimize processes by maintaining process variable near the desired values (1). Generally, anything that requires continuous monitoring of an operation involve the role of a process engineer. In years past the monitoring of these processes was done at the unit and were maintained locally by operator and engineers. Today many chemical plants have gone to full automation, which means that engineers and operators are helped by DCS that communicates with the instruments in the field.

What are the benefits of Process Control?

The benefits of controlling or automating process are in a number of distinct area in the operation of a unit or chemical plant. Safety of workers and the community around a plant is probably concern number one or should be for most engineers as they begin to design their processes. Chemical plants have a great potential to do severe damage if something goes wrong and it is inherent the setup of process control to set boundaries on specific unit so that they don't injure or kill workers or individuals in the community.

The Objectives of Control

A control system is required to perform either one or both task:

1. Maintain the process at the operational conditions and set points

Many processes should work at steady state conditions or in a state in which it satisfies all the benefits for a company such as budget, yield, safety, and other quality objectives. In many real-life situations, a process may not always remain static under these conditions and therefore can cause substantial losses to the process. One of the ways a process can wander away from these conditions is by the system becoming unstable, meaning process variables oscillate from its physical boundaries over a limited time span. An example of this would be a water tank in a heating and cooling process without any drainage and is being constantly filled with water. The water level in the tank will continue to rise and eventually overflow. This uncontrolled system can be controlled simply by adding control valves and level sensors in the tank that can tell the engineer or technician the level of water in the tank. Another way a process can stray away from steady state conditions can be due to various changes in the environmental conditions, such as composition of a feed, temperature conditions, or flow rate.

2. Transition the process from one operational condition to another

In real-life situations, engineers may change the process operational conditions for a variety of different reasons, such as customer specifications or environment specifications. Although, transitioning a process from one operational condition to another can be detrimental to a process, it also can be beneficial depending on the company and consumer demands.

Examples of why a process may be moved from one operational set point to another:

1. Economics
2. Product specifications
3. Operational constraints
4. Environmental regulations
5. Consumer/Customer specifications
6. Environmental regulations
7. Safety precautions

Definitions and Terminology

In controlling a process there exist two types of classes of variables.

1. **Input Variable** – This variable shows the effect of the surroundings on the process. It normally refers to those factors that influence the process. An example of this would be the flow rate of the steam through a heat exchanger that would change the amount of energy put into the process. There are effects of the surrounding that are controllable and some that are not. These are broken down into two types of inputs.

a. *Manipulated inputs:* variable in the surroundings can be control by an operator or the control system in place.
b. *Disturbances:* inputs that can not be controlled by an operator or control system. There exist both measurable and immeasurable disturbances.

2. **Output variable**- Also known as the *control variable* These are the variables that are process outputs that effect the surroundings. An example of this would be the amount of CO2 gas that comes out of a combustion reaction. These variables may or may not be measured.

As we consider a controls problem. We are able to look at two major control structures.

1. Single input-Single Output (SISO)- for one control(output) variable there exist one manipulate (input) variable that is used to affect the process
2. Multiple input-multiple output(MIMO)- There are several control (output) variable that are affected by several manipulated (input) variables used in a given process.

- **Cascade:** A control system with 2 or more controllers, a "Master" and "Slave" loop. The output of the "Master" controller is the setpoint for the "Slave" controller.
- **Dead Time:** The amount of time it takes for a process to start changing after a disturbance in the system.
- **Derivative Control:** The "D" part of a PID controller. With derivative action the controller output is proportional to the rate of change of the process variable or error.*
- **Error:** In process controls, error is defined as: Error = setpoint - process variable.
- **Integral Control:** The "I" part of a PID controller. With integral action the controller output is proportional to the amount and duration of the error signal.
- **PID Controller:** PID controllers are designed to eliminate the need for continuous operator attention. They are used to automatically adjust system variables to hold a process variable at a setpoint. Error is defined above as the difference between setpoint and process variable.
- **Proportional Control:** The "P" part of a PID controller. With proportional action the controller output is proportional to the amount of the error signal.
- **Setpoint:** The setpoint is where you would like a controlled process variable to be.

Design Methodology for Process Control

1. **Understand the process:** Before attempting to control a process it is necessary to understand how the process works and what it does.

2. **Identify the operating parameters:** Once the process is well understood, operating parameters such as temperatures, pressures, flow rates, and other variables specific to the process must be identified for its control.
3. **Identify the hazardous conditions:** In order to maintain a safe and hazard-free facility, variables that may cause safety concerns must be identified and may require additional control.
4. **Identify the measurables:** It is important to identify the measurables that correspond with the operating parameters in order to control the process.

Measurables for process systems include:

- Temperature
- Pressure
- Flow rate
- pH
- Humidity
- Level
- Concentration
- Viscosity
- Conductivity
- Turbidity
- Redox/potential
- Electrical behavior
- Flammability

5. **Identify the points of measurement:** Once the measurables are identified, it is important locate where they will be measured so that the system can be accurately controlled.

6. **Select measurement methods:** Selecting the proper type of measurement device specific to the process will ensure that the most accurate, stable, and cost-effective method is chosen. There are several different signal types that can detect different things.

These signal types include:

- Electric
- Pneumatic
- Light
- Radiowaves
- Infrared (IR)
- Nuclear

7. **Select control method:** In order to control the operating parameters, the proper control method is vital to control the process effectively. On/off is one control method and the other is continuous control. Continuous control involves Proportional (P), Integral (I), and Derivative (D) methods or some combination of those three.

8. **Select control system:** Choosing between a local or distributed control system that fits well with the process effects both the cost and efficacy of the overall control.

9. **Set control limits:** Understanding the operating parameters allows the ability to define the limits of the measurable parameters in the control system.

10. **Define control logic:** Choosing between feed-forward, feed-backward, cascade, ratio, or other control logic is a necessary decision based on the specific design and safety parameters of the system.

11. **Create a redundancy system:** Even the best control system will have failure points; therefore it is important to design a redundancy system to avoid catastrophic failures by having back-up controls in place.

12. **Define a fail-safe:** Fail-safes allow a system to return to a safe state after a breakdown of the control. This fail-safe allows the process to avoid hazardous conditions that may otherwise occur.

13. **Set lead/lag criteria:** Depending on the control logic used in the process, there may be lag times associated with the

measurement of the operating parameters. Setting lead/lag times compensates for this effect and allow for accurate control.

14. **Investigate effects of changes before/after:** By investigating changes made by implementing the control system, unforeseen problems can be identified and corrected before they create hazardous conditions in the facility.

15. **Integrate and test with other systems:** The proper integration of a new control system with existing process systems avoids conflicts between multiple systems.

References

1. Romagnoli, Jose A. Introduction to Process Control. s.l. : CRC press, 2006.

This page titled 1.1: Introduction to Controls- Background and design methodology is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

1.2: Introduction to DCS- Control system hardware

Introduction

Digital Control Systems(DCS) also known as Distributed Control System is the brain of the control system. It is used mainly for the automation of a manufacturing process and manages the logic that exist for major unit operations. A DCS in the past was tailor made for the process, plant or company that intended to use the structure to control and model it's process. Before the beginning of the DCS era there were pneumatic devices that controlled process and engineers manually turned valves on the site. Modeling of the systems was made possible by DCS as it allowed the ability to record and manage process from comfort of a computer screen. Because of DCS we are able to control processes remotely and gain a better understanding of how the process operate and how they can be improved to both increase safety and increase profit possibilities.

Control Systems are collectively named as "ICSS" Integrated Control and Safety System. Distinctly identified as "BPCS" Basic Process Control System. "SIS" Safety Instrumentation System. "F&G" Fire and Gas System.

DCS is employed in BPCS as well as used and prevalent control system. How does a DCS work?

In the field you have sensors and gauges that give and recieve information. They convert this information into a electric signal that is sent to a control room somewhere in the field. This control room has programmed logic that is able to converts the signal into a pressure, flow rate, concentration, temperature, or level. This logic also contains the information that controls the process and takes the signal compares it with the set point sent from the operator may or may not be in the field and sends a signal to the manipulated variables in the field. The DCS covers all of the computer logic from the operator screen to the field box that contain the logic.

Shutdown systems

Shutdown systems are the emergency setting of the logic to make sure the process can be contained and is environmentally safe. These setting are important for emergency response of the system. It is the job of the DCS to contain the logic for the shutdown system and be able to operate when a process exceed a certain limit.

This page titled 1.2: Introduction to DCS- Control system hardware is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

1.3: Current Significance- Process controls and you

Authors: (December 12, 2009) Steve Dzul, Steve Pankratz, Derrick Boroski

Introduction

Industrial processes are central to the chemical engineering discipline. Generally, processes are controlled in order to do things such as maximize safety, minimize cost, or limit effects on the environment. This course aims to help undergraduate engineering students understand the mechanisms used to moderate these processes, such as to control their output.

Automation

Generally, process controls are designed to be automated. This means that given a change in system response, the control system can act on its own to account for it. In order to minimize cost, automated systems have become widespread throughout industry. Before automation, a huge amount of labor would be required to run even the simplest processes. For example, a technician might be hired to monitor the temperature in a reaction vessel, and operate a valve to manipulate the cooling water flow rate in the jacket. In a sense, this technician operated as a control system. If the temperature reading is too high, the technician will manipulate the system in order to bring the temperature down. Via automation, this simple, arduous labor can be done by an algorithm.

By designing an effective control system, even the most complicated of processes can be run with minimal worker supervision. Telephone operators, for example, have largely been replaced by automated telephone switch boards. Removing the need for telephone operators decreases operating cost for phone companies, thereby allowing the general consumer to pay less for phone service. Automated process controls, therefore, are enormously important in the modern world.

This page titled 1.3: Current Significance- Process controls and you is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

1.4: Failures in Process Control- Bhopal, Three Mile Island

Authors: (December 12, 2009) Steve Dzul, Steve Pankratz, Derrick Boroski

Process controls can have a huge impact on surrounding communities, as well as the environment. An engineer of a large-scale process, therefore, has an important ethical responsibility to operate a process safely and properly. These responsibilities extend well beyond the scope of merely the company for which they work. Catastrophic failures in process control remind us of the importance of control systems and engineering in today's world.

Bhopal, India Disaster

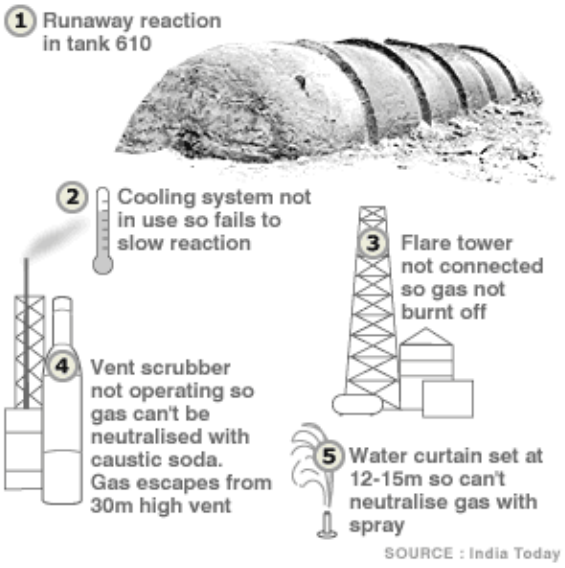
Overview

The Bhopal Gas Tragedy in Bhopal, India on December 3, 1984 was a large toxic gas leak that killed thousands of people in the surrounding area. A tank with 42 tons of methyl isocyanate(MIC) was contaminated with water. This in turn caused a run away reaction that greatly increased the pressure and temperatures in the tank, which forced the emergency venting of the toxic gases to the atmosphere.

This tragedy was largely due to the failure or lack of safety controls:

1. Runaway reaction as temperature and pressure increased without regulation
2. MIC was supposed to be cooled, however in the Bhopal plant the refrigeration system was not turned on. Temperature control on the tank could have greatly hindered the runaway reaction that ensued with the addition of water.
3. Flare tower to handle the leakage of toxic gases was not functional
4. The plant also had vent scrubbers, which were also not functional
5. Water curtain, which would neutralize some escaping gas, not designed properly. It was not tall enough to reach the top of the flare tower, making it essentially worthless.
6. Alarms that would have alerted to a malfunction in the tank had not been operational for 4 years

The figure below illustrates some of these failures:



Had at least some of these been functioning the amount of toxic gas released would have been substantially reduced.

Results

From this tragedy we can see that if the plant had proper safety controls the effects of the disaster would have been greatly reduced. Therefore as a chemical engineer it is our responsibility to society to provide sufficient safety controls to chemical processes in order to prevent disasters such as the Bhopal Gas Tragedy from happening. Unfortunately, industrial negligence is still a problem in many third-world countries.

Reference

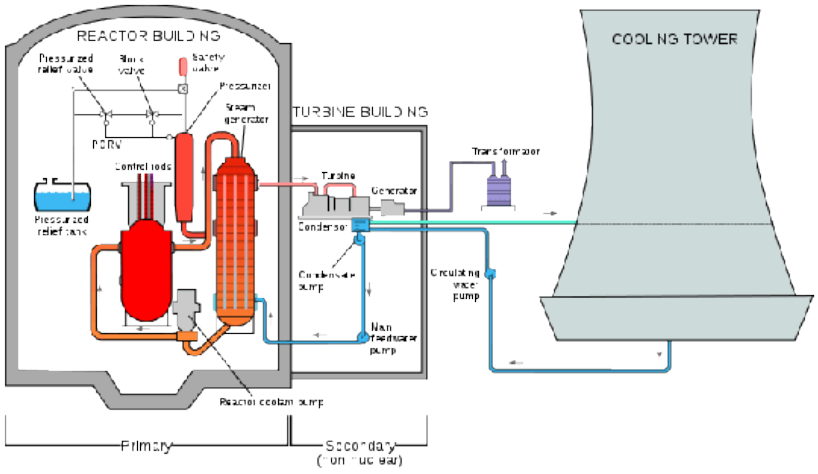
http://www.corrosion-doctors.org/Pollution/bhopal.htm

Three Mile Island Disaster

Overview

One of the largest and most far reaching plant failures in United States history took place at a nuclear power plant on Three Mile Island in March 1979. The event was caused by either a mechanical or electrical failure of the main feed water pumps causing the power plant to begin to overheat. As the heat increased, the control scheme caused the turbine and reactor to shut down. This caused a pressure increase in the primary system (nuclear portion of the plant) and a relief valve automatically opened to release some of the pressure to prevent the reactor from blowing. All of these actions were well designed to prevent a significant event from happening. The problem was that the release valve did not close properly when the pressure in the reactor was relieved. As a result, when the reactor started back up, coolant in the core of the reactor was lost through the pressure relief valve. Because there was no control mechanism that measured the level of the coolant in the reactor, the operators, who only judged the water level by the pressure in the reactor, actually decreased coolant flow to the reactor.

The figure below is a simplified diagram of the TMI-2-plant:



The result of the control design failure that prevented the operators from cooling the reactor was that the rods that held the nuclear fuel melted causing the fuel to also melt. This is the worst thing to have happen in a nuclear power plant and is what happened to cause the disaster at Chernobyl. Thankfully, the accident was largely contained and although the entire nation watched for 3 days as the threat of an explosion or breach of containment loomed, 0 deaths or injuries resulted. In fact, corrective steps were so successful that the average increase in radiation to the surrounding population was around 1% and the maximum increase at the boundary to the site is estimated to be less than 100% of the natural background radiation present in the region.

Results

The accident at Three Mile Island showed the importance of proper design of control systems. As a result the US Nuclear Regulatory Commission took steps to tighten their regulation and increase the safety requirements on Nuclear Power Plants. These included revamping operator training as well as increasing the design and equipment requirements. This also brought the dangers of all industrial processes to the forefront and reminded people of the importance of the safety of the communities surrounding chemical and power plants.

Unfortunately, the incident also inspired intense fear of nuclear power in the general population and is partially responsible for the reduced build rate for new nuclear power plants since that time. Although control failures can be corrected fairly quickly, after one safety issue it is difficult to convince the general public that engineers have fixed the problem and that it will not happen again.

Reference

References: www.nrc.gov/reading-rm/doc-collections/fact-sheets/3mile-isle.html: The United States Nuclear Regulatory Commission

Texas City Disaster

Authors: (December 14, 2009) Virgil Humes, Fred Garner

Overview

In March of 2005, 15 people were killed at the of a refinery explosion in Texas City, TX. The refinery is the third largest in the country with the capability to process over 400,000 barrels of crude oil daily. On this particular day, the isomerization unit (whose purpose is to boost the octane rating of fuels) was being started up. As part of the normal start-up sequence, operators began feeding hydrocarbon liquid into the tower. However, a discharge valve that should have been opened was overlooked. A high level alarm was ignored and a second high-level alarm had not yet been re-enabled from the various maintenance activities that had been going on while the unit was down. Upon realizing that the tower was filling, a discharge valve was opened to release the hot fluid from the bottom of the vessel. This fluid then passed through a heat exchanger, warming the fluid that was still being pumped into the tower. The boiling that resulted in the bottom of the tower caused liquids to spill over the top of the stack and into a blow down drum with an atmospheric vent and no flare. The blow down drum could not contain the volume and began ejecting liquid and vapor hydrocarbon from the atmospheric vent. The vapors were soon ignited by a diesel truck with its ignition on. Fifteen people in a nearby trailer were killed in the resulting explosion. Inadequacies in the written start-up procedures, operator training, and the design of the safety relief system led to tragic and unnecessary loss of life. But examples such as these serve to make process environments even safer places to work in and around.

An animation detailing the circumstances surrounding the explosion was released by the U.S. Chemical Safety Board during the investigation. Some accounts of the incident vary in a few details, but all are in agreement that the written procedures, operator training, and some aspects of the safety system design were to blame.

Results

After the incident at the refinery, representatives from BP said the company would eliminate all blow-down drums/vent stack systems in flammable service. Also, the Chemical Safety Board (CSB) recommended that BP commission an independent panel to investigate the safety culture and management systems of BP. The findings of the panel showed that BP management had not distinguished between occupational safety and process safety. This led to new implementations of process safety throughout BP North America.

Reference

CSB Final Report into March 2005 Texas City Disaster

Wikipedia - Texas City Refinery (BP)

This page titled 1.4: Failures in Process Control- Bhopal, Three Mile Island is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

1.5: Process Controls in Everyday Life- Applying process control thinking to everyday situations

Written By: Andrew Sereno, Randy Tin (FA09)

Introduction

Have you ever wondered if the abstract and arbitrary subject of chemical engineering process controls has any relevance outside of Dow 1013? Below are a few examples of how the concepts learned in process controls can be applied to situations in your everyday life.

Examples of Process Control for the Common Man

Consumption

Spicing Up Soup

Pavlo LaBalle just finished scaling up a mixing tank for use in the manufacture of horse glue. Scaling up mixing tanks for horse glue is hard work - enough to make any man crave a hearty bowl of spicy liver soup. Fortunately, Mrs. LaBalle has a batch of liver soup ready for Pavlo's return home. Using his built-in composition sensor, the tongue, Pavlo realizes the liver soup spice concentration is below his specified setpoint - the soup is not spicy enough. Having never spiced his own soup, but not wanting to offend Mrs. LaBalle, Pavlo attempts to increase the spiciness of the soup using his favorite hot sauce, Sriracha. In order to prevent catastrophic overspicing, Pavlo adds a drop of Sriracha (a differential amount) and immediately begins sampling the soup after addition. In this way, Pavlo is able to determine the effect of a small disturbance (the introduction of Sriracha) on the process output (the soup's spiciness). Using his new knowledge of Sriracha's effect on the soup, Pavlo is able to optimize his addition of Sriracha so as to not catastrophically overspice the soup. Pavlo then adds a new amount of Sriracha estimated to bring the soup closer to the spice setpoint and samples the soup shortly after. By tasting the soup after the addition of spice, Pavlo is able to bring the soup closer to his desired spice setpoint through the use of feed backward spice control. By applying his education in chemical engineering process controls, Pavlo has enabled himself to enjoy his spicy liver soup.

Imbibing Alcohol

Jimmy Johnson just finished a presentation and a final report for Chemical Engineering 460. Unfortunately, during the presentation, Jimmy's personal record of 80 straight hours without sleep came to an abrupt end as he blacked out mid-line, and he was subsequently criticized for insufficiently displaying the iron resolve befitting of a Michigan engineer. Jimmy decided to drown his sorrows at Good Time Charley's through the careful ingestion of mood-improving fluids. Jimmy planned on drinking until he felt a moderate buzz, then leaving to finish his graduate school application to pursue a masters degree in English.

Jimmy, inexperienced with the process of satiating his thirst with such beverages, planned to continuously drink until he felt adequately affected. However, Jimmy failed to take into account the dead time inherent in his body's response to alcohol. Jimmy was unable to control his body's response as he had planned, resulting in a vast overshoot into drunken stupor.

Jimmy woke up with no pants the next morning. Fortunately, by applying his education in chemical engineering, he earned the necessary experience for tuning the amount he should drink in the future.

Purchasing Food at U-Go's

Rachel Malta was busy typing her technical report on the design space of a distillation column in a pilot horse glue plant. It was 11:50 p.m., and Rachel was making good progress until she felt the odd sensation of bubbling acids corroding through her stomach wall, at which point she decided to visit U-Go's to find something to eat.

Confronted with the shop's wide array of processed foods and saturated fats, a difficult problem lay before Rachel. If she purchased too much food, her hunger set point would be reached, but she would be left with bags of uneaten popcorn and sour

patch kids. If she purchased too little, she would remain hungry and would be unable to return to purchase more food after the store closed. Rachel needed to make a prediction of how much food her digestion process would require to reach her desired fullness. Based on her previous experiences with U-Go's "food," her predictive control (her memory) decided to purchase a bag of barbecue chips, one apple, and a bagel. By applying her education in chemical engineering, Rachel was able to successfully predict how much food would satisfy her hunger without being forced to find more food.

Hygiene

Filling a Bathtub

Lan Ri has been in the Duderstadt Center for the past 3 nights working on his ChE 487 project. Having not bathed for 60 hours, Mr. Ri decides it is time to wash himself for the sake of his fellow Duderstadt dwellers, however Mr. Ri has forgotten how to stand up (he's been sitting at a computer for 60 hours). The ever creative Mr. Ri decides to take a bath instead of his usual standing shower. At the bath, Mr. Ri finds himself confronted with two flow rate controllers: one "hot water feed" controller and one "cold water feed" controller. Turning the controls, he realizes that they control the flow rate of the bathtub feed streams. Yet Mr. Ri has a problem - the hot water feed temperature, as measured by his built-in temperature sensor (his skin), is far too high. The cold water feed temperature, also measured using his skin-type temperature sensor, is too low for a comfortable bath. Accordingly, Mr. Ri performs a quick energy balance to determine expected ratio of the feeds necessary to achieve a bath temperature close to what his desired setpoint. However, because Mr. Ri has poor control of the actual flow rate of each stream (hot and cold), the actual resulting temperature is still below that which Mr. Ri desires. Mr. Ri then readjusts the hot-to-cold feed ratio by increasing the hot water feed flow rate. As the bathtub fills up, Mr. Ri uses his built-in level sensor (his eyeballs) to turn off the valves controlling the feed stream flow when the bathtub level is at his calculated level. Mr. Ri was able to calculate the level to which the bathtub should be filled by using his knowledge of static fluid systems. By applying his education in chemical engineering process controls, Mr. Ri has enabled himself to wash away the smell of the Duderstadt Center.

[Can use some diagrams, pictures, and/or equations to describe these problems - RZ].

This page titled 1.5: Process Controls in Everyday Life- Applying process control thinking to everyday situations is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

2: MODELING BASICS

2.1: Verbal Modeling- process description, control specifications, and connections
2.2: Degrees of Freedom- importance, calculation procedure, and examples
2.3: Incidence Graphs -Interpretations, Consistency, and Inconsistency
2.4: Excel Modeling - logical models, optimization with solver for nonlinear regression, sampling random numbers
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles
2.6: Numerical ODE solving in Excel- Euler's method, Runge Kutta, Dead time in ODE solving
2.7: Solving ODEs with Mathematica- How to find numerical and analytical solutions to ODEs with Mathematica
2.8: Fitting ODE parameters to data using Excel- Using regression to fit complex models in Excel
2.9: Helpful Mathematica Syntax- Hints on how to use Mathematica to model chemical processes

This page titled 2: Modeling Basics is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

2.1: Verbal Modeling- process description, control specifications, and connections

Authors: (September 8, 2006) Brian McQuillan, Crystal Miranda, Brandon Quigley, and John Zhang

Stewards: (September 5, 2007) Kevin Luchi, Mike Nappo, Matt Neff, Lisa Schauman

Introduction

Every process requires a great deal of planning in order to successfully accomplish the goals laid out by its designers and operators. In order to accomplish these goals, however, personnel that are not familiar with the design must fully understand the process and the functions of the control systems. Control systems consist of equipment (measuring devices, valves, etc.) and human intervention (plant operators and designers). Control systems are used to satisfy three basic needs of every process:

1. Reduce the influence of external disturbances
2. Promote the stability of the process
3. Enhance the performance of the process

Verbal modeling is used for creating and also understanding a process control system. Verbal modeling consists of first receiving and then gathering information about the process. A step-by-step process is then used to describe the control systems used to satisfy constraints and objectives that have been outlined. In the following sections you will read what requirements are generally outlined for process control and the step-by-step method used to meet these requirements.

Prerequisite Information Regarding a Process

For the sake of this article it is assumed that a process has already been designed and that certain restraints and criteria are provided by either a customer, management, or the government. The goal of this section is to classify the types of criteria that are usually given. These criteria then become the conditions that the control systems employed must satisfy. In general, there will be five sets of criteria, often coming from different people and institutions. By gathering all of these criteria you will be able to describe the control system. If you do not have a complete list of these criteria you must research the process to determine these constraints before beginning the step-by-step process below.

Safety

The safe operation of a process is the biggest concern of those working in the plant and those that live in the surrounding community. The temperatures, pressures, and concentrations within the system should all fall within acceptable limits, and these limits can be dictated by either government agencies or company policy.

Production Objectives

The production objectives usually include both the amount and purity of the desired product. This criterion is generally set by the company or customer.

Environmental Regulations

These come in the form of restrictions on the temperature, concentration of chemicals, and flowrate of streams exiting a plant. State and federal laws, for instance, may dictate the exit temperature of a cooling water stream into a lake in order to prevent harm to aquatic wildlife.

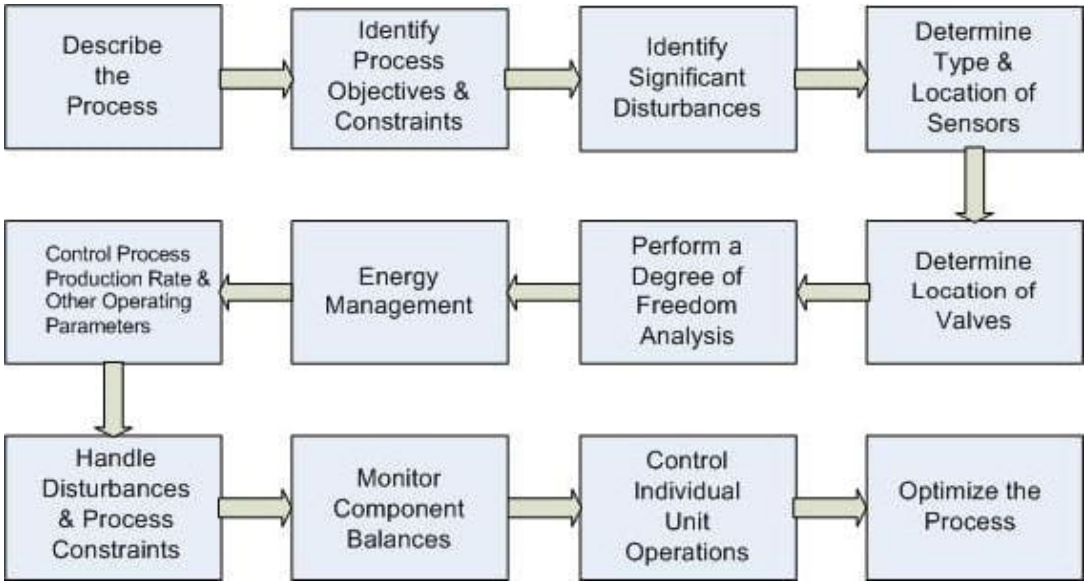
Operational Constraints

Equipment found in the plant may have their own unique limitations, such as temperature or pressure that require proper control and monitoring. For instance, a thermocouple may be damaged at extremely high temperatures, thus the location of the thermocouple must be accounted for.

Economics

In general, a company will operate so that its profits are maximized. The process conditions that maximize these profits are determined by way of optimization. Many costs must be considered when optimizing process conditions. Some of these costs are fixed, or will not change with process variables (i.e. equipment costs) and others are variable, or do depend on process variables (i.e. energy costs). The overall process is usually limited by certain factors including availability of raw materials and market demand for the final product. Therefore, the economics of a process must be well understood before process changes are enforced.

Step-by-Step Method For Describing Controls and Their Purpose



1.) Describe the Process

A brief description of the general process is needed while not dwelling on the details and calculations involved. The major steps of the process, as well as inputs and outputs of the process, should be stated. A simple diagram should be provided detailing the chemical process to help visualize the process.

2.) Identify Process Objectives and Constraints

The objectives and constraints of the process must be identified before process control actions can be performed.

The process objectives include the type, quantity, and quality of the product that is to be produced from the process. The economic objectives, such as the desired levels of raw material usage, costs of energy, costs of reactants, and price of products, should also be identified.

The process constraints include three different categories: operational, safety, and environmental limitations. *Operational constraints* refer to the limits of the equipment used in the process. For instance, a liquid storage tank can only hold a certain volume. *Safety constraints* describe the limits when the people or the equipment may be in danger. An example would be a pressure limitation on a reactor, which if exceeded, could result in an explosion. *Environmental constraints* limit how the process can affect the immediate surroundings. For example the amount of harmful chemicals that can be released before damage is done to nearby water supplies. All of these constraints should be mentioned to build a robust control system.

Careful reading of the information provided to you by the customer, management, and government is required in order to properly identify each constraint and objective. Often times, the process objectives will be very clearly laid out by the needs of the customer or management. Operational constraints, or the limitations of the equipment being used, must be researched for each piece of equipment used in the process. Generally, by satisfying the operational constraints a good portion of safety constraints are satisfied as well, but additional safety constraints may exist and must be investigated by researching company policy and governmental regulations. Environmental regulations also have to be researched through resources such as the EPA and Clean

Air Act. Satisfying the economic aspect is largely determined by manipulating additional variables after all other constraints and objectives have been met.

3.) Identify Significant Disturbances

Disturbances, in the sense of process description, are defined as inputs or external conditions from the surrounding environment that have certain properties that cannot be controlled by the plant personnel. Examples of disturbances include ambient air temperature, feed temperature, feed flow rate, feed composition, steam pressure changes, and cooling water temperature changes. Disturbances can drastically affect the operation of a unit. A control system should be able to effectively handle all process disturbances. As such, all possible disturbances must be identified and these disturbances need to be accounted for by the development of contingency plans within the process.

4.) Determine Type and Location of Sensors

A proper design must ensure that adequate measurements of the system are obtained to monitor the process. To meet this goal, sensors must be chosen to accurately, reliably, and promptly measure system parameters. Such parameters include temperature, flow rate, composition, and pressure. Placement of sensors is important both in the usefulness of measurements as well as the cost of the system. Sensors should be placed such that the measured quantities are appropriate in addressing control objectives.

5.) Determine the Location of Control Valves

Valves must be placed in a location to control variables that impact the control objectives. For example, control of the temperature of a reactor could be obtained by placing a valve on either the stream of heating / cooling fluids or by placing a valve on the feed stream to the reactor. One must determine which streams should be manipulated to meet process objectives.

6.) Perform a Degree of Freedom Analysis

The degrees of freedom in a system are equal to the number of manipulated streams (determined in step 5) minus the number of control objectives and control restraints (determined in step 2). A degree of freedom analysis is used to determine if a system is being under- or over-specified by the process objectives. The degrees of freedom come from the number of knowns and unknowns that are specified within the system. If there are extra degrees of freedom present in a system, unused manipulated variables can be used to optimize the process. If there are negative degrees of freedom, a system is over-specified because more objectives and restraints exist than manipulated streams. In this case, all objectives cannot necessarily be met simultaneously and the least important objectives must be neglected. A system with zero degrees of freedom is fully specified. All objectives can be met, but there is no room for optimization.

7.) Energy Management

In any system with exothermic or endothermic reactions, distillation columns, or heat exchangers, energy management becomes a factor that must be accounted for. Heat must be removed from exothermic reactions in order to prevent reactor runaway, and heat must be supplied to endothermic reactions to ensure desired production rates. Strategies such as pre-heating feed streams with the excess heat from a product stream are helpful in maintaining efficient usage of energy, however, they also result in more complex processes that may require more intricate control systems.

8.) Control Process Production Rate and Other Operating Parameters

The production rate can be controlled by a variety of manipulated variables. One manipulated variable may be the feed rate. The plant feed rate can be changed and each subsequent unit can use its controls to accommodate this change, ultimately resulting in a change in the final production rate. Other manipulated variables may also include reactor conditions, such as temperature and pressure. Temperature and pressure affect reaction rates and can be used to alter the final production rate. It is important to choose the most suitable manipulated variable to control production rate.

In addition to the production rate, other control objectives must be effectively managed by manipulated variables. For example, temperature of an exothermic reactor may be controlled by the flow of a coolant stream passing over it in order to avoid dangerous high temperatures. The pressure of a reactor may be controlled by the flow of feed gas in order to comply with the pressure limitations of the vessel.

9.) Handle Disturbances and Process Constraints

The effects of disturbances should be minimized as much as possible, in order to maintain the system at desired conditions

and meet all process objectives and constraints. Feedback or feedforward are specific control techniques and are common ways to overcome disturbances. A feedback control works by studying the downstream data and then altering the upstream process. The actions executed are reactive. Feedback can be viewed as an if-then statement: if a feed's temperature is detected to be lower than desired, then steam can be used to preheat the feed. Feedforward is a more proactive approach in that it adjusts a manipulated variable before the disturbance is felt in the process. Hence, if a sensor indicates low temperatures upstream of the feed, the feedforward control will counteract the effect of the cooler upstream temperatures by preheating the feed before the feed temperature is effected. Note that a disturbance must be detectable and measurable in order for the feedforward control to fix the anticipated disturbance before the system is effected.

Additionally, if constraints are reached during the process, controls should be implemented to avoid safety, operational, or environmental hazards. This can also be done with feedback and feedforward controls on manipulated variables.

10.) Monitor Component Balances

Every component within a process, whether it is inert or not, should be accounted for at every step of the system in order to prevent accumulation. This step is more crucial in processes that involve recycle streams. If such a stream is present, a purge stream is often necessary to remove unwanted components. In addition, component balances are used to monitor yield and conversion or reveal locations in the process where loss may be occurring. In order to monitor component balances, composition sensors are used.

11.) Control Individual Unit Operations

Most systems used today in industry employ the use of multiple unit operations. Each of these unit operations, however, needs to be fully controllable in the sense that it has a control system that can adjust manipulated variables in order to maintain other parameters. For instance, if an absorber is present, the system must be able to control the liquid solvent feed as some ratio to the gas feed. Another example is a crystallizer. The refrigeration load of the crystallizer must be controllable in order to control the temperature.

12.) Optimize the Process

In most cases, there will be certain aspects of a process that will not be dictated to a designer and can be changed to make the overall process more economical for the company. These are referred to as "unaccounted for" degrees of freedom and can be implemented as new control valves or adjustable controller setpoints.

Alternative Method of Verbal Modeling

This alternative method is also described in Peter Woolf's Recorded_Lectures Lecture 3. For purposes of this description, it is scaled down to a single unit process. However, this method can easily be applied to describe an entire system of unit processes.

1. Describe the process in words

Some of the important questions to answer before delving deeper into a model are:

- What are the components entering the system?
- How do they enter? Separately? Combined stream? What physical states are they in?
- What happens inside the unit process and what comes out at each exit point?

Remember to keep this part simple. There is no need to include chemical formulations or equations of any sort. Just lay out the basic flow of material.

2. Define the primary goal of the process

The primary goal should be simple. Often, it is to maintain a specific measured variable above a minimum or below a maximum. In this step, the only thing that needs to be determined is what the main goal is, and a few supporting details about why this is an important goal to achieve.

For example, a primary goal could be to minimize the concentration of Compound Y in orange juice because studies show Compound Y gives the juice a bad aftertaste.

3. Identify secondary processes that influence the primary goal

In a typical unit process, the primary goal will be directly influenced by one or two other aspects of the system. These can include temperature, pressure, inlet conditions, and more and can occur at various points in the process.

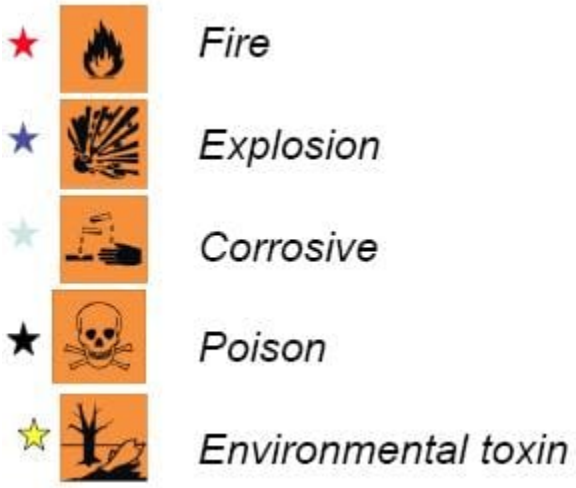
The goal of this step is to determine which of these other process variables will be most likely to influence the primary goal and to step down from there.

For example, the temperature of the orange juice mixer could have the greatest influence on production of Compound Y.

4. Identify safety and environmental risks

Next, you need to identify all of the points in the process that represent any type of risk. This will be important later in determining which system variables need to be monitored.

Step through your process and identify any points that pose a significant risk of the hazards shown in the following figure.



Examples include: Boilers represent fire and explosion risks. Any stream with a dangerous chemical can represent corrosive, poison, environmental, or all three risks.

5. Identify major costs associated with the process

How much something costs to produce is obviously a big deal in manufacturing. Identifying the largest sources of cost is critical in finding ways to reduce cost overall. Typical places to start identifying costs are at inlet streams (what is the cost of raw materials) and at any portion of the process where heat is added or removed.

It is important to include the high costs that can be associated with the risks identified in Step 4. Often the high cost of failure and risk exposure will determine what other seemingly costly steps must be taken to ensure the safety of the process.

6. Identify variables you can directly manipulate

The basics of the process have been laid out, and now it's important to determine what variables you can actually control. Typically, you only have direct control over the simplest of variables: switches and valves. Essentially, this just means that you cannot, in fact, choose a temperature for your system and implement it. What you can do, is control a valve or switch that activates heating or cooling to control the temperature.

During this step, you must decide where it is important to place these valves and/or switches. Use the information acquired previously about the primary goal and secondary effects to determine what variables are worth controlling. Remember that you don't need to put valves EVERYWHERE! Control valves are not costless and can also add unwanted complexity to your system. If you need to isolate equipment, you can install manual valves. Keep the control valves to the needed level.

7. Identify sources of variation

In order to write a control scheme, you need to know what values in your system will change and why. Some common causes of variation include:

- Environment: ambient temperature
- Other processes upstream or downstream: variable inlet conditions or outlet demand
- Economic forces: product worth, material costs
- Operators

Identifying what aspects of your process can be affected by these forces will allow you to assemble a more complete control scheme.

8. Describe your control system in words

Before you start trying to write everything out in computer code and mathematical equations, take the time to lay out your controls in words. This is much like preparing an outline before writing a paper. It can save you from many headaches later on.

One example of generic, simple syntax for verbal modeling is: Maintain [system variable] at specified level by adjusting [variable I can control].

The Barkel Method of Verbal Modeling

This method is an elaboration upon the steps outlined in Mr. Barry Barkel's lecture on September 29, 2009.

"*If you design a system, you have the ethical responsibility to control it.*"

1) Understand the Process

Before you can control anything, you have to understand the process and how different parts of the system interact with each other. Make sure that the overall process is understood. This includes inputs and outputs as well as major steps, however specifics are not necessary. You should also be able to construct a diagram to help explain the process.

2) Identify Operating Parameters

Operating parameters can include temperature, pressure, flow, level, etc. Choose the parameters to manipulate in your system that will safely result in the desired output.

3) Identify hazardous conditions

Consider all possible dangerous aspects of your process when designing your system. This could include a chemical overheating or a vessel overflowing. It is imperative that you ensure the safety of your operators and being aware of all hazardous conditions can aid in this.

4) Identify measurables

The main three measureables addressed in this class are: temperature, pressure, and flow. However, there are many more measureables, some more common than others. Here are some more: pH, humidity, level, concentration, viscosity, conductivity, turbidity, redox/potential, electrical behavior, and flammability.

5) Identify points of measurement

It is important to place sensors in locations so that efficient and reliable measurements are taken of the system to monitor the process. For example, in a distillation column, temperature sensors will display different temperatures at different locations down the tower. The sensors must be positioned so that accurate readings of the system are given. Also, it is necessary to place sensors in an area of constant phase.

6) Measurement methods (thermo couple? choose for range)

After identifying what is to be measured and where it will be measured, you have to decide how it will be measured. For example, a thermocouple can be used to measure the temperature. When choosing the equipment, be sure to check that the conditions of

use fall within the recommended range of operation for the equipment.

7) Select control methods

Decide whether to use feedback, feed-forward, cascade, or other types of control methods.

8) Select control system

A control system is a set of devices that will manipulate the actions of other devices in the system. A control system ranges from having an operator manually open and close a valve to running a system with feedback such as with PID controllers. Control systems can vary from relatively cheap to expensive. When picking a system, it would be most economic to choose the cheapest one that gets the job done. An example of a control system is a PIC, or programmable interface controller.

9) Select control limits

When choosing setpoints for controllers, you will also have to decide on a range that the values are allowed to fluctuate between before a corrective change is made. When selecting these limits, keep in mind that "equal" and "zero" do not exist due to the infinite number of decimals that electronics are now able to handle. So you must further define what "equal" means, when is it "close enough" to count as "equal" And when is the number small enough to count as "zero", is 0.1 or 0.01 or 0.0000000001 count as "zero"?

10) Define control logic

As every process is different, a customized code for each process must be written to tell the system what to do. For example, when a level control is a tank has reached a critically high point, the logic should spell out the necessary changes needed to bring the tank level back down. For example, this could be partially closing a valve upstream of the tank or partially opening a valve downstream of the tank.

11) Create redundancy system

In the real world, you must balance cost and efficiency/safety. On one hand, you don't want an out-of-control system if one control fails. But on the other hand, you can't afford to order two of everything. The critical point to keep in mind is to optimize the safety while minimizing the cost.

12) Define "fail-safe"

A fail safe is a set up in the control logic to ensure that in the event of a failure of a control method, the system will automatically reach a safe condition so that there is little to no harm done to other equipment or personnel.

13) Set lead/lag criteria

Valves and other equipment do not necessarily open/close or turn on/off at the exact instant a button is pressed or the control logic kicks in. There is often lag time associated with each controller. You must determine how long this lag time is so that you can take it into account.

14) Investigate effects of change before/after

Be sure to investigate effects of changing each controller. For example, what are the effects of closing/opening this valve?

15) Integrate all systems

Ensure that all systems are working together and that there are no holes in the system. Make sure that information does not fall through any cracks in the system.

[Note - we can use an example of the Barkel method - RZ]

Common Errors

1. Impossible direct manipulations e.g., Change the concentration of salt in a tank

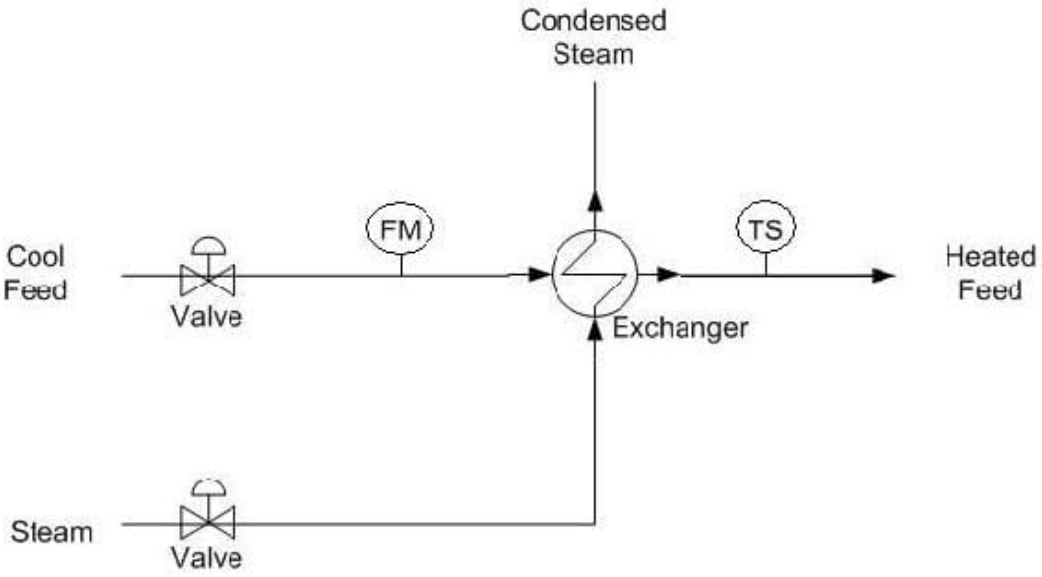
2. Missing the forest for the trees e.g., Sacrificing product quality for tight level control on a tank

3. Excessive or insufficient control e.g., Control every variable because you can or ignore the possibility of significant disturbances

Worked out Example 1

PLEASE NOTE: ALL VALUES IN THE FOLLOWING PROBLEMS ARE FICTICIOUS BUT MEANT TO BE LOGICAL

A heat exchanger uses steam to heat a stream of water from 50°F to 80°F. The water enters at a flow rate of 20 gallons per minute from a nearby lake. The process costs $65 per hour and yields a profit of $2 per gallon of product. The steam is provided by the plant and is 1000°F, which is a temperature that the pipes can sustain. For safety reasons, the exchanger may only run for 12 consecutive hours and requires 4 hours to cool down. Using more than 10,000 gallons of water per hour would cause an environmental disturbance to the water source. The diagram is shown below. Verbally model this system.



== Solution ==

1.) Describe the Process

The purpose of the process is to heat an incoming stream of water from a temperature of 50°F to a temperature of 80°F. The main equipment involved is a shell-and-tube heat exchanger.

2.) Identify Process Objectives and Constraints

The product specification of the process is water at a flow of 20 gallons per minute and a temperature of 80°F.

Economically, the process costs $65 per hour to operate. There are no costs for the raw materials, as the only inputs to the system are water and steam. The finished product produces a profit of $2 per gallon. The economic objective is to reduce process costs while producing sufficient product.

The operational constraints and safety concerns are due to the pipes. The pipes can only sustain a temperature of 1000°F. Safety is a concern because attempting to heat the incoming water to a certain temperature may cause the heat exchanger to malfunction, leading to equipment damage and possible burn injuries to nearby personnel. The system may only operate for 12 consecutive hours, after which the system will need to be cooled down for 4 hours to avoid the aforementioned hazards. A simplified assumption is that there are no constraints on steam because it is provided by the plant and causes no safety issues. The only environmental constraints involve the incoming water stream. The incoming water is gathered from the nearby lake, and a stream of greater than 10000 gallons per hour would cause a disturbance in the equilibrium of the lake.

3.) Identify Significant Disturbances

Significant disturbances can be found in the ambient air temperature, variable flow rates of the feed, and the temperature of the steam.

4.) Determine the Type and Location of Sensors

A flow sensor (FM) is placed at the incoming water stream. A temperature sensor (TS) is located on the product water stream. A flow sensor is not needed for the steam stream for this problem because this value is not needed for control. A sensor could be placed here but the information is not needed for this problem.

5.) Determine the Location of Control Valves

A flow valve is placed at the entrance of the incoming water stream. A flow valve is placed at the entrance of the steam.

6.) Perform a Degree-of-Freedom Analysis

There are two manipulated variables: the flow of the water feed stream and the flow of the incoming steam. There are two control objectives: the flow of the feed stream, monitored by the flow sensor, and the temperature of the product, monitored by the temperature sensor. Therefore the system has zero degrees of freedom.

7.) Energy Management

The incoming steam is used to transfer heat to the cool water feed. The temperature sensor on the product stream determines the applicable setting on the steam flow valve.

8.) Control Process Production Rate and Other Operating Parameters

The process production rate is controlled by the flow valve on the entering water stream. The water temperature is controlled by the flow valve on the incoming steam.

9.) Handle Disturbances and Process Constraints

Changes in the ambient air temperature can be detected by the temperature sensor, and can be corrected by the flow valve on the incoming steam stream. Variable flow rates of the water feed stream can be detected by the flow sensor and compensated by adjustments on the flow valve on the water feed stream. Changes in the temperature of the steam can be detected by the temperature sensor. The flow valve on the steam stream can be adjusted to increase or decrease the flow of steam and subsequently the amount of heat exchanged.

10.) Monitor Component Balances

A vent is located on the heat exchanger to release excess steam from the system. Aside from that, any accumulation is unlikely and can be neglected.

11.) Control Individual Unit Operations

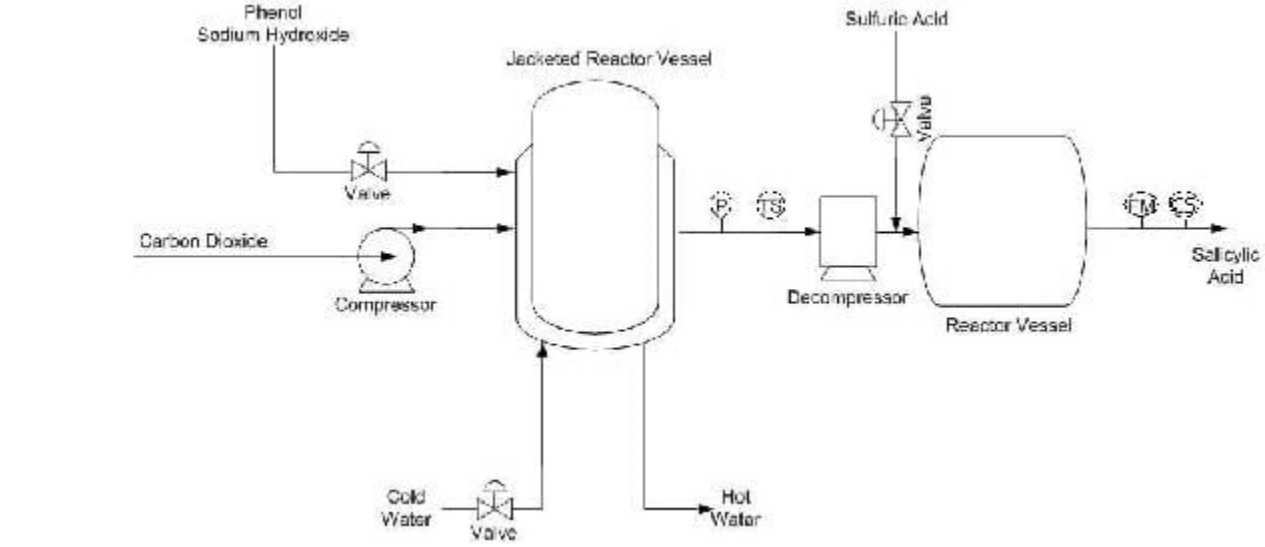
The outlet temperature of the product stream is controlled by the flow valve on the steam feed stream. The flow of the incoming water stream is controlled by the flow valve on the water feed stream.

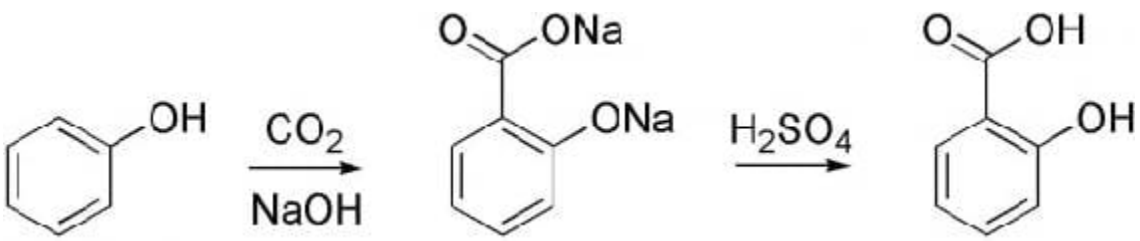
12.) Optimize the Process

One might notice that the process is only using 1,200 gal/hr of water, well below the 10,000 gal/hr environmental constraint. If the profit of the process is linear with the flow-rate of water, then increasing the flow-rate of water will increase the profits for the company. (With the constraints specified, this is a Linear Programming optimization problem. The optimal setpoint falls on a boundary condition.) However, the flow-rate of water entering the system is already specified, which results in zero degrees of freedom. (Zero degrees of freedom implies there are no further control valves or setpoints.) Further investigation should be conducted to determine the reason for the flow-rate specification. When considering increasing the flow-rate of water into the system, one should also check that the other constraints are not violated.

Worked out Example 2

A process converts phenol into salicylic acid through a series of two reactors. Phenol and NaOH are fed in the liquid phase into the first reactor where it reacts with gaseous carbon dioxide that is pumped in. Assume constant fresh feed temperature and that the feed flow rate is within operational constraints. Management has dictated that salicylic acid production must be 200 moles per hour. Also, management would like the product stream to have a molar composition of 80% salicylic acid. Due to environmental concerns, a maximum flow rate of 10000 gallons per hour of cold water can be used to cool the first reaction chamber. The valve controlling the flow of cold water does not allow a flow rate in excess of 7500 gallons of water per hour. The salicylic acid product is used in another process to produce aspirin, which has a market value of $10 per mole. The first reactor can be operated at pressures up to 200 atm, while the second can be operated at pressures up to 10 atm. The first reaction is exothermic, while the second reaction is assumed to generate negligible heat. A diagram of this process is shown below, as well as the reaction scheme. Verbally model this system.





== Solution ==

1.) Describe the Process

The purpose of the process is to convert phenol into salicylic acid, the precursor for aspirin. First, the phenol is reacted with gaseous carbon dioxide (CO2) and sodium hydroxide (NaOH) under high pressure and temperature in the first reactor. The product of this reaction is then combined with sulfuric acid to form salicylic acid in the second reactor. The reaction scheme is shown above.

2.) Identify Process Objectives and Constraints

The process is expected to produce 200 moles per hour of salicylic acid. The product stream must contain at least 80% by moles of salicylic acid. The equipment used in the process dictates the operational limitations. The first reactor vessel can be operated up to a pressure of 200 atm, while the second reactor vessel has a 10 atm upward pressure limit. As such, pressures in excess of these limits must be avoided. Since the first reactor will generate a significant amount of heat, the heat must be removed to avoid damage to equipment and possible runaway reactions. Therefore, a heat exchanger (in the form of a reactor jacket in this case) with cool water should be included to decrease the temperature of the reactor. Economic concerns demand that phenol, sodium

hydroxide, and sulfuric acid should not be used in extreme excess. The costs of these materials and the energy costs required to process them affect the overall profitability, so these compounds should not be wasted. Environmental regulations limit the use of water to cool the reactor at 10000 gallons per hour, however the valve constraints limits the amount of water to only 7500 gallons per hour.

3.) Identify Significant Disturbances

The amount of cold water available to cool the reactor can be considered a disturbance because it comes from a reservoir outside of our control. The ambient temperature is also a disturbance. If it drastically increases, the amount of cold water needed to cool the reactor would need to increase as well. Composition of the feed streams will be assumed to be constant in this example. Therefore, they are not considered disturbances.

4.) Determine the Type and Location of Sensors

A temperature sensor (TS) and pressure sensor (P) are located on the stream exiting the first reactor vessel. A flow meter (FM) is located on the product stream leaving the second reactor. A composition sensor (CS) will also be located on the product stream leaving the second reactor. The pressure drop can be controlled through the decompressor and thus is a control.

5.) Determine the Location of Control Valves

Control valves are located on the feed stream containing the phenol and sodium hydroxide, the incoming cold water to the first heat exchanger, and the sulfuric acid feed stream. There is also a pump located on the carbon dioxide stream that enters the reactor.

6.) Perform a Degree of Freedom Analysis

There are 3 valves, 1 pump, and 1 decompressor but 5 objectives. This results in zero degrees of freedom. The valve located on the sulfuric acid feed stream is meant to meet the composition constraint placed on the product stream leaving the second reactor. The valve located on the feed stream carrying the reactants is set to satisfy production requirements. The valve on the cold water stream is used to maintain reactor temperature, which satisfies an operational constraint. The pump is to ensure the correct pressure is achieved in the reactor, also satisfying an operational constraint. The decompresser is to maintain a pressure of less than 10 atm in the second reactor, thus satisfying another operational constraint.

7.) Energy Management

The heat from the exothermic reaction in the first reactor is transferred to the cold water stream. The hot water stream exiting the reactor vessel jacket could be used to heat streams on other processes. The second reactor is assumed to generate negligible heat during the reaction, thus any release of heat from the reactor will be considered safe to release into the environment surrounding the process.

8.) Control Process Production Rate and Other Operating Parameters

The production rate is measured by the flow sensor on the product stream and this signals the control valve on the feed stream through a feedback mechanism to change the production rate as necessary.

9.) Handle Disturbances and Process Constraints

If the temperature sensor on the reactor exit stream exceeds a certain level due to a diminished cold water supply, the feed stream valve would decrease the amount of reactants entering the reactor. The amount of feed would also be decreased if more than 7500 gallons per hour of cooling water were needed, as this is an operational constraint. If the pressure gauge controlling the pump begins to read higher than allowed pressures, the pump would decrease the flow of the carbon dioxide entering the reactor. Also, if the pressure gauge reads out a pressure that will be too high for the second reactor, the decompresser will be allowed to disperse more pressure. If ambient air temperature drastically increases, the temperature sensor would open the cold water valve allowing more cooling water to enter the reactor vessel jacket. If the composition of the product stream falls below 80 mole percent of salicylic acid, then the valve controlling the sulfuric acid feed would allow more sulfuric acid into the second reactor to increase the conversion of reactants.

10.) Monitor Component Balances

The composition sensor and flow meter on the product stream leaving the second reactor will account for every species to ensure

that there is no accumulation or loss within the system.

11.) Control Individual Unit Operations

The first reactor vessel's pressure is fully controlled by the pressure gauge and pump system and its temperature is fully controlled by the temperature sensor which controls the reactant feed valve and the cool water valve. The second reactor's pressure is fully controlled by the same pressure gauge and the decompresser system, and its temperature will be highly dependent on the amount of cooling water used to cool the product exiting the first reactor.

12.) Optimize the Process

Since there are no unaccounted degrees of freedom, there are no valves to adjust in order to optimize the process. It should be noted, however, that if there was no constraint on the composition of the product stream, the sulfuric acid feed valve would have become an unaccounted for degree of freedom. If this had been the case, the valve could be adjusted to maximize the profit of the process. In order to maximize the profit, the benefits of having higher conversion and more product would have to be weighed against the increase costs of using more sulfuric acid feed.

Sage's Corner

An example of verbal modeling
http://www.youtube.com/watch?v=YTs_IWccnvw&feature=player_embedded

Powerpoint Slides from the example

References

Luyben, William L., Tyreus, Bjorn D., and Luyben, Michael L. “Chapter 8: Eastman Process” in Plantwide Process Control, McGraw-Hill, New York, pp. 251-272.

Luyben, Michael L., Tyreus, Bjorn D., and Luyben, William L., "Plantwide Control Design Procedure" in AIChE Journal Dec. 1997, Vol. 43, No. 12 pp. 3161-3174.

Riggs, James B. and Karim, M. Nazmul. “Chapter 17: Multiunit Controller Design” in Chemical and Bio-Process Control, Ferret Publishing, pp. 491-504.

Stephanopoulos, George. Chemical Process Control: An Introduction to Theory and Practice, Prentice Hall, New Jersey, pp. 1-41.

This page titled 2.1: Verbal Modeling- process description, control specifications, and connections is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

2.2: Degrees of Freedom- importance, calculation procedure, and examples

Authors: (13 December 2009) Jon Butler, Matthew J LaVelle

Introduction

In control engineering, a degree of freedom analysis is necessary to determine the regulatable variables within the chemical process. These variables include descriptions of state such as pressure or temperature as well as compositions and flow rates of streams.

Definition

The number of process variables over which the operator or designer may exert control. Specifically, control degrees of freedom include:

1. The number of process variables that may be manipulated once design specifications are set
2. The number of said manipulated variables used in control loops
3. The number of single-input, single-output control loops
4. The number of regulated variables contained in control loops

The following procedure identifies potential variables for manipulation.

The Process

The method we will discuss is the Kwauk method, developed by Kwauk and refined by Smith. The general equation follows:

Degrees of freedom = unknowns - equations

Unknowns are associated with mass or energy streams and include pressure, temperature, or composition. If a unit had **Ni** inlet streams, **No** outlets, and **C** components, then for design degrees of freedom, C+2 unknowns can be associated with each stream. This means that the designer would be manipulating the temperature, pressure, and stream composition.

This sums to an equation of

Total Unknowns = Ni*(C+2) + No*(C+2)

If the process involves an energy stream there is one unknown associated with it, which is added to this value.

Equations may be of several different types, including mass or energy balances and equations of state such as the Ideal Gas Law.

After **Degrees of Freedom** are determined, the operator assigns controls. Carrying out a DOF analysis allows planning and understanding of the chemical process and is useful in systems design.

Applications

Single phase systems

- All outlet streams have the same composition, and can be assumed to have the same temperature and pressure

Multiple phase systems

- An additional (C-1) composition variable exists for each phase

Complete Process

- When connecting units which share streams, one degree of freedom is lost from the total of the individual units

Example 1: Blender

This example (from the ECOSSE Control Course web page listed below) investigates degrees of freedom in a simple vapor mixing unit. Two gaseous streams enter a vessel and exit as a single well-mixed stream (Figure 1). We will apply the above equation to determine degrees of freedom.

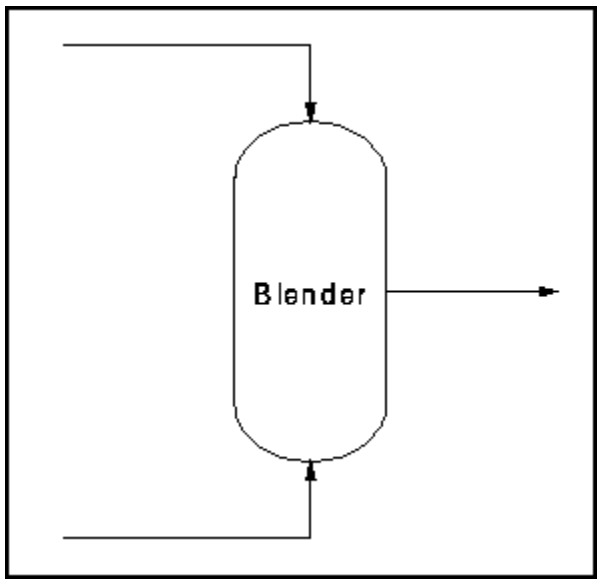


Figure 1: Blender Schematic

Here, there are 3 streams, each with C+2 unknowns for a total of **3C+6 Unknowns**.

We have **C** mass balances and **1** energy balance for a total of **2C+1** equations. We also know composition, pressure, and temperature for the incoming streams. This adds **2C+2** to the equation. Putting everything together gives:

Degrees of freedom = 3C+6 - (2C+1 + 2C+2)

Hence, the system has **3** degrees of freedom. Therefore, we can fix outlet composition, pressure, and flow rate. Figure 2 shows an example control scheme:

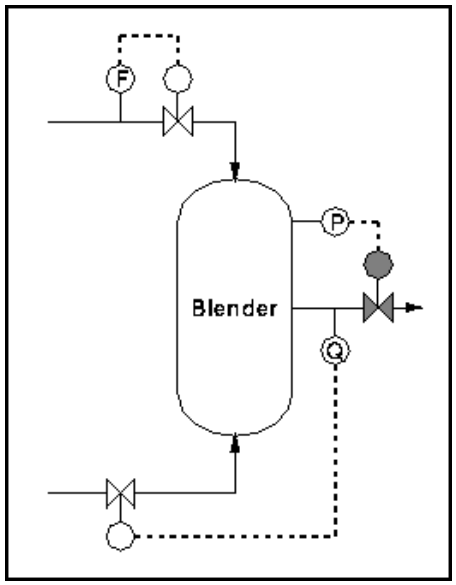


Figure 2: Blender with controls

References

Ponton JW, 1994, Degrees of Freedom Analysis in Process Control, Chemical Engineering Science, Vol. 49, No. 13, pp 1089 - 1095.

eweb.chemeng.ed.ac.uk/courses/control/restricted/course/third/course/formal.html

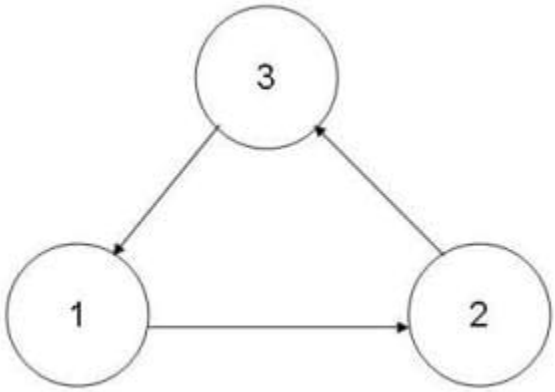
This page titled 2.2: Degrees of Freedom- importance, calculation procedure, and examples is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

2.3: Incidence Graphs -Interpretations, Consistency, and Inconsistency

3.1 Introduction

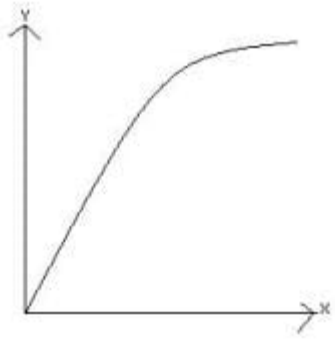
Incidence graphs are useful tools for chemical engineers to model various relationships in a process. They are used to systematically map an entire chain of processes and controllers to describe the effect, which each element has on the others. This helps to visualize the possible process pathways or a chain of effects.

Below is an example of an incidence graph. The circles are nodes which represent a particular device or controller, e.g. a temperature sensor. The arrows indicate the directional pathway of influence. For instance, a temperature sensor node would be connected to an adjacent node (such as a heat exchanger). This heat exchanger node would in turn be connected to other nodes representing devices or controllers. This type of diagram could be extremely useful in identifying the redundancies within a control system. A more elaborate application of incidence graphs will be discussed in the Worked Out Examples section.

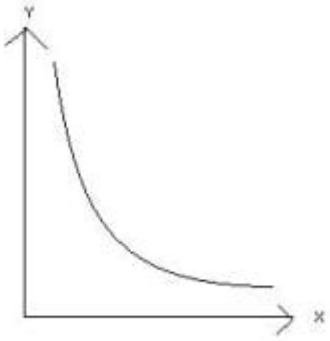


Monotonicity

Before proceeding to the next few sections, it is imperative to understand the basics of monotone systems. A monotonic function defines a function that preserves a given order. Monotonic functions are classified as monotonic increasing or monotonic decreasing. Taking the example of a function f(x), for a monotonic increasing function an increase in variable ‘x’ would result in an increase in f(x). On the other hand, for a monotonic decreasing function an increase in variable ‘x’ would result in a decrease in f(x).



Monotonic Increase : If x≥y, then f(x)≥f(y)

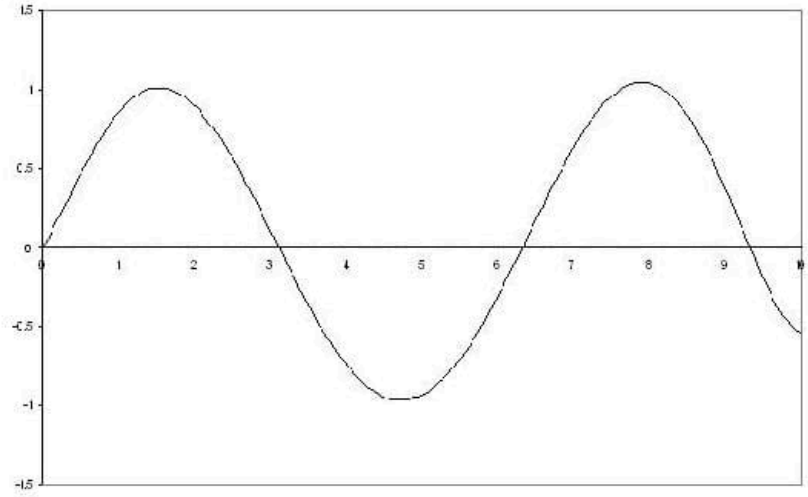


Monotonic Decrease : If x≥y, then f(x)≤f(y)

As long as a function is either always monotone increasing or monotone decreasing, it is referred to as a monotone function.

Monotonic functions are ideal systems to control because changes in one variable/device lead to known changes in other variables/devices. A non-monotonic function on the other hand, does not yield a straightforward output for a particular input. In these cases IF-THEN-ELSE statements are used to illustrate the controls for the nodes. For instances, IF node 1 is activated, THEN node 2 is repressed and so on.

An example of a possible non-monotonic graph



Incidence Graphs

Incidence graphs (also “Levi” or “Causation” graphs) consist of monotonic connections between nodes, whereby each connection indicates a particular effect of one factor on another. The term “monotonic” means that a change in the initiating device influences only one aspect of the target device. This simplification facilitates describing how control systems work through the use of incidence diagrams.

In reality, not all interactions are monotonic. Some control devices can simultaneously affect several aspects of other devices or variables. This may result in a non-monotonic relationship, in which one factor increases another over one region, but diminishes it over another (imagine a parabola, which is a non-monotonic curve). Temperature often has such a relationship with the rate of a reaction: for an exothermic reversible reaction, increasing the temperature will increase yield up to a certain point; however, further increase will diminish the yield.

Incidence graphs can only be used to represent monotonic relationships. Therefore, over simplification of a system can cause problems in constructing an incidence graph. Similarly a non-monotonic relationship can be resolved by breaking down a factor into separate components, each of which has only monotonic relationships with other factors.

When two factors are connected by multiple signal paths, incoherence between the results of different paths leads to *inconsistency*. In other words, the outcome cannot be predicted because the incidence graph does not communicate quantitative information. For instance, in the reaction rate example, reaction rate could be separated into the variables which affect it: concentrations, rate constant, and equilibrium constant. An incidence graph can then be constructed to demonstrate that temperature and reaction rate have an *inconsistent* relationship.

A diagram is consistent (to be discussed later in detail) if all signal paths between two nodes, for every pair of nodes, are in agreement with one another (i.e. they have the same result).

Traversing

Every incidence graph establishes a particular way to depict the direct effects of one device on its neighbors. Typically, a line will connect a node to its neighbors--these lines are terminated by either a bar or an arrowhead.

An arrowhead at the end indicates that the initiator (from which the line is coming) is either increasing the value of the target (where the arrow is pointing) or simply activating it (turning it on). The context of the connection will determine whether something is being increased or turned on. For example, it would be illogical to assume that a drop in temperature would cause a temperature controller to *deactivate* a pressure controller. Instead, it would be more prudent to assume that the temperature controller would *adjust* the pressure sensor.

arrowhead bar

A perpendicular bar indicates that the initiating device is either decreasing the value of the target device or inhibiting it (turn off). Note that these indications are not necessarily strictly followed. Once again, context of the situation determines which interpretation should be used.

perpendicular bar

Graphs consisted of the arrowhead bar and the perpendicular bar allow you to see the result of an increase/decrease in one aspect of the initial device on a designated device in the system. Everything done after the initial device comes as a result of changing the initial device. In addition, these graphs provide the possibility of simplifying complicated relationships between devices into one conceptual unit. The relationship between the initiating device and target nodes can be described by the following set --> {initiator,target1,target2,etc}. It is read as "the increase or decrease in the initiating device affects the target nodes in the following manner... (+ for increase/activate and - for decrease/inhibit). The initiator can be any device or variable in the system as long as the arrows support the designation. In other words, you can show the effects of any device on any other device, as long as the pathway is supported by the diagram. There are often multiple routes that a signal can travel.

2 helpful tips to keep in mind as you interpret an incidence graph:

1. Always travel from **tail to head** (or source to target) when traversing the graph from the initiating device to the designated device. Keep in mind that it is possible to have arrowhead and/or perpendicular bars pointing in both directions in between two devices.

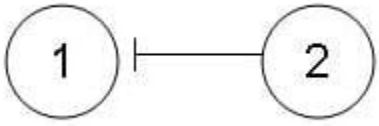
2. Arrowhead bars don't always mean increase as perpendicular bars don't always mean decrease. Arrowhead bars indicate that the effect of performing an action to the device at the tail of the bar would yield the **same** effect to the device at the head of the bar. Penpendicular bars indicate that the effect of performing an action to the device at the tail of the bar would yield the **reverse** effect to the device at the head of the bar.

The following is a simple example of an incidence graph. We will refer to the circles below as "nodes" from this point onward. Nodes are the selected devices or variables in a control system.

This incidence graph indicates that:

a. A change in an aspect of 1 will change some aspect of device 2 in the same direction -- otherwise written as **{+,+}**

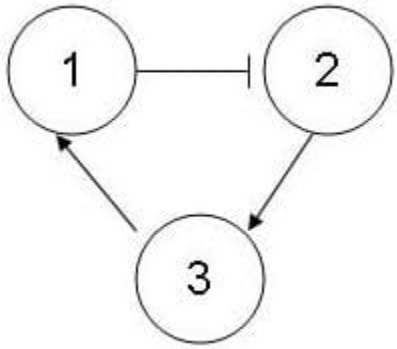
The ends indicate the direction of influence as well as the type of influence.



This incidence graph indicates that:

a. increase in an aspect of 2 (initiator) decreases some aspect of 1 (target)
b. indicated by {+,-}

This next example builds on the previous two...



This incidence graph indicates that:

Given an arbitrary 3,1,2 order, an increase in 3 (initiator) will make 1 (target) go up. 1's increase will then make 2 decrease. The symbolic representation for this would be {+,+,-}. One can see that this diagram shows how node 3 influences everything directly and indirectly connected to it.

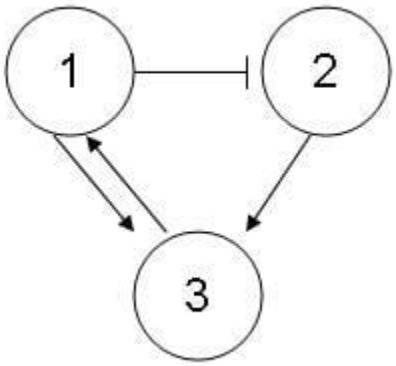
Now, given a 1,2,3 order, node one becomes the initiator. The symbolic representation then changes to {+,-,-}. **NOTE: the last node is negatively affected because 2's impact on 3 is shown to increase when 2 increases. But since 1 decreases 2, then the opposite is true!**

The main objective of these diagrams is to determine the effect that a change in the starting node has on another node in the system. However, in some cases, you can follow a pathway that ends at the starting node in order to check for ambiguity. For example, the 1,2,3,1 order yields a result of {+,-,-,-}. Since increasing node 1 cannot cause a decrease in itself, it is ambiguous.

Incidence graphs for complex controls systems will usually indicate that there are several ways for a device to influence another device in the system.

Multiple paths

Remember these are done with respect to increases in a node's value. So, if a path leads you to a point where you decrease a node, then you have to do the opposite of what the diagram is indicating for all steps after that (for that route).



The incidence graph indicates:

1. For an order {1,2,3}: you observe {+,-,-}... remember not to count beginning node twice...

2. For an order {1,3}: you observe {+,+}

Consistent, Inconsistent, and Partially Consistent

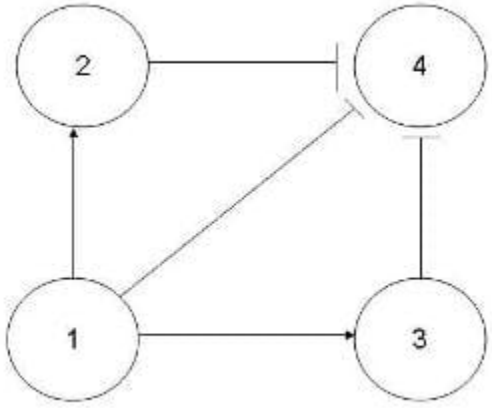
When designing controls for systems it is always important to keep in mind the need for redundancy and feedback control. Redundancy in a system will ensure that a system is monitored by multiple controllers. So, if one controller was to fail there would be another to ensure the same desired outcome. Consistent graphs can be used to describe redundancy. Feedback control is also very important and it is not good to have the system overreact to external stimuli. A controlled response is desired. Inconsistent graphs can be used to describe feedback control. Additionally, systems exist that both need to be redundant and feedback controlled and these can be described with partially consistent graphs.

Consistent Graphs

Consistent pathways happen when all pathways from the same starting node, that lead to the same target node, give the same output.

For example,

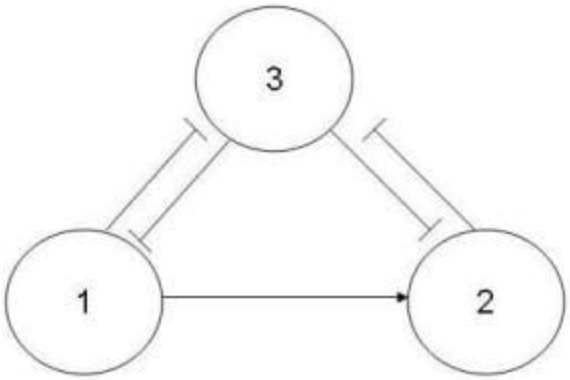
All possible pathways from node 1 to target node 4 yield the same output. These are called consistent pathways.



Paths(s)	Sign series
1,4	(+,-)
1,2,4	(+,+,-)
1,3,4	(+,+,-)

From the table and graph above, all pathways from node 1 to node 4 yield the same outcome. Node 4 is being decreased.

A consistent graph is made up of only consistent pathways for all possible pathways between any combination of two nodes. The graph below is an example of a consistent graph.



Paths(s)	Sign series
1,2,3	(+,+,-)
1,3	(+,-)
3,1,2	(+,-,-)
3,2	(+,-)
1,3,1	(+,-,+)
1,2,3,1	(+,+,-,+)
3,2,3	(+,-,+)
3,1,3	(+,-,+)
3,1,2,3	(+,-,-,+)

All pathways leading from node 1 to node 3 -

1,2,3 : (+,+,-)
1,3 : (+,-)

Since the two pathways cause the same change to node 3, this sub-pathway is consistent.
All pathways leading from node 3 and to node 2 -

3,1,2 : (+,-,-)
3,2 : (+,-)

Since the two pathways cause same change to node 2, this sub-pathway is consistent.

All pathways leading from node 1 and back to node 1 -

1,3,1 : (+,-,+)
1,2,3,1: (+,+,-,+)

Since the two pathways result in the same change to node 1, this sub-pathway is consistent.

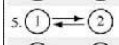
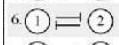
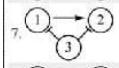
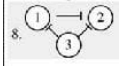
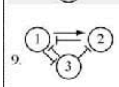
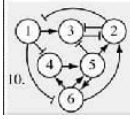
All pathways leading from node 3 and back to node 3 -

3,2,3 : (+,-,+)
3,1,3 : (+,-,+)
3,1,2,3: (+,-,-,+)

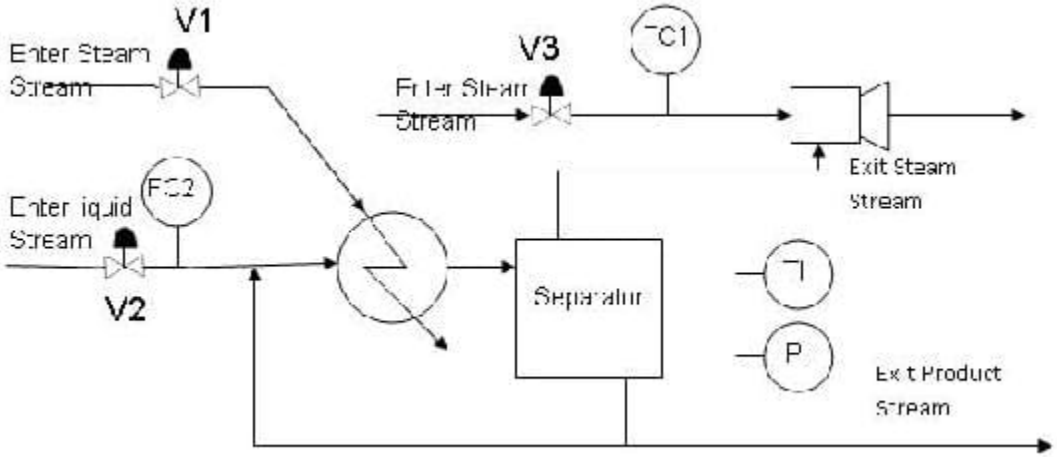
Since the three pathways give the same effects to node 3, this sub-pathway is consistent.

Because all sub-pathways of this incidence graph are consistent, the entire process is consistent.

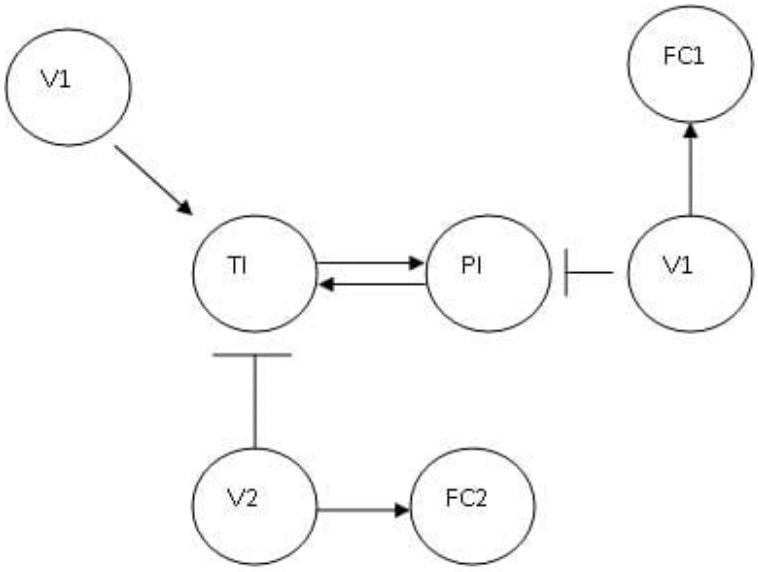
The table shown below gives examples of many different incidence diagrams and shows whether or not they are consistent.

Number	Graph	Consistent?	Paths(s)	Sign series	Interpretation
1.		Yes	{1}	+	1 influences itself
2.		Yes	{1,2}	+,+	1+ influences 2
3.		Yes	{1,2}	+,-	1 - influences 2
4.		No	{1,2,1}	+,+,-	influence of 1 on 1 via 2 is ambiguous
5.		Yes	{1,2,1}	+,+,+	1 + influences 2, which + influences 1
6.		Yes	{1,2,1}	+,-,+	1 - influences 2, which - influences 1
7.		Yes	{3,1,2} {3,2}	+,- +,-	both paths from 3 to 2 show that 3 - influences 2
8.		No	{3,1,2} {3,2}	+,-,+ +,-	influence of 3 on 2 is ambiguous
9.		No	{3,1,2} {3,2} {1,3,1} {1,2,1} {3,2,1} {3,1}	+,-,- +,- +,-,+ +,+,- +,-,+ +,-	influence of 1 on 1 via 2 is ambiguous and influence of 3 on 1 is ambiguous
10.		Yes	{1,4,5,3} {1,3} {6,4,5,3,2} {6,5,2} (and many many more possible paths)	+,-,-,+ +,+ +,+,+,-,+ +,+,+	Each variable 1,2,3,4,5,6 net has a simple relationship, monotonic between any other variable in the set.

Another type of consistent graph is one in which the valves of the system are not dependent on one another for feedback. In other words, there is no proportion of the output signal from a particular valve of a system is passed to the input of another valve. An example of this would be to consider the following process diagram:



This process is of a distillation column in which the bottom columns collect high intensity liquids and the liquid is removed from a series of evaporators. The evaporator is a single-effect evaporator equipped with a simple heat exchanger, a separation vessel and a steam jet ejector to supply the vacuum. The product is then sent to an oven where it is dried and sent to the manufacturer. It is the chemical engineer's task to create a consistent incidence graph in which the valves in the diagram do not depend on each other's feedback. An example of this is below:



Paths(s)	Sign series
V1, TI, PI, TI	(+,+,+,+)
V2, TI, PI, TI	(+,-,-,-)
V3, PI, TI, PI	(+,-,-,-)

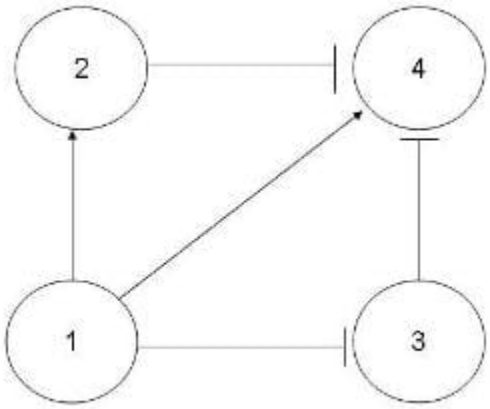
Based on the table above, the incidence graph is consistent because the valves are not dependent on one another for feedback.

Inconsistent Graphs

Inconsistent pathways happen when pathways from the same starting node, that lead to the same target node, give different output.

For example,

Pathways from node 1 to target node 4 do not yield the same output. These are called inconsistent pathways.

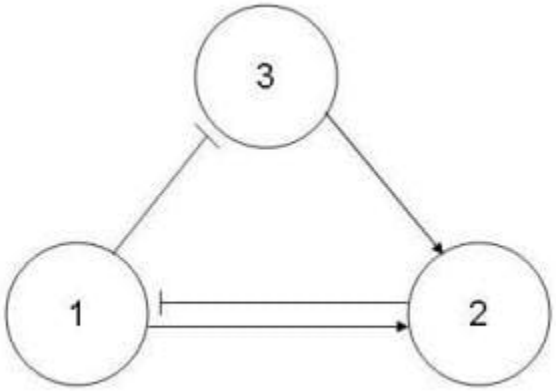


Paths(s)	Sign series
1,4	(+,+)

1,2,4	(+,+,-)
1,3,4	(+,-,+)

From the table and graph above, all pathways from node 1 to node 4 do not yield the same outcome. Node 4 is being decreased while simultaneously being increased.

An inconsistent graph is made up of only inconsistent pathways for all possible pathways between any combinations of two nodes. The graph below is an example of an inconsistent graph.



Paths(s)	**Sign series**
1,2	(+,+)
1,3,2	(+,-,-)
1,2,1	(+,+,-)
1,3,2,1	(+,-,-,+)

All pathways leading from node 1 to node 2 -

1,2 : (+,+)
1,3,2 : (+,-,-)

Since the two pathways cause different changes to node 2, this sub-pathway is inconsistent.

All pathways leading from node 1 and back to node 1 -

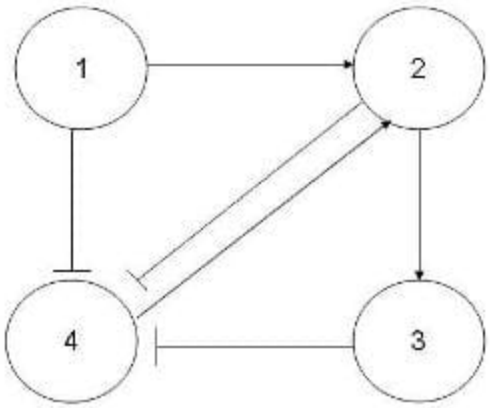
1,2,1 : (+,+,-)
1,3,2,1: (+,-,-,+)

Since the two pathways cause different changes to node 1, this sub-pathway is inconsistent.

Because all sub-pathways in this incidence graph are inconsistent, the incidence graph is inconsistent.

Partially Consistent Graphs

Partially consistent graphs are made up of both consistent and inconsistent pathways. If you increase node 1, all paths leading to node 4 cause a decrease in the node which means it is a consistent pathway. If node 1 is increased, one path leading to node 2 causes an increase and another path causes a decrease in the node. This is an inconsistent pathway. Since this model contains both at least one consistent and one inconsistent pathway, it is a partially consistent model.



Paths(s)	Sign series
1,2,3,4	(+,+,+,-)
1,2,4	(+,+,-)
1,4	(+,-)
1,2	(+,+)
1,4,2	(+,-,-)

All pathways leading from node 1 to node 4 -

1,2,3,4: (+,+,+,-)
1,2,4 : (+,+,-)
1,4 : (+,-)

Since the three pathways cause the same change to node 4, this sub-pathway is consistent.

All pathways leading from node 1 and to node 2 -

1,2 : (+,+)
1,4,2 : (+,-,-)

Since the two pathways cause different changes to node 2, this sub-pathway is inconsistent.

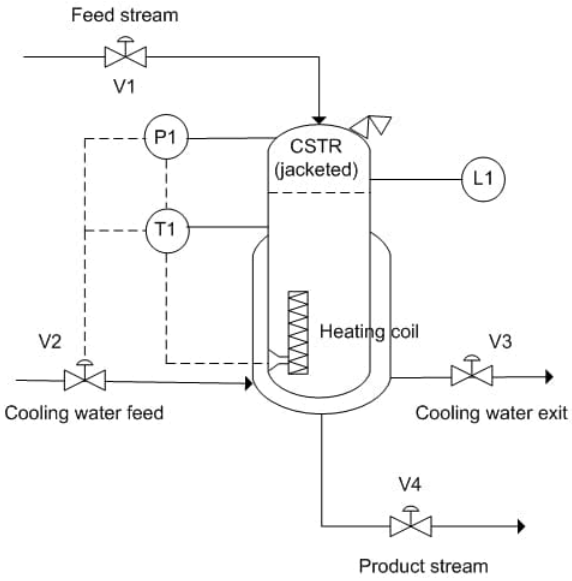
In this graph there are sub-pathways that are consistent and inconsistent, therefore the incidence graph is partially consistent.

Summary

Just because a process is inconsistent does not mean that the desired results cannot be achieved. All it means is that you cannot consistently achieve the desired results. Take the example of the flow of a final product. In a consistent process, no matter which path is taken, the flow of the final product will always respond the same, whether desirable or not. In an inconsistent process, changing the path will not always cause the flow of the final product to change the same way, but rather depends on the path taken.

In all, inconsistent processes are not bad. Often, processes will be very complex, containing many variables, and inconsistent processes will be unavoidable. Care must just be taken with complex, inconsistent processes to achieve desirable results.

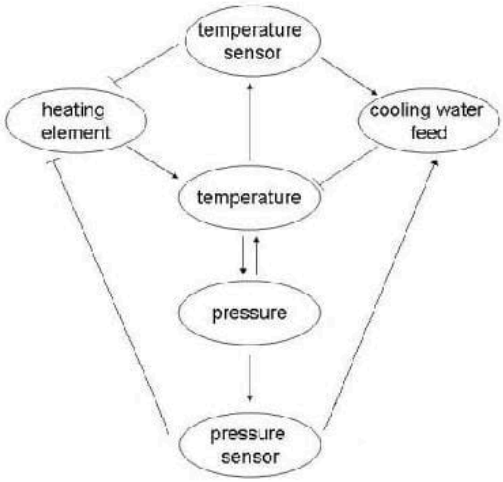
Worked out Example 1



You are a process engineer and you have been assigned to devise a process control for a new reactor. The optimum operating condition for the reaction inside the reactor is 220 ^{o}C ± 10°C. The reaction is exothermic and often goes way past the optimum temperature. It is imperative that the reaction should not go above 250°C. The reactor has a built in heater (H1) to keep temperature at its optimum temperature and a cooling water feed to cool the reactor in case temperature goes past 250°C. The reactor also has temperature sensor installed to monitor the temperature. If temperature goes above 250°C, the desired mechanism is to increase the flow of water in the cooling jacket surrounding the reactor and to turn off the heating element inside the reactor until it reaches its optimum temperature. As a control redundancy mechanism, a pressure sensor is also installed inside the reactor. The pressure control will send signal to the cold water feed if pressure reaches 1.2 atm which is the critical pressure. Draw a causation graph governing the situation when the temperature gets too hot inside the reactor. Base your graph on the following variables:

Note: (There may be more than one correct solution to the assigned problem)

a. temperature
b. temperature sensor (T1)
c. cooling water feed (V3)
d. heating element (H1)
e. pressure
f. pressure sensor (P1)



High temperature will 'activate' the temperature sensor, telling it that the temperature has exceeded the optimal temperature and an action need to be done to bring it back to the desired temperature. The temperature sensor would then open the valve for the cooling water feed which would cause the water to flow through the jacket. This results in the temperature inside the reactor to decrease. The reverse happens for the heating element. An activation of the temperature sensor would translate to an inhibition of the heating element in which the heating element would stop increasing the temperature of the reactor. Temperature and pressure are directly related by ideal gas law and thus affects each other in both ways. Thus, for the control redundancy mechanism, an increase in pressure would activate the pressure sensor which would then activate the cooling water feed to cool the temperature for the reactor.

Logic control:

- IF T1 =< 220 ^{o}C, THEN H1 is on, ELSE H1 is off
- IF T1 >= 250 ^{o}C, THEN V2 and V3 is open, ELSE V2 and V3 is close
- IF P1 >= 1.2 atm, THEN V2 and V3 is open, ELSE V2 and V3 is close

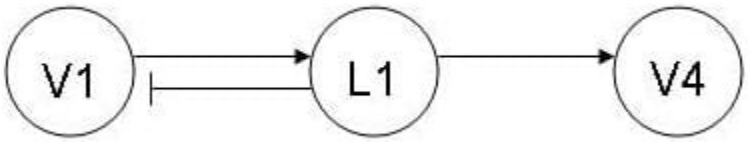
Worked out Example 2

After completing the assignment of devising the temperature control, your boss asked you to devise a mechanism to control the flow of raw materials and products. Draw a causation graph showing negative feedback relations between L1, V1 and V4 when the level inside the reactor is increasing (this is an example of an inconsistent pathway). The normal operating conditions are shown below:

V1 is 50 gal/min, V1 is usually open

V4 is 50 gal/min, V4 is usually open

L1 is 20 meters, L1 should not go above 30 meters and should not be below 10 meters



The level sensor monitors the fluid height inside the jacketed CSTR. Assuming the level inside the reactor is increasing steadily, the level sensor will send a negative feedback signal to V1 to decrease the input flow.

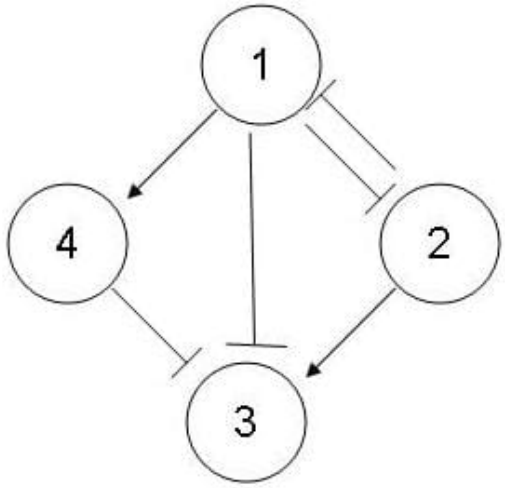
Logic control:

IF L1 > 30 meters, THEN V1 is closed

IF L1 < 10 meters, THEN V4 is closed

Worked out Example 3

Determine whether the following causation graph is consistent, inconsistent or partially consistent. Include your reasoning.



Solution:

All relationships are consistent. Each variable 1,2,3,4 has a simple relationship which is monotonic among the other variables in the set.

Paths(s)	Sign series
{1,2}	(+,-)
{1,2,3}	(+,-,-)
{1,3}	(+,-)
{1,4,3}	(+,+,-)
{1,4}	(+,+)
{2,1}	(+,-)
{2,3}	(+,+)
{2,1,3}	(+,-,+)
{2,1,4}	(+,-,-)
{4,3}	(+,-)

References

- Woolf, Peter (2006). A Mini-Guide to Consistent Graphs.
- Woolf, Peter (2006). Terms Review.

- Angeli, David & Sontag, Eduardo D. (2003). Monotone Control Systems.

Contributors and Attributions

- Authors: (September 26, 2006) Christopher Garcia, Anwar Stephens, Winardi Kusumaatmaja, Meng Yang Ng
- Stewards: (September 5, 2007) Alexander Voice, Andrew Wilkins, Ibrahim Oraiqat, Rohan Parambi

This page titled 2.3: Incidence Graphs -Interpretations, Consistency, and Inconsistency is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

2.4: Excel Modeling - logical models, optimization with solver for nonlinear regression, sampling random numbers

Introduction

Microsoft Excel program is one of the most popular and useful computer programs for a wide variety of numerical applications. Excel features several different functions, interfaces and graphing tools which can be applied in many fields. The following functions might be especially useful for logical programming in Excel:

- Logical functions (ex: IF, OR and AND): These functions can be used in control analysis, particularly in safety regulation (for example, if the temperature exceeds X degrees, shut down the reactor and send cooling water to the jacket)
- Solver: This function can be used to maximize, minimize or try to obtain an input value in a cell by varying referenced cells.
- Random Number Generator. This can be used as a tool for probability or randomly selecting data points for analysis from a large pool of points.

While it might be less accurate than other modeling programs, the Excel model provides good estimates, and facilitates observation and understanding of process behavior while requiring minimum computer knowledge. The following sections will provide you with a better understanding of these common Excel functions, followed by examples in engineering applications.

Logical Methods

A logical test is any expression that can be evaluated as TRUE or FALSE with specified input conditions. Commonly used logical tests are IF, OR and AND statements. All three statements can be used in conjunction with each other for more complex and complete modeling.

Note: all functions and formulas in Excel require an equal (**=**) sign in front of the syntax.

Excel's IF statement

The IF function in Excel tests values and formulas in the target cells and returns one specified value "if" the condition input evaluates to be TRUE and another value if it evaluates to be FALSE. The IF statement in Excel is called out with the following syntax:

- IF(logical_test,value_if_true,value_if_false)

Logical_test: any expression or correlation of values that can be evaluated as TRUE or FALSE.

Value_if_true: the value or expression that is returned IF the condition in the logical test returns TRUE.

Value_if_false: the value or expression that is returned IF the condition in the logical test returns FALSE.

The **value_if_true** and the **value_if_false** part in the syntax may vary and can be numerical values, expressions, models, and formulas. If either of the values is left blank, the test will return 0 (zero) correspondingly.

Sample coding:

- IF(1+1=2, "Correct!", "Incorrect!")

For the logical test part, 1+1 is equal to 2, and the iteration will return TRUE. The **value_if_true** is called for the TRUE response, and "Correct!" will be output in the corresponding cell.

It is possible to nest multiple IF functions (called **Nested IF**) within one Excel formula. Such a formula is useful in returning multiple answers, compared to a a single IF function which can return only two possible answers.

A Nested IF statement has the following syntax:

- IF(logical_test1, value_if_true1, IF(logical_test2, value_if_true2, value_if_false))

The following example shows how a Nested IF function is used to make a spreadsheet that assigns a letter grade to different scores.

Suppose that grade distribution is:

@ A= Greater than or equal to 80

@ B= Greater than or equal to 60, but less than 80

@ C= Less than 60

This logic function can be written in several different ways. If the score is tested to be the letter grade C first, then checked to see if it is a B or an A, the syntax could be written as follows:

- IF(A5<60,"C",IF(A5<80,"B","A"))

Note that the Nested IF statement can be placed in either the **value_if_false** OR the **value_if_true** positions. In the example above, it was placed in the **value_if_false** position because of how the conditional was written.

Additional letter grades could be added by including more nested IF functions inside the existing function. One formula can include up to 7 nested IFs, which means you can return up to 8 different results.

Excel's OR statement

The OR statement in Excel is similar to the IF statement. It returns TRUE if **any** argument or expression of values is TRUE, and it returns FALSE if **all** arguments or expression of values are FALSE. The syntax for OR is:

- OR(logical1,logical2,...)

logical1: there can be up to 30 arguments for the logical expression to test and return either TRUE or FALSE. If all of the arguments are false, the statement returns FALSE for the entire expression.

Sample coding:

- OR(1+1=2, 2+2=5)

In the logical1 part, 1+1 is equal to 2 and will return TRUE. In the logical2 part, 2+2 is not equal to 5 and will return FALSE. However, since one of the test returns TRUE (logical1), the function returns TRUE even if logical2 is FALSE. **TRUE** will be output in the corresponding function cell.

Excel's AND statement

The AND statement returns TRUE if **all** its arguments or values are TRUE and returns FALSE if one or more argument or value is FALSE. The AND statement syntax is:

- AND(logical1,logical2,...)

logical1: similar to OR statement with up to 30 arguments to test that could return either TRUE or FALSE.

Sample coding:

- AND(1+1=2, 2+2=5)

Logical1 will return TRUE for the correct calculation while logical2 will return FALSE for the incorrect calculation. However, since one of the expressions is FALSE, the AND logical test will return FALSE even if one of the tests is true.

It is common and useful to understand how to implement different logical functions into each other. A good example would be:

- IF(AND(2+2=4,3+3=5),"Right Calculation","Wrong Calculation")

During iteration, the AND statement should return FALSE since one of the expression/calculation is incorrect. The FALSE return will output "Wrong Calculation" for the IF statement iteration.

Other Useful Excel Functions

Excel has many useful functions that can help any engineer who is using Excel to model or control a system. An engineer may want to know the average flow rate, number of times the liquid in a tank reaches a certain level, or maybe the number of time steps that a valve is completely open or completely closed. All of these situations and many more can be addressed by many of the simple built-in functions of Excel. Below are some of these functions and a brief definition of what they compute and how to use them. They are followed by an image of a sample excel file with an example of each function in use. To see a list of all the built-in functions Excel has to offer, click shift-enter on the cell.

AVERAGE(number1,number2,...) - Returns the average or arithmetic mean of the arguments. The arguments can be numbers, arrays, or references that contain numbers.

CEILING(number1, significance) - Rounds a number up to the nearest integer or multiple of significance.

CHIDIST(x, degrees_freedom) - Returns the probability of the chi-squared distribution.

COUNT(value1,value2,...) - Counts the number of cells that contains numbers.

COUNTIF(range, criteria) - Counts the number of cells within the range that meet the given criteria.

COUNTIFS(criteria_range1,criteria1,criteria_range2,criteria2,...) - Counts the number of cells within multiple ranges that meet multiple criteria.

FLOOR(number1,significance) - Rounds a number down toward zero, to the nearest multiple of significance.

FREQUENCY(data_array,bins_array) - Calculates how often a value occurs within a range of values (bins). The function will return a vertical array of numbers having one more element than the bins_array. Once the formula is entered into the target cell, highlight the target cell making a vertical array one cell larger than bins_array. Once selected press F2 to open the formula of the target cell and then press CTRL+SHIFT+ENTER to execute your function.

LARGE(array,k) - Returns the k-th largest value in a data set. For example, the 2nd largest number.

MAX(number1,number2,...) - Returns the largest number in a set of values.

MEDIAN(number1,number2,...) - Returns the median of a set of values.

MIN(number1,number2,...) - Returns the minimum value in a set of values.

MODE(number1,number2,...) - Returns the most frequently occurring value in an array or range of values.

PERCENTILE(array,k) - Returns the k-th percentile of values in a range

ROUND(number1, num_digits) - Rounds a number to the specified number of digits.

SMALL(array,k) - Returns the k-th smallest value in a data set. For example, the 7th smallest number.

STDEV(number1,number2,...) - Returns the standard deviation of a given set of values ignoring text and logical values.

SUM(number1,number2,...) - Adds all of the values in a range of cells.

SUMIF(range,criteria,range,...) - Adds values of cells in a range meeting criteria.

VAR(number1,number2,...) - Estimates the variance of a given set of values.

	A	B	C	D	E
1					
2				Range	Bins
3	AVERAGE(D3:D26)	4.125		1	2
4	COUNT(D3:D26)	24		1	5
5	COUNTIF(D3:D26,3)	3		1	9
6	FREQUENCY(D3:D26,E3:E6)	7		1	
7		10		2	
8		7		2	
9		0		2	
10	LARGE(D3:D26,3)	7		3	
11	MAX(D3:D26)	8		3	
12	MEDIAN(D3:D26)	4		3	
13	MIN(D3:D26)	1		4	
14	MODE(D3:D26)	1		4	
15	SMALL(D3:D26,6)	2		4	
16	STDEV(D3:D26)	2.251810865		5	
17	SUM(D3:D26)	99		5	
18	VAR(D3:D26)	5.070652174		5	
19				5	
20				6	
21				6	
22				6	
23				7	
24				7	
25				8	
26				8	
27					
28					

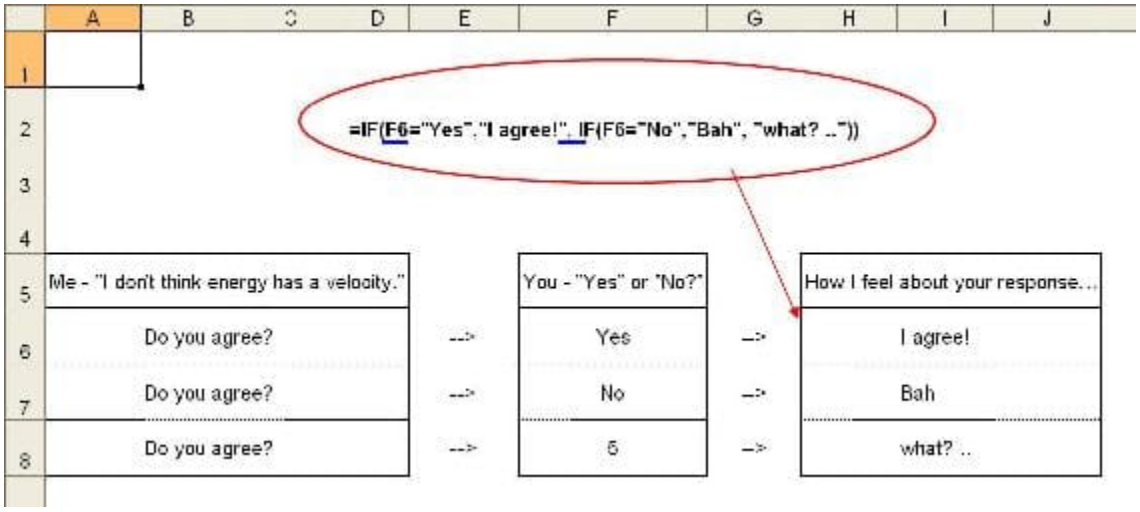
Nesting

Though not complex or difficult, placing logical statements within one another can be a powerful practice. This is referred to as **nesting**.

An example demonstrating the use of a simple **IF** statement **nested** within another **IF** statement is shown below:

If one were to ask three individuals whether or not they agree that energy has a velocity, these individual will hypothetically respond in one of three ways: "Yes","No", or (other). In this case, if the individual were to respond "Yes", then the questioner would respond "I agree!". If the individual were to respond "No", the questioner would bitterly respond with a "Bah". And if the individual were to respond with a nonsensical answer, then the questioner would respond "what .." in confusion.

The important point to get out of this example is that **nested** statements gives us the potential to process more information in one cell (or line of programming code).



Excel's Data Analysis Tools

Excel has several useful built in add-ins that can make data analysis much easier. Of these we will cover the solver tool, as well as the ANOVA analysis.

Solver Tool

The **solver** in Excel is a built-in tool with multiple capabilities. It could be used to optimize (find the maximum or the minimum value) a specific cell (the target cell) by varying the values in other cells (the adjustable cells). It may also be used to solve a system of non-linear equations. Solver formulas embedded in the spreadsheet tie the value in the target cell to the values in the adjustable cells. The user can also specify numerical range constraints for the cells involved.

Optimization Model

In optimization, one seeks to maximize or minimize the value of a real function. Just as a student in a calculus class might use optimization techniques to find a local maxima or minima of a complex function, a CEO might use similar techniques to find how much product their company should manufacture to maximize profits. There are numerous optimization techniques, many specialized for specific types problems. For the most part, MS Excel uses gradient-based techniques.

The solver tool is located in the **Tools** menu. If it does not appear there, it must be added in by selecting **Add-ins** and selecting the appropriate check box.

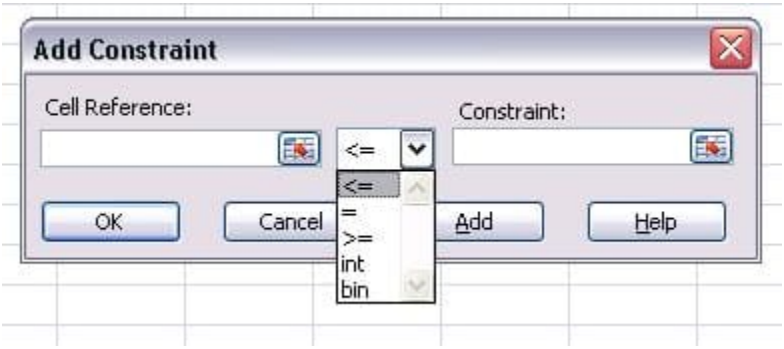
The solver window contains the following fields:

- **Set Target Cell:** This is the cell containing the value you want to optimize.

- **Equal To:** Choose whether to maximize the value, minimize the value, or set it to a specific number.
- **By Changing Cells:** Here you specify all the cells that can be varied when optimizing the target cell value.
- **Subject to the Constraints:** Constraints are optional, but can be added by clicking **Add** and defining each constraint.
- **List of Available Constraints**

- **<= (Less than or equal to)** Stipulates that a selected cell must be less than or equal to a certain value.
- **= (Equal to)** Stipulates that a selected cell must be equal to a certain value.
- **>= (Greater than or equal to)** Stipulates that a selected cell must be greater than or equal to a certain value.
- **int (Integer)** Stipulates that a selected cell must be an integer
- **bin (Binary)** Stipulates that a selected cell must be equal to 1 or 0.

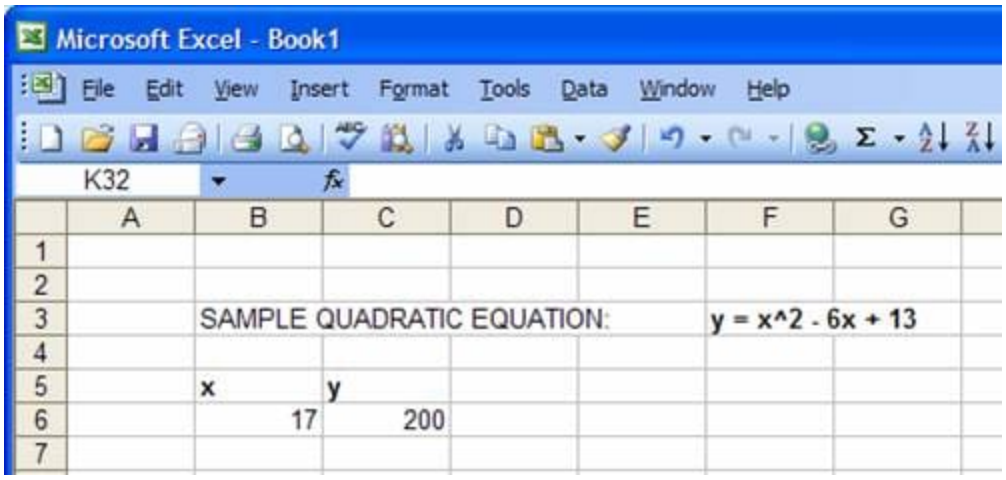
The window where these constraints can be entered is shown below.



It is important to make sure that constraints do not violate each other otherwise solver will not work without reporting an error. For example, if a cell must be great than 2 and less than 1 Solver will not change anything.

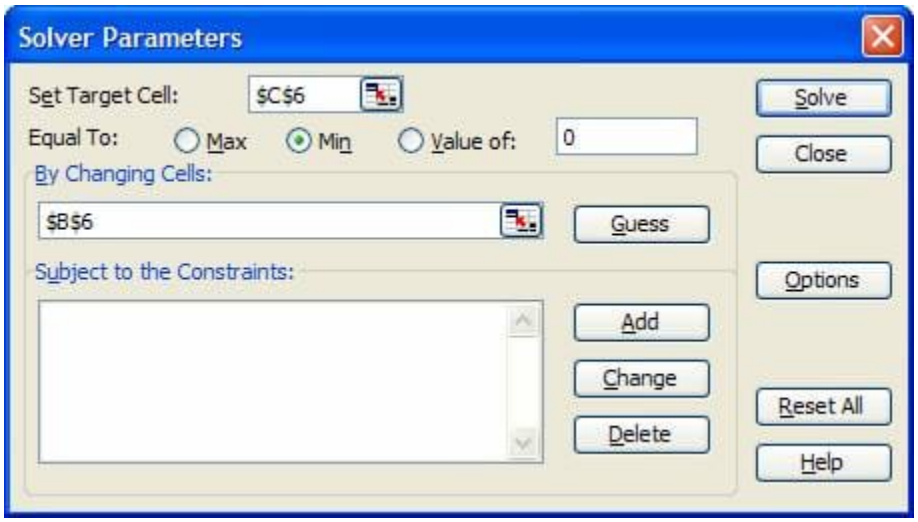
Once all the fields in the Solver window are completed appropriately, click **Solve**. If solver can find a solution a window will pop up telling you so, and the values will appear in the target and adjustable cells. If solver cannot find a solution it will still display the values it found but will state that it could not find a feasible solution. Be aware that Excel solver is not a powerful enough tool to find a solution to nonlinear equations, where programs such as Matlab might be a better choice.

*** Guided Example**: Suppose you have an Excel spreadsheet set up as shown below, with cell C6 depending on cell B6 according to the quadratic relationship shown, and you want to minimize the value in C6 by varying the value in B6.



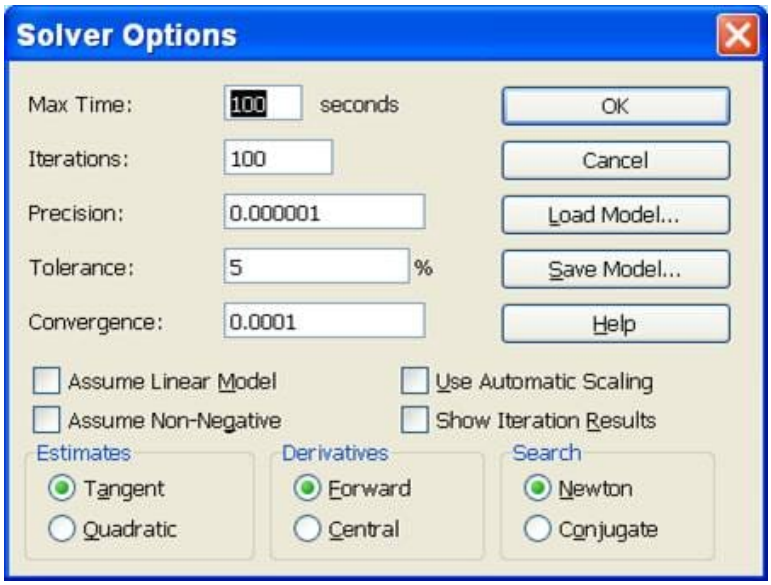
1) Open up the solver window.

2) Input C6 in the **set target cell** field, check the **min** button, and input B6 in the **by changing cells** field:



3) Click **solve** to find the solution.

You can place further constraints or further specify how you want your task solved under "Options." Some of the most commonly used specifications include:



- **Max Time:** This is the maximum amount of time excel will spend trying to find a converging value before giving up.
- **Iterations:** This is the number of iterations excel will perform to converge to a value.
- **Precision:** This is related to how close excel has to get before a value is "acceptable." In general, making this value several orders of magnitude smaller will yield a more accurate result, but will take more time to do so, and may have a harder time converging.
- **Tolerance:** This is also related to how accurate a solution is, this is related to the percent error. Making this smaller should also yield a more accurate result, but with more difficulty finding a solution.
- **Convergence:** Controls the amount of relative change that occurs during the last five iterations that Solver performs. The smaller the number of the convergence the less change that will occur.
- **Box Options** Click on the box to select the following options:
- **Assume Linear Model:** This will speed up the time to find a solution because it will make Solver assume that the model is linear.
- **Assume Non-Negative:** Solver will make the lower limit of adjust zero for all the cells that are programmed to be adjusted in the model that are not limited by a chosen constraint.
- **Use Automatic Scaling:** Used for large differences in magnitude to scale the inputs and outputs.
- **Show Iteration Results:** Solver will pause between each iteration to show the results of that iteration in the model.
- **Estimates** Specifies to Solver on how to use initial estimates for iterations
- **Tangent:** Uses linear extrapolation by utilizing a tanget vector.
- **Quadratic:** Uses quadratic extrapolation. Best to use for nonlinear models.
- **Derivatives** Specifies what type of differencing is used for partial derivatives for both the objective and constraints.
- **Forward:** Constrained values change slowly between iterations.

Nonlinear Regression

Excel's solver function can also be used to find a solution for two-variable non-linear regression. The description of the data by a function is carried out by the process of iterative nonlinear regression. This process minimizes the value of the squared sum of the difference between the actual data and the predicted value (residual).

There are three basic requirements for using solver including:

1. Raw data
2. A model equation to find predicted values
3. Initial guesses for varying values

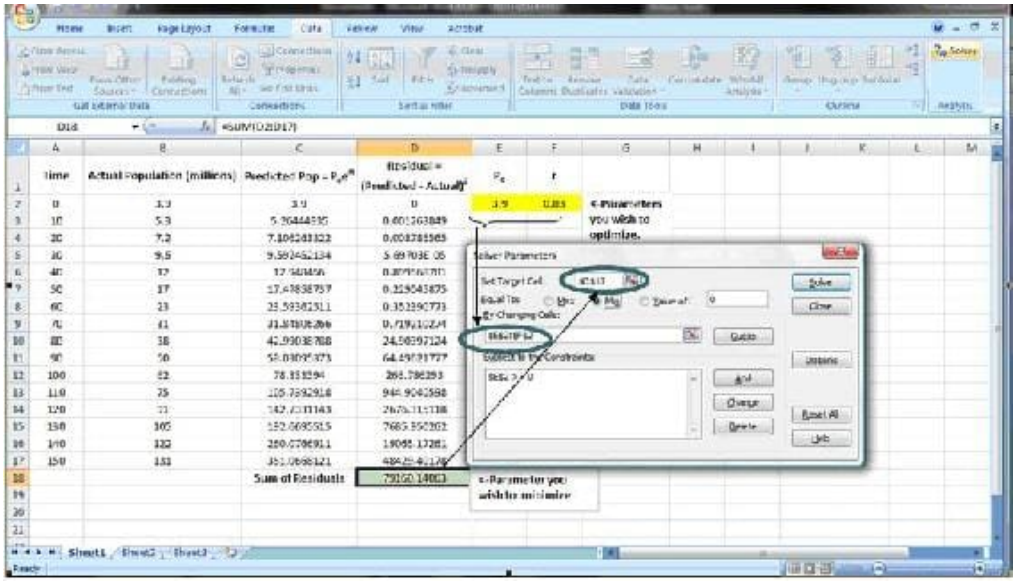
Before using the solver function, the raw data should be organized into separate columns for the independent variable (first column) and dependent variable (second column). A third column will show the predicted value that is obtained by using the model equation mentioned earlier, and must reference various single cells that have initial values that Excel will use as a starting point for the regression. A fourth column should be created with the square of the difference between the predicted value and the

actual experimental value (raw data); this is also known as the square of the residual. A cell at the bottom of this fourth column should include the sum of all the squared difference values above it. Two more cells anywhere else in the worksheet should be designated as the initial guesses for the values to be varied to find the solution.

It is important to choose a reasonable value for the initial guess. This will not increase the chance that Excel arrives at a global optimization rather than a local optimization, but also reduces the time it takes for Excel to solve the system of non-linear equations.

Now, solver can be applied to the spreadsheet to optimize the regression. Refer to section 3.1 for how to access and use solver. The bottom most cell of the fourth column that has the sum of the squares of the residuals will be the target cell. Since the residuals are essentially the error, the sum of the residuals squared should be as small as possible, so the min circle should be selected. The initial guesses should be entered into the "by changing cells". Constraints can be entered if the values of the cells are limited to a certain range. Solver will produce the values of the varying parameters that yield the best fit to the nonlinear data.

- Guided Example: A screen shot of a simple population problem is shown below. The yellow highlighted cells are the initial guesses for the variables. In this case, the initial population (P_0, cell E2) is entered as a constraint.



ANOVA Analysis

ANOVA is another tool on Excel that can be used for data analysis purposes. It is used primarily to compare continuous measurements to determine if they are sampled from the same or different distributions. This is usually done on set of data that consists of numerous samples and/or groups. Single factor ANOVA analyzes the difference between "groups" which can be different units, sensors, etc. On the other hand 2 way ANOVA analyzes the differences on the basis of both samples (trials) and groups, allowing you to compare, which one most affects the system. Two-way ANOVA can be employed for two different cases: data with or data without replicates. Data without replicates is used when collecting a single data point for a specified condition, while data with replicates is used when collecting multiple data points for a specific condition.

The telling result of ANOVA analysis is a p-value telling you if there is a statistically significant difference between data sets. A more in depth discussion of the uses and implementation of ANOVA can be seen here.[1]

Random Number Sampling

A useful numerical technique is the random number generator. This generator uses random number sampling to create a predictive model. This can be especially useful in probability problems.

A typical random number generator produces any number between 0 and 1 with equal frequency. At each step a random number is generated that picks a possible outcome. A typical Excel statement would appear as the following: **IF (RAND(). This means that X percent of the time, you would get a value1 and 1-X percent of the time, you would get a value2. For example,**

suppose a situation exists with two possible outcomes, A and B, where A takes place 70% of the time. If a random number is generated that is greater than 0.3, then the outcome is A. Inversely if the random number is less than or equal to 0.3, then B is the outcome. This would be represented as IF(RAND()<0.7, A, B). Each step within the generator does not affect the next step, i.e. each step is determined individually. A process in which previous outcomes do not affect later outcomes is known as a Stochastic method or "Monte Carlo" method. A Monte Carlo simulation uses a model that takes random input that follows an assumed distribution (in this case, 70% A and 30% B) and produces output that models some phenomenon. The following is a simple example of such a simulation:

The Excel is equipped with a random number generator which can be called out using the RAND function. A random number between 0 and 1 can be generated in a cell by the following command:

In Microsoft Excel 2007 the RAND function returns a new random number every time the worksheet is calculated or when F9 is pressed. If the same random number is needed to be referred to more than once in an Excel IF, AND, OR statement, one can create a column of random numbers using =RAND(), and refer back to these cells in an adjacent column.

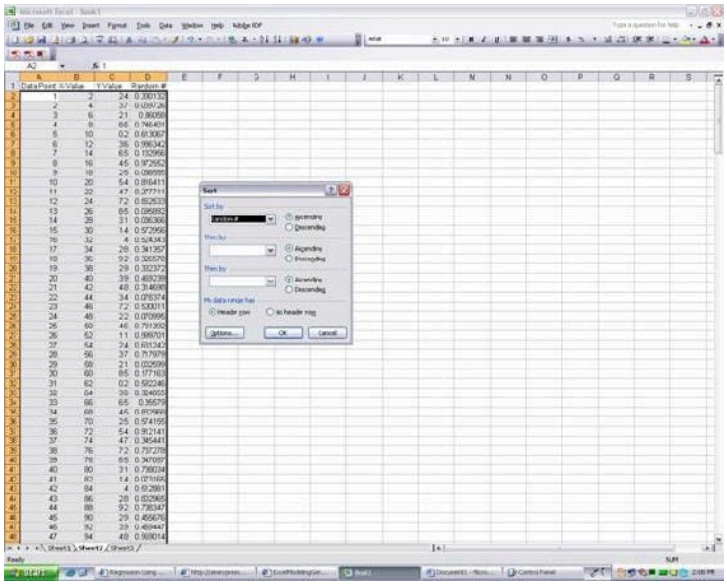
There are many ways to manipulate the random numbers that Excel produces. A random number can be generated within a different range by multiplying the RAND function by a constant, i.e. multiply by 100 to change a decimal to percent. Another option to generate numbers within a different range is the RANDBETWEEN function. RANDBETWEEN generates a random integer within a specified range. The following command generates a random integer between -6 and 21:

It is also possible to modify your random numbers by nesting the RAND function within another function. For example, one can square the results of one distribution to create a different distribution of numbers.

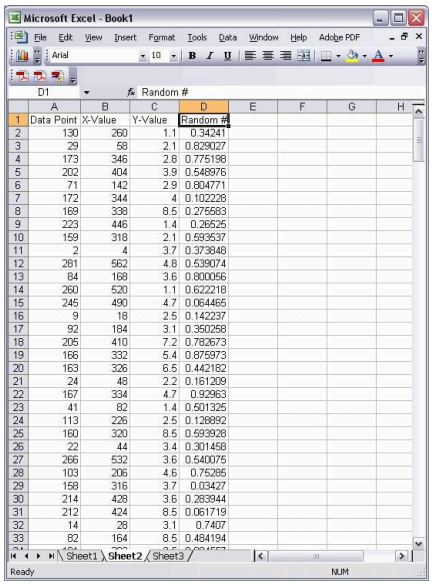
Another way to utilize the RAND function is to take a random sampling of a data set. If one has a spreadsheet with several hundred data points, and would like to analyze only 100 random points, another column can be created titled "Random Number" or something similar. In the first cell of this column, type the =RAND() function. Then drag this function all the way to the end of the data set. At this point, the user can go to the DATA drag-down menu and select Sort... In this menu, select your random number column in the SORT BY drag down menu. Either ascending or descending can be selected, as desired. Once the spreadsheet is sorted by the random numbered cells, the first 100 data points can be used as a statistically random cross-section of the data.

The sort tool can be useful in cases when you are trying to find out the probability of having certain values above a certain threshold. For example, if the 100 data points that you have represent the temperature variations occurring within a system, you can use sort tool to determine how many temperature values exceed your threshold. From that, you can calculate approximately the probability of the system temperature exceeding the threshold. Two screenshots of the sort process are shown below. Microsoft Excel 2007 uses a different menu, allowing for one to choose between sorting from Largest to Smallest or vice versa. A screenshot below shows this new menu.

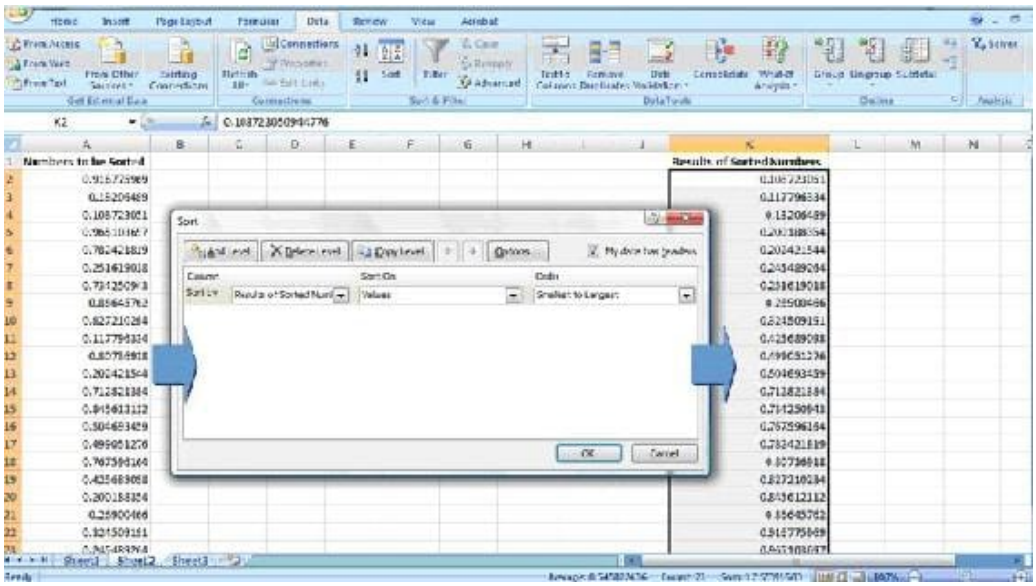
Sorting Setup:



Results to Randomization:



Microsoft Excel 2007:



Worked out Example 1

The Arrhenius equation defines the relationship between a reaction rate constant k and temperature:

$$k(T) = Ae^{-E/RT}$$

where *T* is the absolute temperature, *A* is the frequency factor, *E* is the activation energy, and *R* is the universal gas constant.

It is frequently desirable to be able to predict reaction rates at a given temperature, but first you need to know the values of *A* and *E*. One way to obtain these values is to determine the reaction rate constant at a few temperatures and then perform a nonlinear regression to fit the data to the Arrhenius equation. Given the following rate constant data for the decomposition of benzene diazonium chloride, determine the frequency factor and activation energy for the reaction.

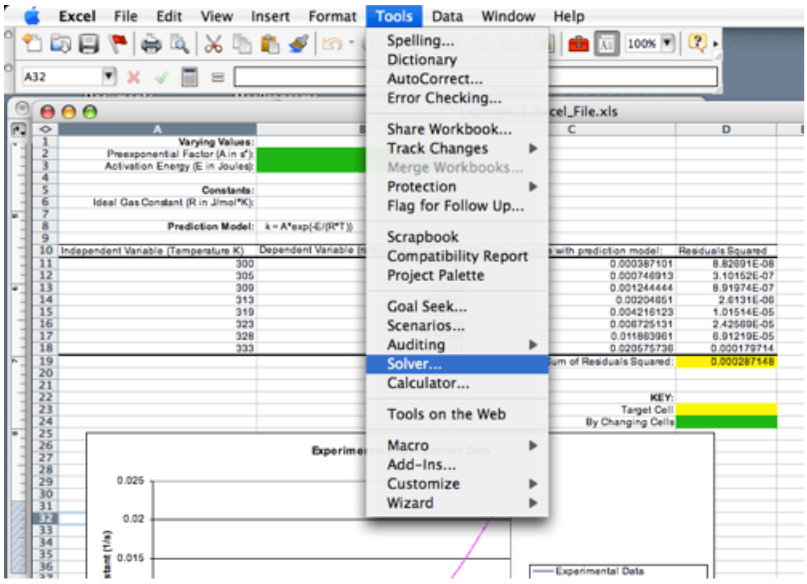
(s^{-1})
0.00043
0.00103
0.00180
0.00355
0.00717

Solution: The following Excel file contains a solution to the problem that uses Excel's solver tool to perform the nonlinear regression and determine the values of *A* and *E*:

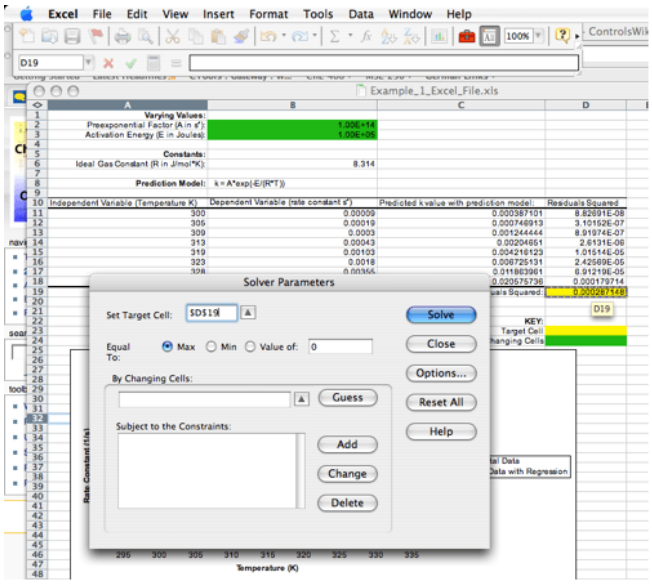
Example 1

The spreadsheet is set up as described in the nonlinear regression section above.

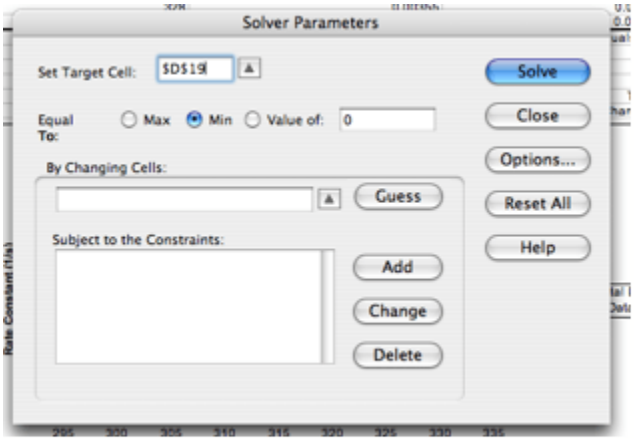
1. Open solver (Tools / Solver)



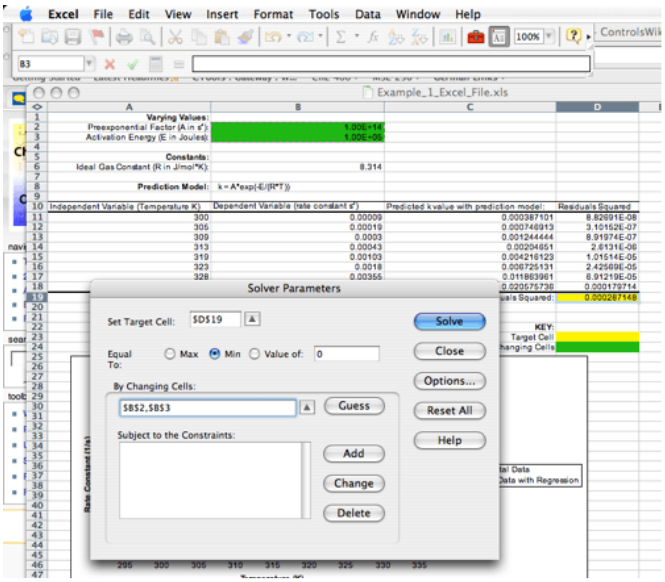
2. Set the Sum of Residuals Squared value cell as the target cell



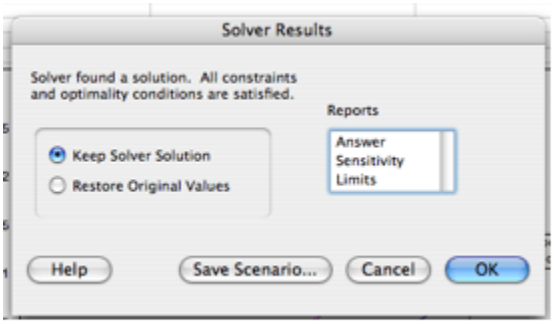
3. Click the Min radio button



4. Set Pre-exponential Factor and Activation Energy as the adjustable cells. For this problem, keep *A* between 1E+13 and 1E+14 s^-1 and *E* between 9.5E+4 and 1.05E+5 J



5. Sometimes, depending on what version of excel you are using, a message will come up like the one below. If this happens, click OK. The graph will update with the new curves.



6. Click Solve and observe the values of A and E calculated by Solver and the changes to the plot showing the rate constant as a function of temperature.

NOTE: If solver is not fitting the data well, click the options button in the solver window and make the precision value smaller by several orders of magnitude.

Worked out Example 2

Kinetics can be used to determine the probability of a reaction taking place between gas-phase molecules. Consider molecules A and B, which have a 60% chance of reacting upon collision. Create an Excel spreadsheet that utilizes a random number generator and an if-statement to model an individual collision.

Solution

In this Excel file, sample coding for Random number generator and if-statement is shown.

Collision_ex

Random number generator is used to output numbers between 0 and 1 for four trials, ten trials, and one hundred trials on sheets one, two, and three, respectively. Formula for the random number generator is employed in the cell where you want the number to be output (in this case, B6 is the cell). An IF logical test is then employed to test whether the number follows the requirement or not. The IF-statement cell will then output the corresponding result based on the logic test. In this collision case, if the random number function generated a number greater than 0.4, the logic test will return TRUE and then output "Reacts" as the result. Assuming that the random number generator produces an

even distribution between 0 and 1, the random number will be greater than 0.4 sixty percent of the time. Coding of the IF-statement is shown:

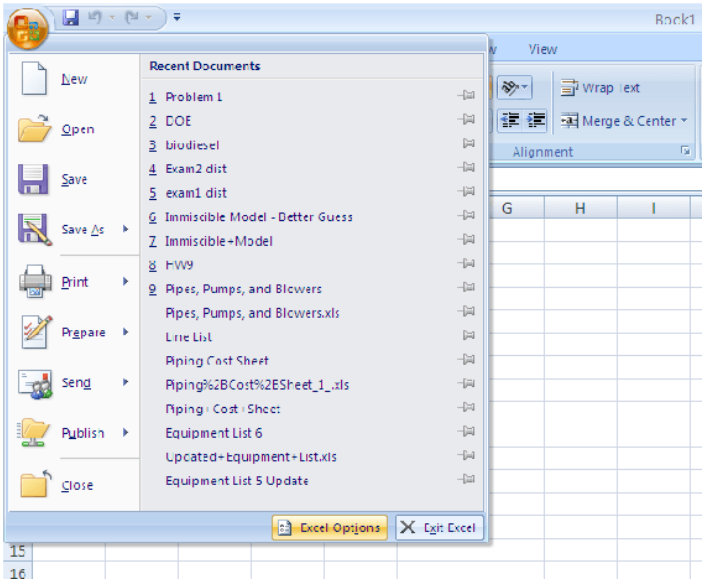
On the other hand, if the number generated is less than 0.4, the logic test returns FALSE and "Does not React" will be shown as the result. Different numbers are generated each time the spreadsheet is updated, thus different result will return for the IF function. You can also modify the numbers and return statement to better understand the operation of the functions.

The number of reactions is summed up and divided by the total number of trials to compare the predicted reaction probability with the value given in the problem. It can be seen by comparing sheets one, two, and three that increasing the number of trials decreases the variation in the predicted reaction probability.

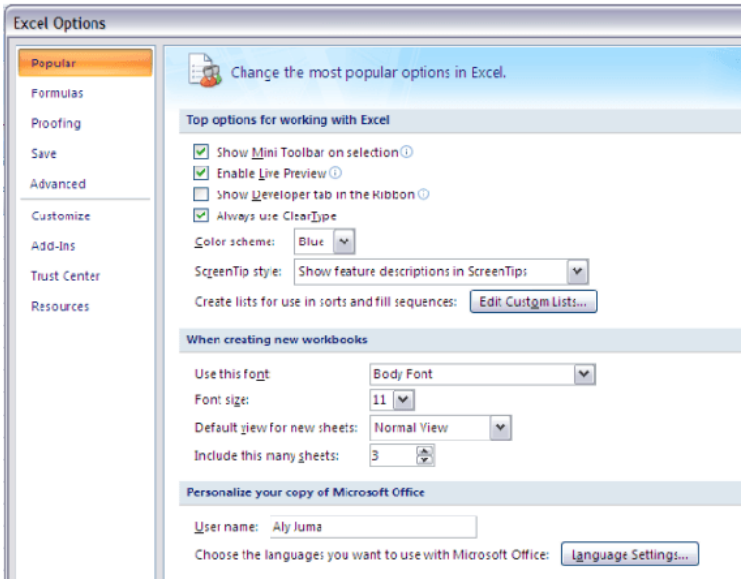
Accessing Excel Tools on Windows

On many computers, excel tools such as solver and the data analysis package are not already included and must be installed before used. This section will go over how to access these add-ins for those using Windows opposed to Macs.

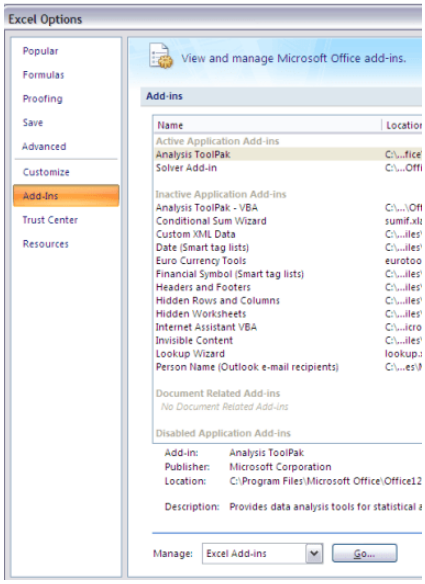
1) Click on the windows symbol in Excel 2007. At the bottom of the opened window, select the Excel options box.



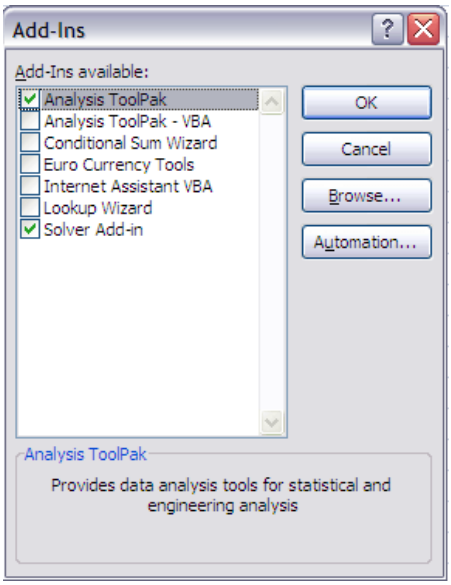
2) On the left side panel you will see a section labeled "Add-ins". Select this.



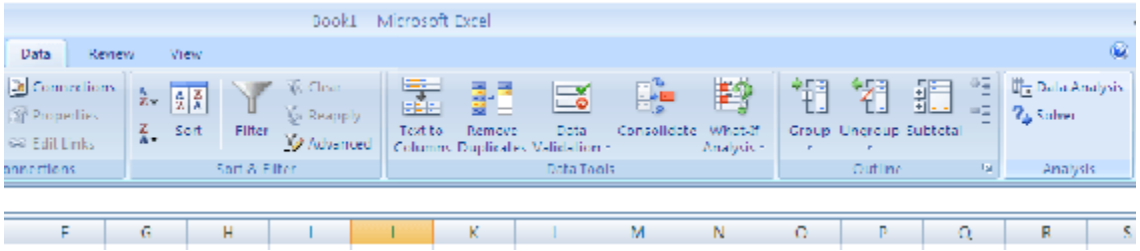
3) Here you will see a list of "Inactive Applications" that are available for excel. Of these the ones we are interested in are the "Solver Add-In" and the "Analysis ToolPak." Select one of these and click "Go" at the bottom of the window.



4) A new window will appear asking you which components of the package you would like to install. Select the needed tools (solver and analysis toolpak). Click "Ok" and windows should install the corresponding tools into excel.



5) Once done, these tools will be found under the data section of Excel, on the far right of the window.



References

1. **=RAND()**

- **=RANDBETWEEEN(-6,21)**
- **Collision Example Excel File:**
- **=IF(B6>=0.4,"Reacts","Does not React")**
- **Bender, E.A. *An Introduction to Mathematical Modeling,* Mineola, NY: Dover Publications, Inc.**
- **Fogler, H.S. (2006). *Elements of Chemical Reaction Engineering,* Upper Saddle River, NJ: Prentice Hall Professional Technical Reference. ISBN 0-13-047394-4**
- **Microsoft, Excel, Proprietary EULA, www.microsoft.com.**

Contributors and Attributions

Authors: (12 September 2006 / Date Revised: 19 September 2006) Curt Longcore, Ben Van Kuiken, Jeffrey Carey, Angela Yeung

Stewards: (13 September 2007) So Hyun Ahn, Kyle Goszyk, Michael Peterson, Samuel Seo

This page titled 2.4: Excel Modeling - logical models, optimization with solver for nonlinear regression, sampling random numbers is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles

Introduction

Noise is all around us in all sorts of forms. In the common use of the word, noise refers to sound. However, noise can more accurately be thought of as random variations that are always present in one or more parts of any entity such as voltage, current, or even data. Rather than thinking of noise only as an acoustic term, it should be thought of more as a random signal. Noise can be the inherent fluctuations in some part in a system (ie. temperature at a given point) or it can be the unavoidable interference on a measurement from outside sources (ie. vibrations from a nearby generator blur measurements from a pressure transducer). The static interference on your radio, the 'snow' on your television, and the unresolved peaks on an infrared spectroscopy report are all examples of noise.

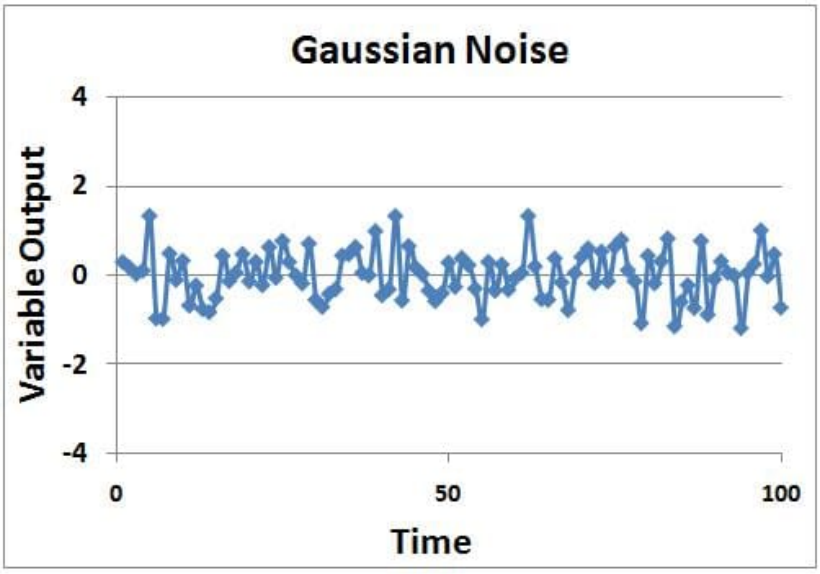
Chemical engineers can use statistical properties to characterize noise so they can understand a current process or develop an optimal process. By characterizing the noise and determining its source, the engineer can devise methods to account for the noise, control a process, or predict the path of a system. For example, when chemical engineers design plants, they use characterizations of noise to determine the best control scheme for each process. Mathematical modeling can be used to characterize and predict the noise of a given system. In modeling, to simplify the representation of general trends that reoccur, noise is classified in two major categories: frequency based and non-frequency based. **Frequency based noise** consists of the colors of noise, and **non-frequency based noise** includes pops, snaps and crackle.

The purpose of the following sections is to give you a qualitative understanding of the sources of noise and how it can be characterized and handled. Examples will be given to help you gain a quantitative understanding of how noise relates to controlling and modeling chemical engineering processes.

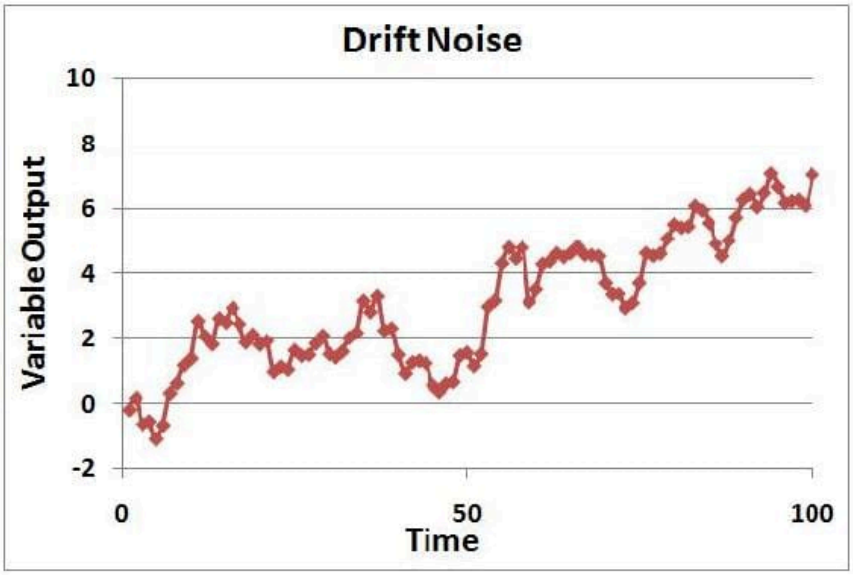
General Types of Noise

Three general types of noise can be categorized in describing a process, and are as follows:

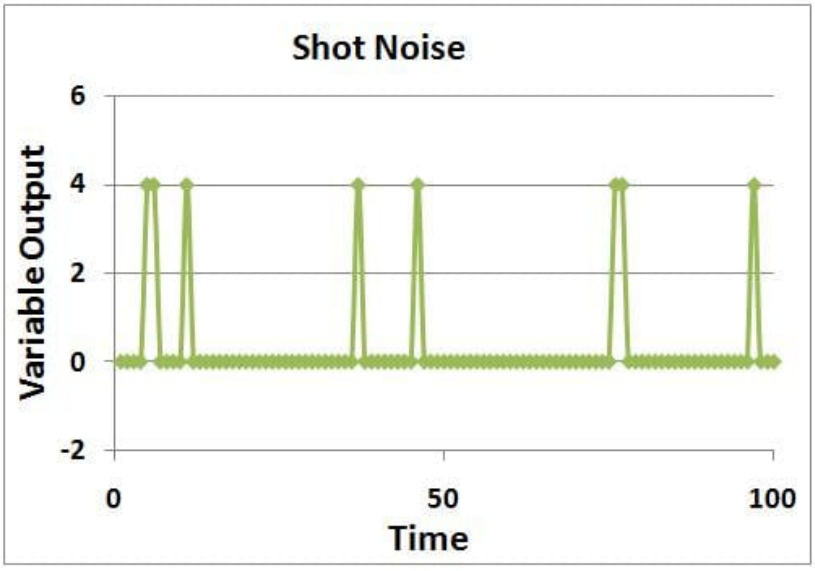
1) Gaussian noise is usually not dependent on time, meaning that it is random and not systematically planned. The amplitude of the frequency can vary, making a crackling notation or sound (see #Crackles below). Some examples of Gaussian noise can include splashing in a tank or unplanned interruption in a sensing device.



2) Drift noise is correlated to time, and has random movement. Examples of drift noise can include fouling or catalyst decay, in which the potency of the substance may decline over time as decaying occurs. Stock price fluctuations also model somewhat like drift noise.



3) Shot noise may be defined as sporadic and short bursts of noise, in which the amplitude is similar among the bursts of noise. Shot noise can be correlated to pops (see #Pops below), in which at random times, the same shot noise is witnessed with the same amplitude. Examples of shot noise include partial clogging or jamming in the process in which the same amplitude will be seen by the noise whenever the clogging or jamming in the process occurs. Another example is customer demand of the product. If the control system (when trying to optimize it) depends on the customer order, and customer demand is not consistent at all times (meaning downtimes for orders) but the order amount is the same when orders are placed, then it will be effected by the shot noise described by the customer order (customer demand).



Microsoft Excel Formulas for 3 Types of Noise

Noise Type	Excel Formula
Gaussian Noise	= [*Original value*] + RAND() - RAND() + RAND() - RAND()[1]

Noise Type	Excel Formula
Drift Noise	= [*Previous value*] + RAND() - RAND() + RAND() - RAND()[1]
Shot Noise	= if(RAND()>0.9, [*Original value*]+[*single error*], [*Original value*])

[1]The more add/subtract cycles of RAND() you use, the closer approximation you'll get for a Gaussian distribution of random errors.

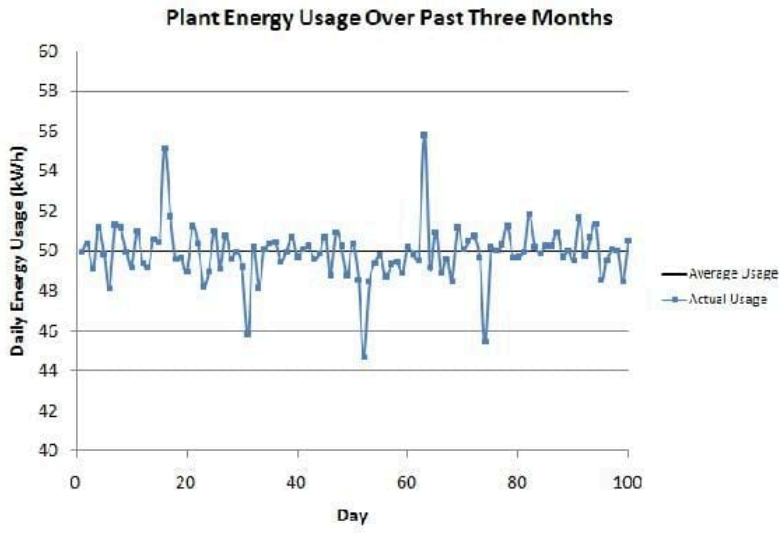
As you would see these in an Excel file:

\	A	B	C	D
1	Time	Gaussian Noise	Drift Noise	Shot Noise
2	1	**original value**	**original value**	**original value**
3	2	= B2 + RAND() - RAND() + RAND() - RAND()	= C2 + RAND() - RAND() + RAND() - RAND()	= if(RAND()>0.9, D2+[*single error value*], D2)
4	3	= B2 + RAND() - RAND() + RAND() - RAND()	= C3 + RAND() - RAND() + RAND() - RAND()	= if(RAND()>0.9, D2+[*single error value*], D2)

Note: These formulas can be extended for any length of time. Longer periods of time are useful for visualizing drift noise.

Combined Types of Noise

Sensors, processes, demands, etc. often do not behave with simple noise models. However, most noise can be derived from the three general types of noise. An example of a process with random and shot noise is shown in the thumbnail.



This is a graph of the total daily electricity usage for a plant during a normal operation period. As we can see, there are minor fluctuations occurring every day, but some days there are large changes of electricity usage. These large changes can be seen as shot noise when the consumption is viewed on this daily scale.

Possible reasons for the minor fluctuations could be due to electric heaters kicking on and off. Changes in the operation of smaller pumps can cause these changes in electricity demand.

The large shots could be due to an energy-intensive part of the process that only operates when needed. The positive shots could be due to this process kicking on and the negative shots could be due to a large process kicking off when not needed.

With this graph, we can say that there is shot noise present with the large changes in consumption, but there is also random noise present with the minor daily fluctuations. We would not say there is drift noise present because the consumption remains around a constant mean value and does not drift away from there.

However, the total consumption when viewed from a different scale may show other forms of noise. If we viewed the hourly usage of electricity on one of the days with shot noise, it may resemble drift noise on the hourly scale.

Colors of Noise

The first category of noise is frequency-based noise, classified by the colors of noise. In understanding the colors of noise, we need to understand the process of converting a given set of data to a point where it can be classified as some color of noise. The section below, "Modeling Noise," goes into more detail as to how this is done. At this point, it is important to understand the main colors of noise which are possible in a system. Noise is classified by the spectral density, which is proportional to the reciprocal of frequency (f) raised to the power of beta. The power spectral density (watts per hertz) illustrates how the power (watts) or intensity (watts per square meter) of a signal varies with frequency (hertz).

$$PSD \propto \frac{1}{f^{\beta}}$$

where $\beta \geq 0$.

Just like colors of light are discriminated by their frequencies, there are different colors of noise. Each beta value corresponds to a different color of noise. For example, when beta equals zero, noise is described as white and when beta equals two, noise is described as brown.

White Noise

One major type of noise is white noise. A defining characteristic of this type of noise is it has a flat power spectral density, meaning it has equal power at any frequency. β equals zero for white noise.

White noise and white light have similar properties. White light consists of all visible colors in the spectrum while white noise is created by combining sounds at all different frequencies. Pure white noise across all frequencies cannot physically exist. It requires an infinite amount of energy and all known energy is finite. White noise can only be created within a specific and defined range of frequencies. Again making an analogy to white light, for a small band of frequencies, *visible* white light has a flat frequency spectrum. Another property of white noise is that it is independent of time. This factor also contributes to the idea that white noise cannot physically exist since it is impossible to have noise be completely independent of time.

A sample power spectral density chart of white noise is shown below. As can be seen from the chart, the power over all displayed frequencies is essentially the same, suggesting the PSD is equal to 1.

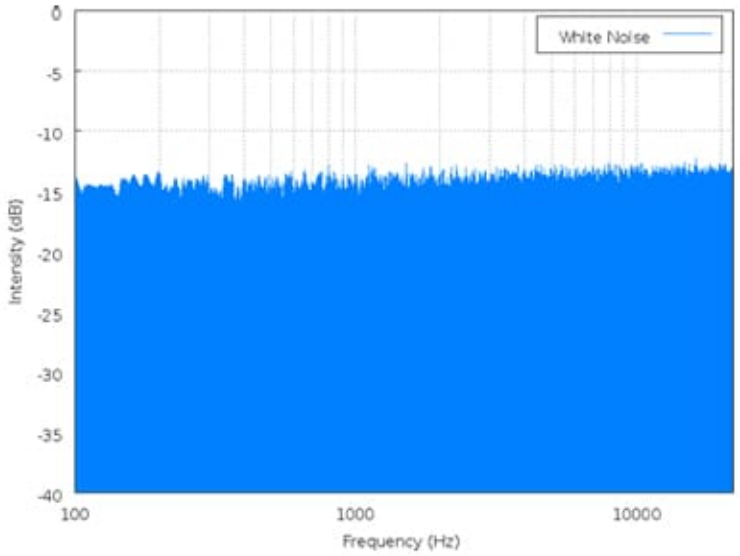


image from en.Wikipedia.org/wiki/White_noise on 9/18/2006

It is important to note that white noise is not always Gaussian noise. Gaussian noise means the probability density function of the noise has a Gaussian distribution, which basically defines the probability of the signal having a certain value. Whereas white noise simply means that the signal power is distributed equally over time. For Gaussian white noise one finds the number of measurements greater than the power spectral line as measurements that are less than the power spectral line with majority falling at the spectral line. Gaussian white noise is often used as a model for background noise in satellite communication. White noise can also come from other distributions, such at the Poisson distribution. Poissonian white noise will look like a normal distribution that has been shifted to the left for small number of measurements. For larger number of measurements Poissonian white noise will look like the normal distribution (also known as gaussian white noise). A Poissonian white noise model is useful for systems where things happen in discrete amounts such as photons being transmitted or the number of cars that go through an intersection.

Pink Noise

Pink noise is a signal whose power spectral density decreases proportionally to the inverse of the frequency, where the β value is equal to one. Because of its decrease in power at lower frequencies when compared to white noise, pink noise often sounds softer and damper than white noise.

Pink noise is actually what is commonly considered audible white noise and is evident in electrical circuits. Pink noise becomes important to measure in things such as carbon composition resistors, where it is found as the excess noise above the produced thermal noise. Below is a power spectral density chart for Pink noise.

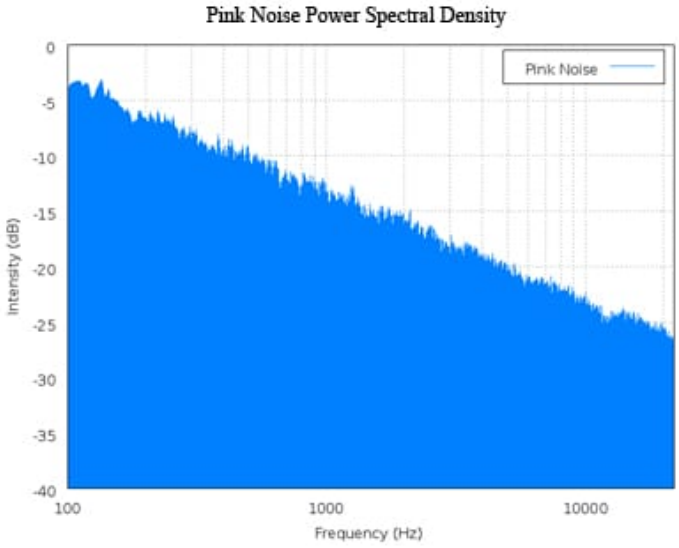


image from en.Wikipedia.org/wiki/Pink_noise on 9/18/2006

Brown Noise

When β equals 2, the noise is Brownian. Brown noise is signal noise created by Brownian, or random, motion. It is a form of unavoidable interference in the transmission of information and can be compared to a random walk that does not have a clearly patterned path. Brown noise has more energy at lower frequencies. Compared to white or pink noise, brown noise is even more soft and damped. It can be generated by integrating white noise. Interference from outside sources, such as vibrations from nearby machinery or background light, in instrument readings usually have a brown noise pattern.

Below is a power spectral density chart for brown noise. From the charts of brown noise and pink noise, it can be observed that brown noise loses power as frequency increases at a much faster rate than that of pink noise.

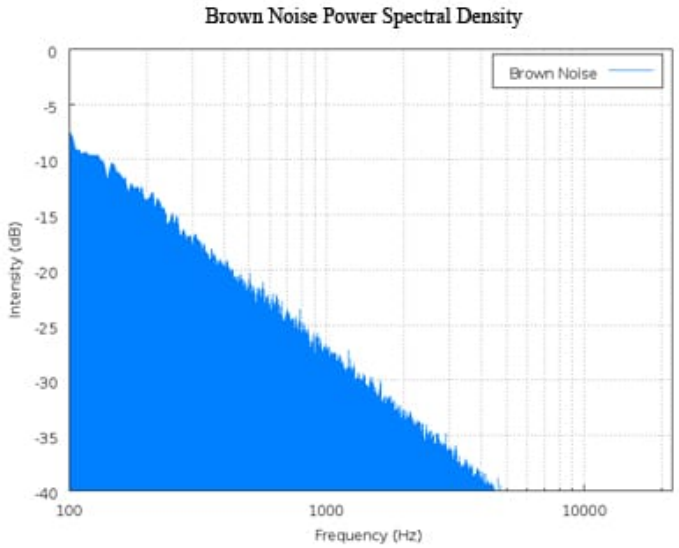


image from en.Wikipedia.org/wiki/Brown_noise on 9/18/2006

Blue Noise

Blue noise is a signal whose power spectral density increases proportionally to the frequency, where the β value is equal to negative one. An example of would be a reversible, exothermic, batch reaction where the products are being accumulated at an overall constant rate, but the reverse reaction occurs when the temperature gets too high. Thus it is overall increasing in product, but has small fluctuations as it increases. Below is a power spectral density chart for blue noise.

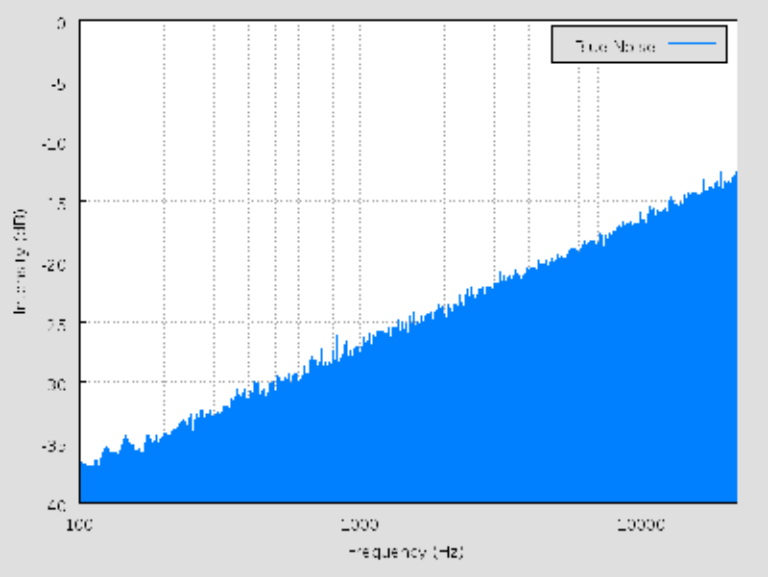


image from en.Wikipedia.org/wiki/Blue_noise#Blue_noise on 12/11/2008

Purple Noise

When β equals -2, the noise is violet. Purple noise has more energy at higher frequencies. It can be generated by differentiating white noise. Below is a power spectral density chart for purple noise. From the charts of purple noise and blue noise, it can be observed that purple noise gains power as frequency increases at a much faster rate than that of blue noise.

Modeling Colors of Noise

Characterizing noisy signals is important to a chemical engineer so that he or she can determine the sources of noise and how to account for it. Sometimes, the noise will be a characteristic of the system. In these cases, the system can not have a very sensitive control because if the sensor for the controlling device is too sensitive, it will respond to the noise of your system and end up controlling your system in a noisy way. Other times, the noise will be from unavoidable disturbances to the system. In these cases, the noise can be damped or accounted for in an appropriate manner.

In order to apply the knowledge of the colors of noise we need to understand the process of converting a given set of data to a point where it can be classified as some color of noise. The general process of classifying noise follows.

$$\text{Data} \Longrightarrow \text{Curve to Fit Data} \Longrightarrow \text{Fourier Transform} \Longrightarrow \text{Power Spectral Density} \Longrightarrow \text{Classification of Noise Color}$$

The classification of noise color becomes important when we need to have an estimate of future data. In other words, given the trends in our current data we can use it to estimate future data. This is done as follows,

$$\text{Noise Color} \Longrightarrow \text{Power Spectral Density} \Longrightarrow \text{Fourier Transform} \Longrightarrow \text{Curve} \Longrightarrow \text{Generate Data From Curve}$$

The power of a noise signal is detected at a certain frequency. Then a plot of the log(power) vs. the log(frequency) can be constructed, and the slope of the line gives the beta value. Following a backward thought process, one can produce a certain color of noise by creating frequency components which have a value generated by a Gaussian distribution and then scaling by the appropriate beta power of frequency.

A general method to characterize and model noise is explained below.

1. DATA

Data is what the signal transmits. The signal is dependent on what you are measuring or modeling. The data can be collected, for example, if you're measuring temperature in a reactor, then your data is the temperature readings from a thermocouple at a certain position in the reactor over a period of time.

2. CURVES

After you collect the data, you plot the data and find a best fit equation, $x(t)$, for that set of data. A math program can be used to find a best fit equation. Microsoft Excel can be used for a simple model and programs such as Polymath or Mathematica can be used for more complex models. A simple equation could be in the form of $Asin(Bt)$. The coefficients A and B can be varied to fit the data. Additional terms can be added for more accurate modeling of complex data. The equation you find can then be used to predict and model future signal noise.

3. FOURIER TRANSFORMS

A Fourier transform is a tool used to convert your data to a function of $\frac{1}{f^{\beta}}$. In this form, the noise can be more easily characterized. Apply a Fourier transform to the curve,$x(t)$, you fit to your data to generate a relation for the power spectral density. The Fourier transform can be performed by a computer depending on the complexity of $x(t)$(or see "Simplifying the Fourier Transform" below). The transform is the integral shown below.

$$X(w) = \int_{-\infty}^{\infty} x(t)e^{-jwt}dt$$

where;

- $x(t)$*is the equation of the curve to fit the data.*
- e^{-jwt} *is the exponential form of writing the relation* $cos(wt) - jsin(wt)$*. The j in the second term here indicates that that term is imaginary.*
- w *is the frequency*

4. POWER SPECTRAL DENSITY

This value is attained by simplifying and squaring the Fourier Transform. Since the series of simplifications needed are fairly complex, the resulting power spectral density equation is displayed below.

$$PSD = \sqrt{\left(\sum x(t)\cos(wt)\right)^2 + \left(\sum x(t)\sin(wt)\right)^2}$$

At this point we attain a numerical value for the PSD at the particular frequency, w. These numerical PSD values can be plotted versus frequency to obtain the PSD chart.

5. CLASSIFICATION OF NOISE

The summation is repeated over and over for different frequencies, ω. A plot of the PSD vs. frequency, w is made with these values. Once this is done a best fit line is applied to the data points which will give you the characterization relation $PSD \propto \frac{1}{f^{\beta}}$. Based on this we can then classify the noise as a color of noise.

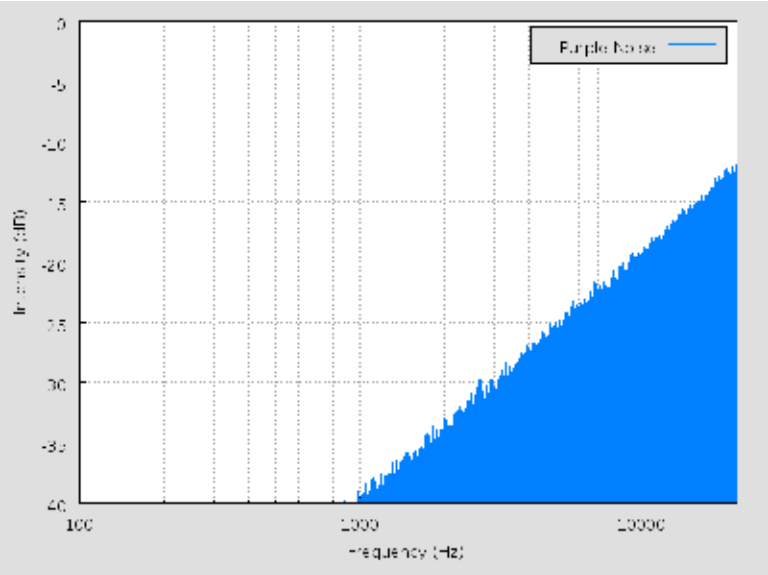


image from en.Wikipedia.org/wiki/Violet_noise#Violet_noise on 12/11/2008

THE REVERSE PROCESS

Knowing how to convert data to a color of noise is only half the problem. What if we know what type of noise is possible and we need data from it for a given process? Knowing the noise color means that we know the power spectral density relation to the frequency. From here onwards we follow the reverse route as that taken to get to the noise color by using the inverse Fourier transform instead.

Inverse Fourier Transform

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(w)e^{jwt}dt$$

where;

- $x(t)$ *is the equation of the curve to produce future the data.*
- e^{jwt} *is the exponential form of writing the relation* $cos(wt) + jsin(wt)$*. The j in the second term here indicates that that term is imaginary.*
- w *is the frequency*

We can either use it to produce a curve or, by making a similar simplification as we did for the Fourier transform, to generate data directly. This reverse process should be trivial to someone who worked through the forward process.

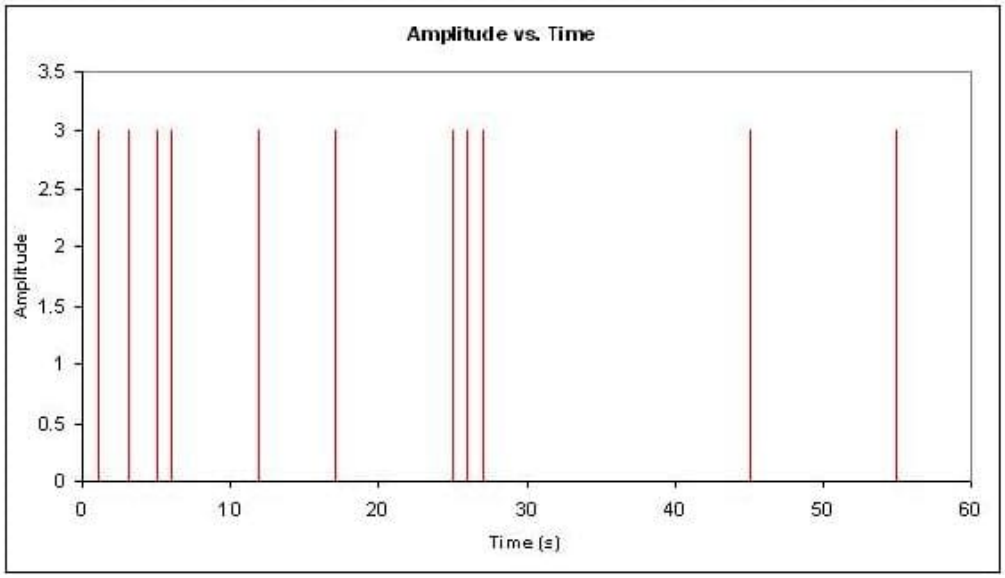
Pops, Snaps & Crackles

The second category of noise is non-frequency based noise. Three examples of this non-frequency noise are pops, snaps and crackles. The study of these noise types is fairly new, and not a lot is known about how to deal with them when they arise in our research or studies. However, one can successfully distinguish and classify these noise types by noticing certain properties which are typical of them.

Pops

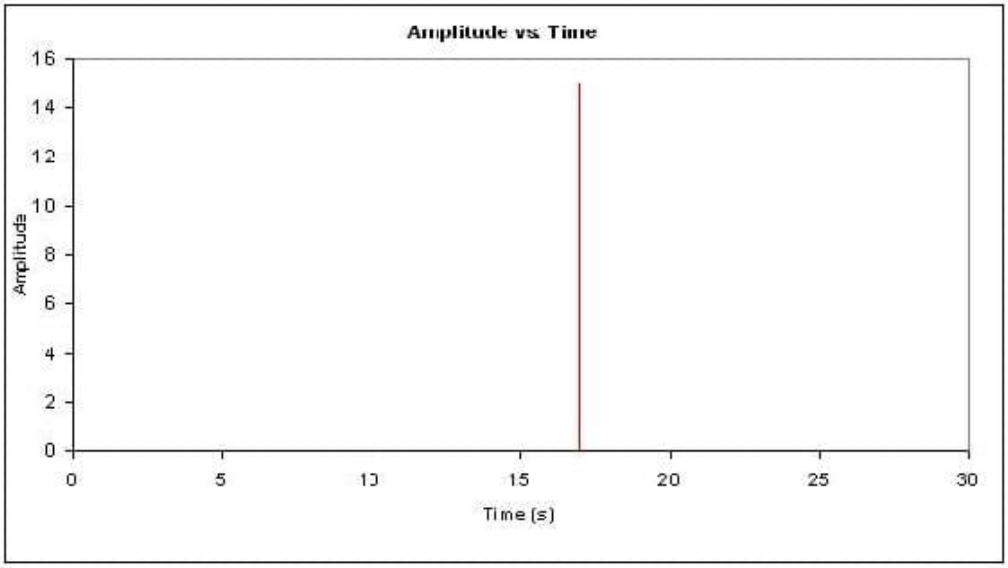
At one end of the extreme of non-frequency noise is what is defined as pops. Pops are infrequent random spikes in noise of approximately the same amplitude. An example of this would be a loose wire that is usually completing the circuit. However, it randomly disconnects for a split second and then recompletes the circuit for normal operation. Chemical engineering processes may inherently have pops that are unpredictable. Safety features can be added to a process, such as pressure relief valves, to

handle the pops when they do occur. The following is an image of what pops might look like in a system:



Snaps

On the other end of the non-frequency noise spectrum there are snaps. Snaps are single independent events that occur only once. When a pencil is bent and it breaks under the force is an example of snapping. Another example would be the snapping of a piece of chalk or exploding a bomb. The “popping” of one’s knuckles is really snapping since it is an independent event, unless you snap all your knuckles all at once or one after the other really fast. Just like pops, snaps are unpredictable, and safety features should be added to the system to handle the snaps. The following is an example of a snap:



Crackles

In between popping and snapping there is crackling. A very common example for crackling is the sound heard coming from a burning piece of wood. Like popping, there is a non-frequency or irregularity in which the crackles occur. In addition, there is also an irregularity of the amplitude of the crackle. In the case of the fire, not only can you not predict when the crackle sound would be heard, but you cannot predict how loud it will be either. Furthermore, there is a universality condition associated with crackling. Regardless of the scale, the similar randomness in repetition and amplitude should be observed.

In dealing with this universality condition the concept of a critical exponent arises. For example, if we are looking at the same crackling effect, S, over a larger period of time the two would have to be equal after scaling the larger one.

$$\langle S\rangle_{\text{small}}(T) = A\langle S\rangle_{\text{large}}\left(\frac{T}{B}\right)$$

for some scaling factors A and B.

where $\langle S\rangle$ is S average.

If the time scale is expanded by a small factor $B = \left(\frac{1}{(1-\delta)}\right)$ then the rescaling of the size will also be small, say $1 + a\delta$.

$$\langle S\rangle(T) = (1 + a\delta) * \langle S\rangle(1 + \delta T)$$

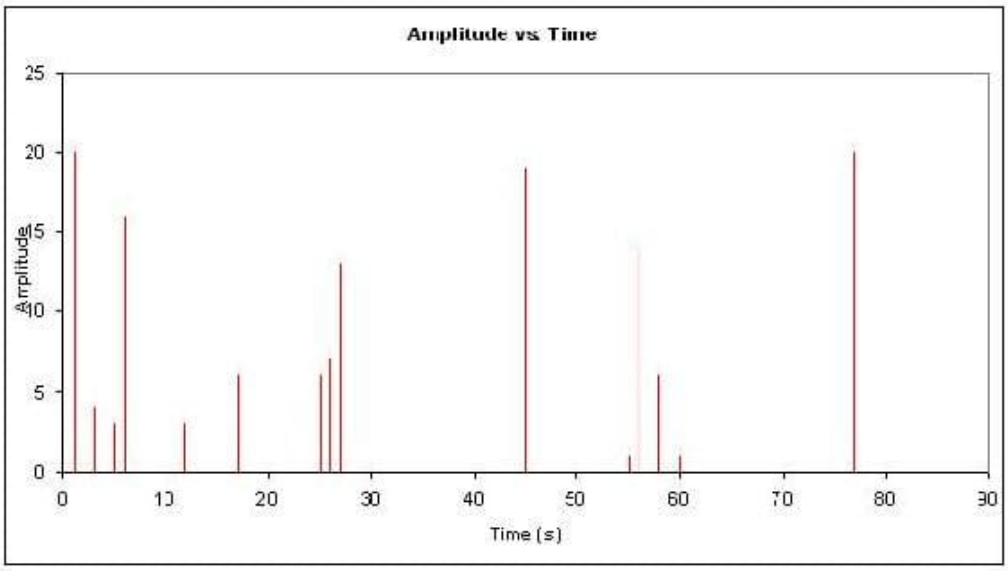
For small δ we have

$$a\langle S\rangle = T\frac{d\langle S\rangle}{dT}$$

Solving gives

$$\langle S\rangle(T) = S_o T^a \tag{2.5.1}$$

The exponent a is called the **critical exponent** and is a universal prediction of a given theory. There are other concepts which are used as a check for universality for times when the critical exponent cannot be used but the most common one is the critical exponent. The following is an example of what a crackle might look like in a system:



Example 2.5.1: Crackle

A lead engineer monitoring the instantaneous flow rate of a coolant used to cool an exothermic reactor gathers data set 'a' over a 30hr period. Later on that week he gathers data set 'b' for the same coolant flow rate. For both sets of data he determined that the way the instantaneous coolant flow rate is reacting to the exothermic reactor is optimal to the reaction process. For his records he would like to represent the noise in the data in a compact form. He wants you to first characterize the noise and then provide any other information, which will generalize the trend in the data. Once you have completed these tasks, report back to him.

Time (hr)	Flow Rate (m^3/s)				
0	0	11	0	21	0
1	0	12	0	22	0
2	0	13	10	23	0
3	0	14	0	24	0
4	0	15	0	25	3
5	2	16	0	26	5.1
6	3.5	17	0	27	0
7	0	18	7.4	28	0
8	0	19	12.8	29	0
9	4	20	0	30	0
10	0				

Data set 'a'

Note that between 0 and 30hr for all intervals of time not represented in this table there is a zero assigned as the instantaneous flow rate.

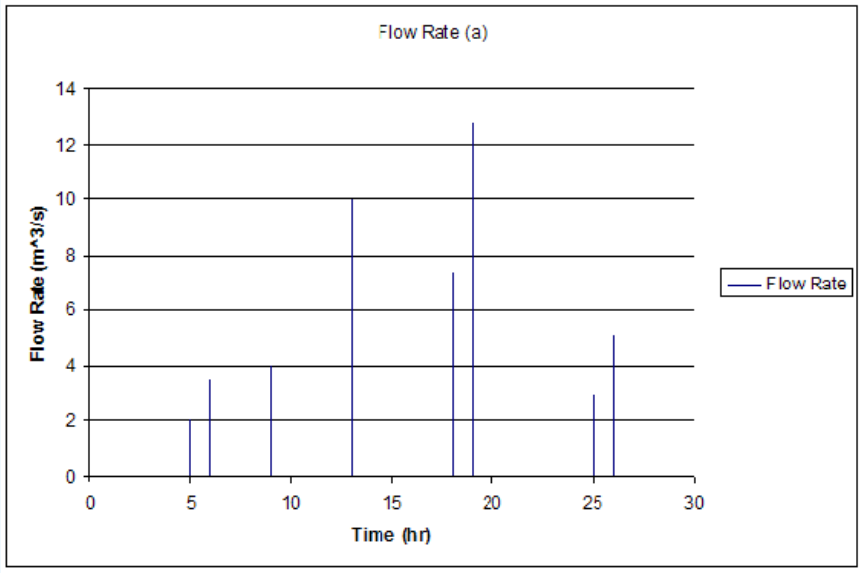
Time (hr)	Flow Rate (m^3/s)				
0	0	16	0	32	0
1	0	17	0	33	0
2	0	18	5.5	34	0
3	0	19	14	35	0
4	0	20	0	36	0
5	3	21	0	37	0
6	4.5	22	0	38	6
7	0	23	0	39	16
8	0	24	0	40	0
9	4	25	3	41	8
10	0	26	7	42	3
11	0	27	16	43	5
12	0	28	0	44	0
13	8	29	0	45	0
14	0	30	0	46	0
15	0	31	5	47	0
				48	0
				49	0
				50	0

Data set 'b'

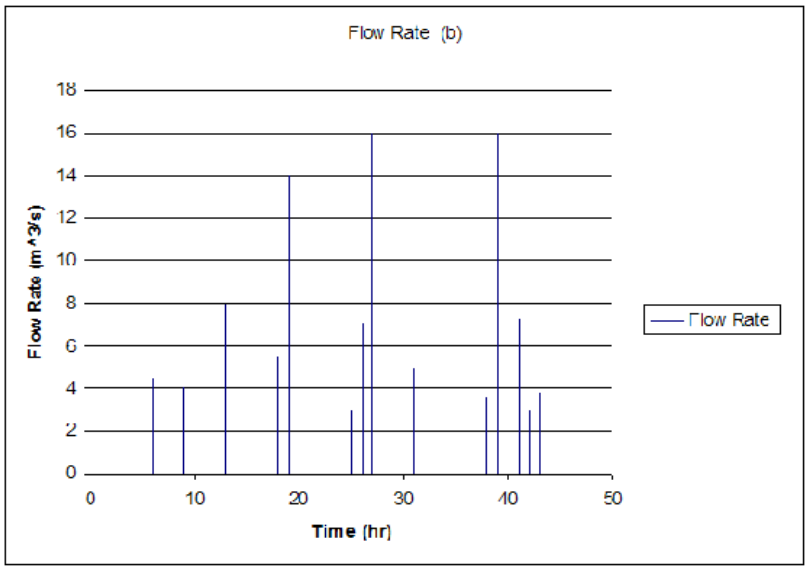
Note that between 0 and 50hr for all intervals of time not represented in this table there is a zero assigned as the instantaneous flow rate.

Solution

Each set of data is plotted, Flow Rate vs. Time. The two graphs show crackling trends.



Graph for data set 'a'



Graph for data set 'b'

Upon observing this, the general noise trend is crackling. To provide a final proof we apply the universality relation (Equation 2.5.1).

We first determine the critical exponent, a, for data set a.

We see that;

~~\[=1.54 \nonumber \]~~

$$S_0 = 2$$

~~this value can be the first value attained or an average of the first 2 or 3 values.~~

$$T = 30$$

Solving for the critical exponent using the relation given above we get;

$$a = -0.08$$

Once again we carry out the same calculation for data set 'b'.

\[=2.12\nonumber \]

$$S_0 = 3$$

this value can be the first value attained or an average of the first 2 or 3 values

$$T = 50$$

Solving for the critical exponent using the relation given above we get;

$$a = -0.08$$

The similarity of the two critical exponent gives further proof that the data the instantaneous flow rate of the coolant over a period of time will show crackling.

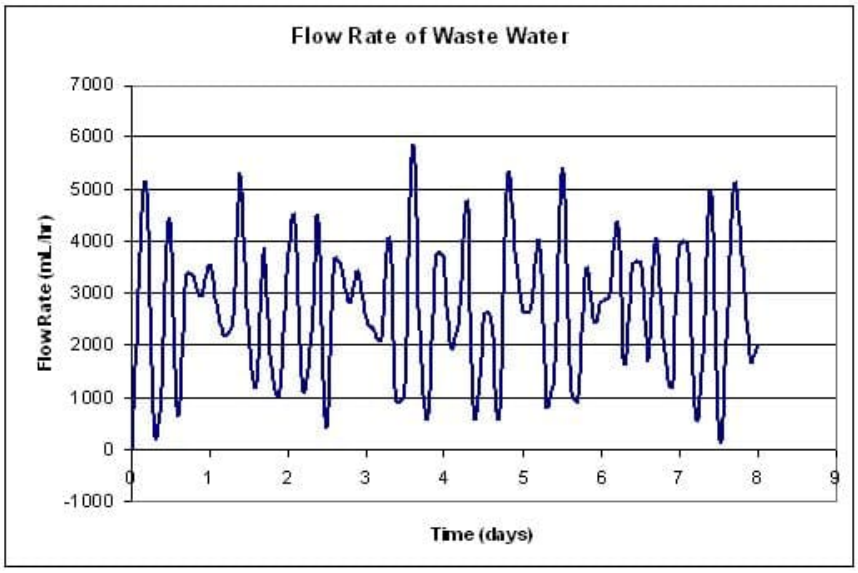
Example 2.5.2: Colors of Noise

A chemical engineer is reading flow rates from flow meter. Every 0.1 day for 8 days a reading was taken and the data is given here: Colors of Noise Example. The data displays the fluctuations from the set flow rate of 3000 liters per hour at a wastewater treatment plant. The specifications for the plant say that the max flow rate is 6000 liters per hour, or the pipes will burst. Also, the flow rate cannot fall below 200 liters per hour, or the system will automatically shut down.

The chemical engineer notices that there were some readings close to these limits. It is the chemical engineer's job to determine if the readings are accurate flow rates or if there is an error with the flow meter. By characterizing the type of noise, the chemical engineer can determine the source of the noise and so take appropriate preventative measures. What type of noise is present and what protective measures should be taken?

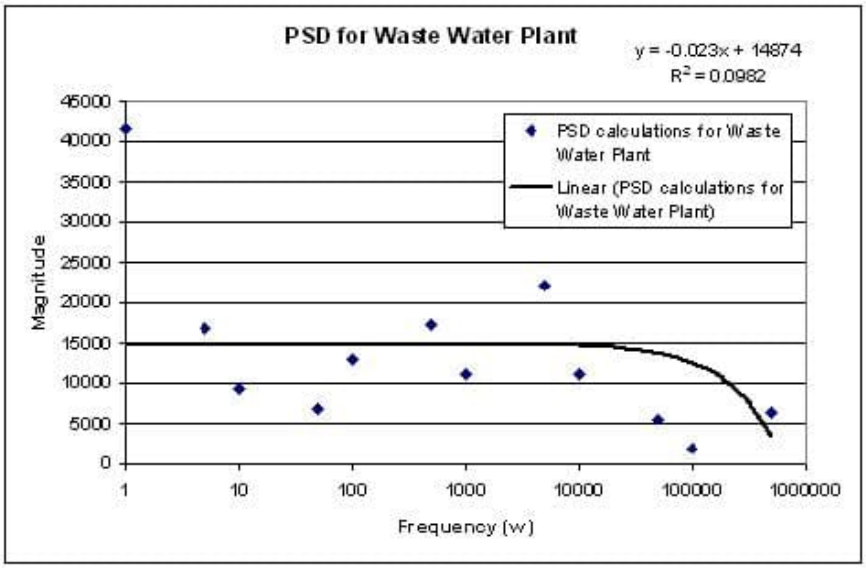
Solution

1) Plot data. From the data presented, a flow rate vs. time chart was graphed to gauge degree of fluctuation in the flow rate.



2) Calculate the power spectral density data. Using the simplified integral derived to calculate the power spectral density, a table was created with the resulting PSD values at varying frequencies. These frequencies were self-defined and were chosen to encompass a broad range. More detailed calculations can be found here: Colors of Noise Example under the worksheet titled "PSD CALC".

3) Plot the power spectral density. The power spectral density for each frequency was then plotted against the frequency, creating the power spectral density plot shown below.



4) Characterize the noise. To determine the β value for the data, a linear trend line was taken on the data. This trend line can be seen in the power spectral density plot above. The slope of this trend line is the β value for the data. In this case, the β=0.023. Since this value is not that of white noise (β=0) nor that of pink noise (β=1), we can say that this noise is somewhere between white and pink noise.

5) Determine the source of noise.

There are two possible major sources of noise. They are the liquid motion in the pipe and noise in the flow meter caused by itself or outside sources. Since it is found earlier that β=0.023, the source of noise is probably from the flow meter. It is *not* from the motion of liquid in the pipe because liquid motion tends to produce brown noise (β=2).

6) Protective measures that can be taken. Knowing this, a correction can be made on the calculations to determine the actual flow rate. A full step by step solution can also be found here: Colors of Noise Example under the worksheet titled "Solution".

Summary

Noise is all around us and is an unavoidable characteristic of monitoring signals. When noise is frequency based, like the colors of noise, it can be characterized and modeled using mathematical tools. When noise is not frequency based, like pops, snaps, and crackles, it can be recognized, but ways to model it are still being devised. Characterizing and modeling noise is important for chemical engineers to accurately handle their data, control processes, and predict future trends.

References

- Kosko, Bart. "White Noise Ain't So White", Noise. ISBN 0670034959
- Ziemer, Rodger E. Elements of Engineering Probability and Statistics, New Jersey: Prentice Hall. ISBN 0024316202
- Papoulis, Athanasios. Probability, Random Variables, and Stochastic Processes, New York: McGraw – Hill Book Company. ISBN 0071199810

- Peebles, Peyton Z. Jr. Probability, Random Variables, and Random Signal Principles, New York: McGraw – Hill, Inc. ISBN 0071181814
- Sethna, James P. Crackling Noise, Nature 410, 242-250 (2001). Also available here.
- en.Wikipedia.org

Contributors and Attributions

- Authors: Danesh Deonarain, Georgina Mang, Teresa Misiti, Carolyn Ehrenberger
- Date Presented: September 08, 2006 Date Revised: September 19, 2006

This page titled 2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

2.6: Numerical ODE solving in Excel- Euler's method, Runge Kutta, Dead time in ODE solving

Authors: (Presented: 9/8/06 /Date Revised: 9/19/06) Aaron Bennick, Bradley Anderson, Michael Salciccioli

Stewards: (9/5/07) Sean Gant, Jay Lee, Lance Dehne, Kelly Martin

Introduction

This article focuses on the modeling of ordinary differential equations (ODEs) of the form:

$$\frac{dy}{dx} = f(x, y)$$

In creating a model, a new value y_{i+1} is generated using the old or initial value y_i, the slope estimate φ, and the step size h. This general formula can be applied in a stepwise fashion to model the solution. All stepwise models will take the following general form:

$$y_{i+1} = y_i + \phi h$$

The modeling methods discussed in this article are Euler's method and the Runge-Kutta methods. The difference between the two methods is the way in which the slope φ is estimated.

Choosing The Right Model and Step Size

The proper numerical modeling method heavily depends on the situation, the available resources, and the desired accuracy of the result. If only a quick estimate of a differential equation is required, the Euler method may provide the simplest solution. If much higher accuracy is required, a fifth-order Runge-Kutta method may be used. Engineers today, with the aid of computers and excel, should be capable of quickly and accurately estimating the solution to ODEs using higher-order Runge-Kutta methods.

In all numerical models, as the step size is decreased, the accuracy of the model is increased. The tradeoff here is that smaller step sizes require more computation and therefore increase the amount of time to obtain a solution. A balance between desired accuracy and time required for producing an answer can be achieved by selecting an appropriate step size. One suggested algorithm for selecting a suitable step size is to produce models using two different methods (possibly a second and third order Runge-Kutta). The steps size can then be systematically cut in half until the difference between both models is acceptably small (effectively creating an error tolerance).

Euler's Method

Euler's method is a simple one-step method used for solving ODEs. In Euler's method, the slope, φ, is estimated in the most basic manner by using the first derivative at x_i. This gives a direct estimate, and Euler's method takes the form of

$$y_{i+1} = y_i + f(x_i, y_i)\, h$$

For demonstration, we will use the basic differential equation $\frac{dy}{dx} = 3x^2 + 2x + 1$ with the initial condition y(0) = 1. If a step size, h, is taken to be 0.5 over the interval 0 to 2, the solutions can be calculated as follows:

x	y_Euler	y_Actual	Error
y(0)	1	1	0
y(0.5)	1 + [3(0)^2 + 2(0) + 1](0.5) = 1.5	1.875	0.375
y(1)	1.5 + [3(0.5)^2 + 2(0.5) + 1](0.5) = 2.875	4	1.125

x	y_Euler	y_Actual	Error
y(1.5)	2.875 + [3(1)^2 + 2(1) + 1](0.5) = 5.875	8.125	3.25
y(2.0)	5.875 + [3(1.5)^2 + 2(1.5) + 1](0.5) = 11.25	15	3.75

The y_actual values in this table were calculated by directly integrating the differential equation, giving the exact solution as:

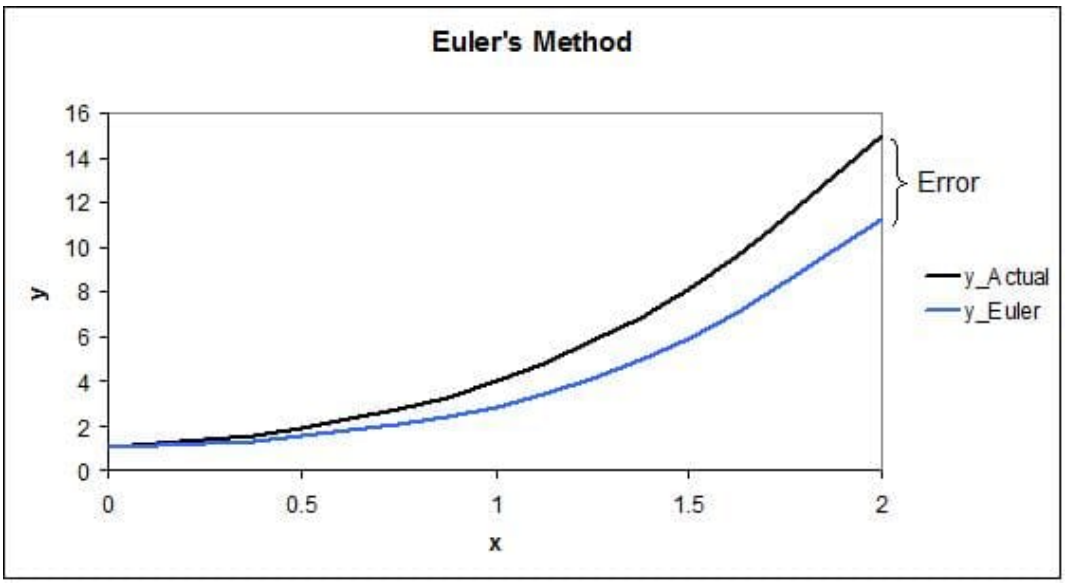
$$y = x^3 + x^2 + x + 1$$

Use of Microsoft Excel

Calculating an ODE solution by hand with Euler's method can be a very tedious process. Fortunately, this process is greatly simplified through the use of Microsoft Excel. Creating a spreadsheet similar to the one above, where the x values are specified and the y_Euler values are recursively calculated from the previous value, makes the calculation rather simple. Any step size and interval can be used.

Drawbacks of Euler's Method

There are several drawbacks to using Euler's method for solving ODEs that must be kept in mind. First of all, this method does not work well on stiff ODEs. A stiff ODE is a differential equation whose solutions are numerically unstable when solved with certain numerical methods. For stiff equations - which are frequently encountered in modeling chemical kinetics - explicit methods like Euler's are usually quite inefficient because the region of stability is so small that the step size must be extremely small to get any accuracy. In a case like this, an implicit method, such as the backwards Euler method, yields a more accurate solution. These implicit methods require more work per step, but the stability region is larger. This allows for a larger step size, making the overall process more efficient than an explicit method. A second drawback to using Euler's Method is that error is introduced into the solution. The error associated with the simple example above is shown in the last column. This error can be seen visually in the graph below.



It can be seen that the two values are identical at the initial condition of y(0)=1, and then the error increases as the x value increases and the error propagates through the solution to x = 2. The error can be decreased by choosing a smaller step size, which can be done quite easily in Excel, or by opting to solve the ODE with the more accurate Runge-Kutta method.

Runge-Kutta Methods

The **Runge-Kutta method** for modeling differential equations builds upon the Euler method to achieve a greater accuracy. Multiple derivative estimates are made and, depending on the specific form of the model, are combined in a weighted average

over the step interval. The order of the Runge-Kutta method can range from second to higher, depending on the amount of derivative estimates made. The second-order Runge-Kutta method labeled Heun's technique estimates derivatives by averaging endpoint measurements of the step size along a function. This averaged value is used as the slope estimate for x_{i+1}. Third and higher power Runge-Kutta methods make mid-point derivative estimations, and deliver a weighted average for the end point derivative at x_{i+1}. As the Runge-Kutta order increases, so does the accuracy of the model.

The general form of the Runge-Kutta method is

$$y_{i+1} = y_i + \phi\left(x_i, y_i, h\right) h$$

Where $\varphi(x_i, y_i, h)$ now represents a weighted average slope over the interval h.

$$\phi\left(x_i, y_i, h\right) = a_1 k_1 + a_2 k_2 + ... + a_n K_n$$

where the *a*'s are constants and the *k*'s are

$$k_1 = f(x_i, y_i)$$

$$k_2 = f(x_i + p_1 h, y_i + q_{11} k_1 h)$$

$$k_3 = f(x_i + p_2 h, y_i + q_{21} k_1 h + q_{22} k_2 h)$$

$$k_n = f(x_i + p_{n-1} h, y_i + q_{n-1,1} k_1 h + q_{n-1,2} k_2 h + ... + q_{n-1,n-1} k_{n-1} h)$$

The constants *a*, *p*, and *q* are solved for with the use of Taylor series expansions once *n* is specified (see bottom of page for derivation). The resulting set of equations have one or more degrees of freedom. This means that for every order of Runge-Kutta method, there is a family of methods. Below are some of the more common Runge-Kutta choices.

Second-Order Runge-Kutta Methods (*n* = 2)

Every second order method described here will produce exactly the same result if the modeled differential equation is constant, linear, or quadratic. Because this is typically not the case, and the differential equation is often more complicated, one method may be more suitable than another.

Heun's Technique

The second-order Runge-Kutta method with one iteration of the slope estimate , also known as Heun's technique, sets the constants

$$a_1 = a_2 = \tfrac{1}{2}$$

and

$$p_1 = q_{11} = 1.$$

Huen determined that defining a_1 and a_2 as $1/2$ will take the average of the slopes of the tangent lines at either end of the desired interval, accounting for the concavity of the function, creating a more accurate result. When substituted into the general form, we find

$$y_{i+1} = y_i + \left(\frac{1}{2} k_1 + \frac{1}{2} k_2\right) h$$

with k_1 and k_2 defined as

$$k_1 = f\left(x_i, y_i\right)$$

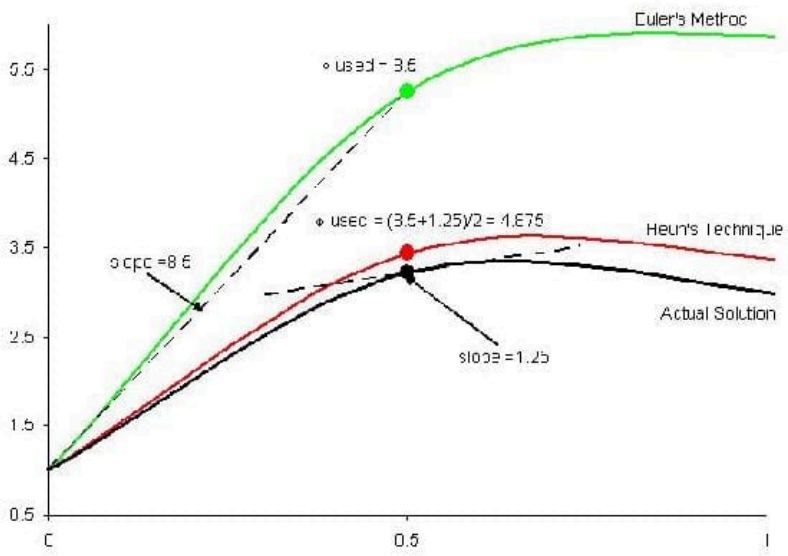
$$k_2 = f\left(x_i + h, y_i + h k_1\right)$$

For demonstration of this second-order Runge-Kutta method, we will use the same basic differential equation $\frac{dy}{dx} = 3x^2 + 2x + 1$ with the initial condition y(0) = 1. If a step size, h, is taken to be 0.5 over the interval 0 to 2, the solutions can be calculated as follows:

Figure 1: second-order Runge-Kutta

x	k_1	k_2	y_Heun	y_Actual	Error
y(0)	3(0)^2 + 2(0) + 1 = 1	3(0.5)^2 + 2(0.5) + 1 = 2.75	1	1	0
y(0.5)	3(0.5)^2 + 2(0.5) + 1 = 2.75	3(1)^2 + 2(1) + 1 = 6	1 + [0.5(1) + 0.5(2.75)](0.5) = 1.9375	1.875	-0.0625
y(1)	3(1)^2 + 2(1) + 1 = 6	3(1.5)^2 + 2(1.5) + 1 = 10.75	2.375 + [0.5(2.75) + 0.5(6)](0.5) = 4.125	4	-0.125
y(1.5)	3(1.5)^2 + 2(1.5) + 1 = 10.75	3(2)^2 + 2(2) + 1 = 17	5.375 + [0.5(6) + 0.5(10.75)](0.5) = 8.3125	8.125	-0.1875
y(2.0)	3(2)^2 + 2(2) + 1 = 17	3(2.5)^2 + 2(2.5) + 1 = 24.75	10.75 + [0.5(10.75) + 0.5(17)](0.5) = 15.25	15	-0.25

When compared to the Euler method demonstration above, it can be seen that the second-order Runge-Kutta Heun's Technique requires a significant increase in effort in order to produce values, but also produces a significant reduction in error. Following Runge-Kutta methods can be worked through a similar manner, adding columns for additional *k* values. Below is a graphical description of how slope is estimated using both Euler's method, and Heun's technique. Observe the increase in accuracy when an average slope across an interval of 0.5 is used instead of just an initial estimate.



Improved Polygon Method

Using the improved polygon method, a_2 is taken to be 1, a_1 as 0, and therefore $p_1 = q_{11} = \frac{1}{2}$. The general form then becomes

$$y_{i+1} = y_i + k_2 h$$

with k_1 and k_2 defined as

$$k_1 = f(x_i, y_i)$$

$$k_2 = f(x_i + \frac{1}{2}h, y_i + \frac{1}{2}hk_1)$$

Ralston's Method

The Ralston method takes a_2 to be $\frac{2}{3}$. Therefore $a_1 = \frac{1}{3}$ and $p_1 = q_{11} = \frac{3}{4}$. It has been determined by Ralston (1962) and Ralston and Rabinowitz (1978) that defining a_2 as $\frac{2}{3}$ will minimize truncation error in second-order Runge-Kutta methods. The general form becomes

$$y_{i+1} = y_i + (\frac{1}{3}k_1 + \frac{2}{3}k_2)h$$

with k_1 and k_2 defined as

$$k_1 = f(x_i, y_i)$$

$$k_2 = f(x_i + \frac{3}{4}h, y_i + \frac{3}{4}hk_1)$$

Third-Order Runge-Kutta Methods (n = 3)

The third-order Runge-Kutta methods, when derived, produce a family of equations to solve for constants with two degrees of freedom. This means an even more variable family of third-order Runge-Kutta methods can be produced. A commonly used general third-order form is

$$y_{i+1} = y_i + [\frac{1}{6}(k_1 + 4k_2 + k_3)]h$$

with

$$k_1 = f(x_i, y_i)$$

$$k_2 = f(x_i + \frac{1}{2}h, y_i + \frac{1}{2}hk_1)$$

$$k_3 = f(x_i + h, y_i - hk_1 + 2hk_2)$$

Fourth-Order Runge-Kutta Methods (n = 4)

The family of fourth-order Runge-Kutta methods have three degrees of freedmon and therefore infinite variability just as the second and third order methods do. The fourth-order versions are most favored among all the Runge-Kutta methods. What is know as the *classical fourth-order Runge-Kutta method* is

$$y_{i+1} = y_i + [\frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)]h$$

with

$$k_1 = f(x_i, y_i)$$

$$k_2 = f(x_i + \frac{1}{2}h, y_i + \frac{1}{2}hk_1)$$

$$k_3 = f(x_i + \frac{1}{2}h, y_i + \frac{1}{2}hk_2)$$

$$k_4 = f(x_i + h, y_i + hk_3)$$

Fifth-Order Runge-Kutta Methods (*n* = 5)

Where very accurate results are required, the fifth-order Runge-Kutta *Butcher*'s (1964) *fifth-order RK method* should be employed:

$$y_{i+1} = y_i + [\frac{1}{90}(7k_1 + 32k_3 + 12k_4 + 32k_5 + 7k_6)]h$$

with

$$k_1 = f(x_i, y_i)$$

$$k_2 = f(x_i + \frac{1}{4}h, y_i + \frac{1}{4}hk_1)$$

$$k_3 = f(x_i + \frac{1}{4}h, y_i + \frac{1}{8}hk_1 + \frac{1}{8}hk_2)$$

$$k_4 = f(x_i + \frac{1}{2}h, y_i - \frac{1}{2}hk_2 + hk_3)$$

$$k_5 = f(x_i + \frac{3}{4}h, y_i - \frac{3}{16}hk_1 + \frac{9}{16}hk_4)$$

$$k_6 = f(x_i + h, y_i - \frac{3}{7}hk_1 + \frac{2}{7}hk_2 + \frac{12}{7}hk_3 - \frac{12}{7}hk_4 + \frac{8}{7}hk_5)$$

The integration of *k*'s within other *k* values suggests the use of a spreadsheet. As with all Runge-Kutta methods, the calculation of values for the fifth-order version would be greatly assisted through the use of Microsoft Excel.

Interactive Model Comparison in Excel

To give a better understanding of the impact between different Runge-Kutta methods, as well as the impact of step size, the interactive excel sheet below will allow you to enter the step size h into a set of models and observe how the models contour to match an example differential equation solution.

Observe the relationship between model type, step size, and relative error.

Dead Time

Dead time, or delay differential equation, occurs when there is a delay or lag in the process of a real life function that is being modeled. This means that the numerical model is not accurate until the delay is over. This concept can come into play for the start up of a reaction process. For example, at the start of a reaction in a CSTR (Continuous Stirred Tank Reactor), there will be reagents at the top of the reactor that have started the reaction, but it will take a given time for these reactant/products to be discharged from the reactor. If one was modeling the concentration of reagents vs. time, time t=0 would have started when the tank was filled, but the concentrations being read would not follow a standard model equation until the residence time was completed and the reactor was in continuous operational mode. The dead time is the time it would take for the readings to start meeting a theoretical equation, or the time it takes for the reactor to be cleared once of the original reagents. Dead time can be determined experimentally and then inserted into modeling equations.

The solution of such a delay differential equation becomes problematic because in order to solve the equation, information from past times (during the delay) is needed in addition to the current time. For example if t_o is the lag time for the given scenario, then the value to use in the ODE becomes (t-t_o) instead of t. When modeling this in excel, (x-t_o) is substituted in for the x value. With this substitution, Euler's method can be used again in the same way to approximate the solution. This substitution will be carried into the differential equation so that the new dead time solution can be approximated using Euler's method. To model the reactor before the initial deadtime is completed, piece-wise functions are often used. This way for the dead time, a given model is used not characteristic of the reactor at normal operating conditions, then once the dead time is completed the modeling equation is taken into effect.

The link below will help to show how to include dead time in a numerical method approximation such as Euler's method. As seen in the excel file, the dead time that is specified by the user in the yellow box will change the delay in the model. The more dead time, the further shifted from the theoretical equation the new model is. To take dead time into account in excel, the x value is simply substituted out for (x-t) where t is equal to the dead time. If you look at the equations entered in the Y cells, you will see that the x value inserted into the differential equation is (x-t), where t is the user specified dead time. It can be seen through this example spreadsheet that the effect of dead time is a simple horizontal shift in the model equation.

Dead Time Interactive Spreadsheet

Error

There are two types of error associated with solving ODEs using stepwise approximation methods in Excel. These errors are also present using other methods or computer programs to solve ODEs. The first, discretization, is the result of the estimated y value that is inherent of using a numerical method to approximate a solution. Discretization errors, also called truncation, occur proportionately over a single step size. Truncation error will propagate over extended results because the approximation from previous steps is used to approximate the next. In essence, a copy of a copy is being made. The accuracy of the next point is a direct result of the accuracy of the previous. Just as the quality of a second copy is dependant on the quality of the first. This error can be reduced by reducing the step size.

Please see ODE model comparison interactive spreadsheet to better learn how step sizes can influence error.

The second types of errors are rounding errors. These errors are dependent on the computer's capacity for retaining significant digits. The more significant digits that a computer can hold, the smaller the rounding error will be.

Estimating Error in Euler's Method

To mathematically represent the error associated with Euler's method, it is first helpful to make a comparison to an infinite Taylor series expansion of the term y_{i+1}. The Taylor series expansion of this term is

$$y_{i+1} = y_i + f(x_i, y_i)\,\hbar + f'(x_i, y_i)\,\frac{h^2}{2} + f''(x_i, y_i)\,\frac{h^3}{3} + \ldots + f^n(x_i, y_i)\,\frac{h^n}{!n}$$

When this expansion is compared to the general form of Euler's method it can be seen that Euler's method lacks every term

beyond $f(x_i, y_i)h$. These missing terms, the difference between the Euler approximation and an infinite Taylor series (taken to be the true solution), is the error in the Euler approximation. Mathematically respresenting the error in higher order Runge-kutta methods is done in a similar fashion.

Estimating and Minimizing Error in Runge Kutta Method

Uniform time steps are good for some cases but not always. Sometimes we deal with problems where varying time steps makes sense. When should you change the step size? If we have an nth order scheme and and (n+1)th order scheme, we can take the difference between these two to be the error in the scheme, and make the step size smaller if we prefer a smaller error, or larger if we can tolerate a larger error. This is fairly simple with Runge Kutta, because we can take a fifth order method and a fourth order method using the same k's. Only a little extra work at each step.

Another way of estimating error in the Runge-Kutta method is to reverse directions at each step of the advancing solution and recompute the previous ordinate. By considering the difference between the newly computed previous ordinate and the originally computed value, you can determine an estimate for the truncation error incurred in advancing the solution over that step. This is a bit more tedious, but does give a good estimate of truncation error.

Euler's Method for Systems of ODEs

In chemical engineering and other related fields, having a method for solving a differential equation is simply not enough. Many real world problems require simultaneously solving systems of ODEs. For problems like these, any of the numerical methods described in this article will still work. The only difference is that for n ODEs, n initial values of y are needed for the initial x value. Then the numerical method of choice is just applied to every equation in each step before going on to the next. This further complicates the step-by-step problem solving methodology, and would require the use of Excel in nearly every application. The ODE solved with Euler's method as an example before is now expanded to include a system of two ODEs below:

$$\frac{dy_1}{dx} = 3x^2 + 2x + 1, y_1(0) = 1$$

$$\frac{dy_2}{dx} = 4y_1 + x, y_2(0) = 2$$

Step size is again 0.5, over an interval 0-2.

x	y_1Euler	y_2Euler	y_1Actual	y_2Actual	y_1 Error	y_2 Error
y(0)	1	2	1	2	0	0
y(0.5)	1 + [3(0.5)^2 + 2(0.5) + 1](0.5) = 2.375	2 + [4(1) + 0.5](0.5) = 4.25	1.875	5.875	-0.5	1.625
y(1)	2.375 + [3(1)^2 + 2(1) + 1](0.5) = 5.375	4.25 + [4(2.375) + 1](0.5) = 9.5	4	18.5	-1.375	9
y(1.5)	5.375 + [3(1.5)^2 + 2(1.5) + 1](0.5) = 10.75	9.5 + [4(5.375) + 1.5](0.5) = 21	8.125	51.875	-2.625	30.875
y(2.0)	10.75 + [3(2)^2 + 2(2) + 1](0.5) = 19.25	21 + [4(10.75) + 2](0.5) = 43.5	15	124	-4.25	80.5

It should be observed that there is more error in the Euler approximation of the second ODE solution. This is because the equation also has y_1 in it. So there is the error introduced by using the Euler approximation to solve the 2nd ODE, as well as the error from the Euler approximation used to find y_1 in the 1st ODE in the same step!

Exact solutions were again obtained from directly integrating the ODEs: $y_1 = x^3 + x^2 + x + 1$ and $y_2 = 4y_1x + \frac{x^2}{2} + 2$

Example 1: Reactor Design

The elementary liquid-phase reaction A --> B is to be carried out in an isothermal, isobaric PFR at 300 degrees C. The feed enters at a concentration of 0.25 mol/L and at a rate of 3 mol/min. The reaction constant is known experimentally to be 0.01 min^{-1} at this temperature. You have been given the task of building a reactor that will be used to carry out this reaction. Using Euler's method with a step size of 0.05, determine how large the reactor must be if a conversion of 80% is desired.

Simplified Design Equation:

$$\frac{dV}{dX} = \frac{F_{A0}}{kC_{A0}(1-X)}$$

Lumping the given flow, concentration, and reaction constant together gives:

$$\frac{dV}{dX} = 1200\frac{1}{1-X}$$

Since no volume is required for a conversion of zero, the initial condition needed is V(0)=0.

Now the information simply has to be entered into Excel. A column for the conversion, X, going from 0 to 0.8 in 0.05 increments is used for the step size, and the first value, V(0), is known to be zero. To get the volume, simply add the previous volume to the constants multiplied by the step size and 1/(1-X), or:

$$V_{i+1} = V_i + 1200 * 0.05 * \frac{1}{1-X}$$

Copying this formula down the column to the final conversion value of 0.8, gives the results shown in the table below:

X	V
V(0)	0
V(0.05)	63
V(0.1)	130
V(0.15)	200
V(0.2)	275
V(0.25)	355
V(0.3)	441
V(0.35)	533
V(0.4)	633
V(0.45)	743
V(0.5)	863
V(0.55)	996
V(0.6)	1146
V(0.65)	1317
V(0.7)	1517
V(0.75)	1757
V(0.8)	2057

The final reactor volume obtained is approximately 2057 L. This compares reasonably well to the exact value of 1931 L. The Excel file used to obtain this solution, along with the exact solution, can be downloaded below.

Example 1

Second-order Runge-Kutta derivation

The following example will take you step by step through the derivation of the second-order Runge-Kutta methods. Setting $n = 2$ results in a general form of

$$y_{i+1} = y_i + (a_1 k_1 + a_2 k_2) h$$

with k_1 and k_2 defined as

$$k_1 = f(x_i, y_i)$$

$$k_2 = f(x_i + p_1 h, y_i + q_{11} k_1 h)$$

The constants in the general form must be defined. To do this we will employ a second-order Taylor series expansion for y_{i+1} in terms of y_i and $f(x_i, y_i)$. This Taylor series is

$$y_{i+1} = y_i + f(x_i, y_i) h + f'(x_i, y_i) \frac{h^2}{2}$$

Expanding $f'(x_i, y_i)$ with the chain rule, and substituting it back into the previous Taylor series expansion gives

$$y_{i+1} = y_i + f(x_i, y_i) h + \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial x} \right) \frac{h^2}{2}$$

The next step is to apply a Taylor series expansion to the k_2 equation. The applied Taylor series expansion rule is

$$g(x + r, y + s) = g(x, y) + r \frac{\partial g}{\partial x} + s \frac{\partial g}{\partial y} + \ldots$$

and the resulting equation is

$$f(x_i + p_1 h, y_i + q_{11} k_1 h) = f(x_i, y_i) + p_1 h \frac{\partial f}{\partial x} + q_{11} k_1 h \frac{\partial f}{\partial y} + O(h^2)$$

where $O(h^2)$ is a measure of the truncation error between model and true solution.

When this Taylor series expansion result of the k_2 equation, along with the k_1 equation, is substituted into the general, and a grouping of like terms is performed, the following results:

$$y_{i+1} = y_i \quad [a_1 f(x_i, y_i) | a_2 f(x_i, y_i)] h \left| \left[a_3 p_1^{\partial f} \right| a_s q_{11} f(x_i, y_i)_{\partial y}^{\partial f} \right| h^2 | O(h^3)$$

Setting this result equal to the substituted general form will allow us to determine that, for these two equations to be equivalent, the following relationships between constants must be true.

$$a_1 + a_2 = 1$$

$$a_2 p_1 = \frac{1}{2}$$

$$a_2 q_{11} = \frac{1}{2}$$

It should be noticed that these three equations, relating necessary constants, have four unknowns. This means that to solve for three constants, one must first be chosen. This results in a family of possible second-order Runge-Kutta methods.

Sage's Corner

Example of usage of Euler's Method:

video.google.com/googleplayer.swf?docId=1095449792523736442

Narrated example of using the Runge-Kutta Method:

video.google.com/googleplayer.swf?docId=-2281777106160743750

Unnarrated example of Using the Runge-Kutta Method:

File:Numerical Solving in Excel, Unnarrated.ppt

References

- "Delay Differential Equation", Wolfram MathWorld, Online: September 8, 2006. Available http://mathworld.wolfram.com/DelayDifferentialEquation.html
- Chapra, Steven C. and Canale, Raymond P. "Numerical Methods for Engineers", New York: McGraw-Hill.
- Franklin, Gene F. et al. "Feedback Control of Dynamic Systems", Addison-Wesley Publishing Company.
- R. England. "Error Estimates for Runge-Kutta Type Solutions to Systems of Ordinary Differential Equations", Research and Development Department, Pressed Steel Fisher Ltd., Cowley, Oxford, UK. October 1968.
- Call, Dickson H. and Reeves, Roy F. "Error Estimation in Runge Kutta Procedures", ACM, New York, NY, USA. September 1958.

This page titled 2.6: Numerical ODE solving in Excel- Euler's method, Runge Kutta, Dead time in ODE solving is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

2.7: Solving ODEs with Mathematica- How to find numerical and analytical solutions to ODEs with Mathematica

Authors: Matthew Baumgartner, Olyvia Dean, Viral Patel, Joel Schweitzer, and Eric Van Beek

Stewards: Brian Hickner, Lennard Gan, Addison Heather, Monique Hutcherson

Date Released: September 6, 2006 /Date Revised: September 8, 2007

Introduction

Mathematica is an advanced mathematics solution program created by Wolfram Research, Inc. One of the most powerful software packages of its kind, Mathematica is able to easily and conveniently solve complicated mathematical formulae, including differential equations. This article focuses on the modeling of first and higher order Ordinary Differential Equations (ODE) in the following forms:

$$\frac{dy}{dx} = f(x,y) \qquad \text{(Basic ODE)}$$

$$\frac{dy^z}{d^z x} = f(x,y) \qquad \text{(Higher Order ODE)}$$

Like all software, Mathematica uses a specific language in order to perform calculations. The names of all functions must be capitalized- this applies to both mathematical functions (such as **Sin** and **Cos**) and built-in functions (such as **Plot** and **DSolve**). For this reason it is common for users to write their own functions in minimized letters. This decreases the chance of overriding or redefining a Mathematica function. Square brackets always follow function calls, and the function's parameters are always enclosed in curved brackets. For example, if a user wished to plot *sin*(*x*) from *x* = 0 to *x* = 1, they would type: **Plot[Sin[x],{x,0,1}]**. The user must type "Shift"+"Enter" to input the function. Typing only "Enter" will add a line to the formula. . For full PC keyboards, the “Enter” key on the far right is the equivalent of “Shift”+”Enter”. If the user does not want to see the output line, it can be suppressed by typing a semi-colon at the end of the expression. Mathematica also recognizes common mathematical constants such as pi (**Pi**), Euler's constant (**E**), and imaginary numbers (**I**). Note that these constants must also be capitalized when entered.

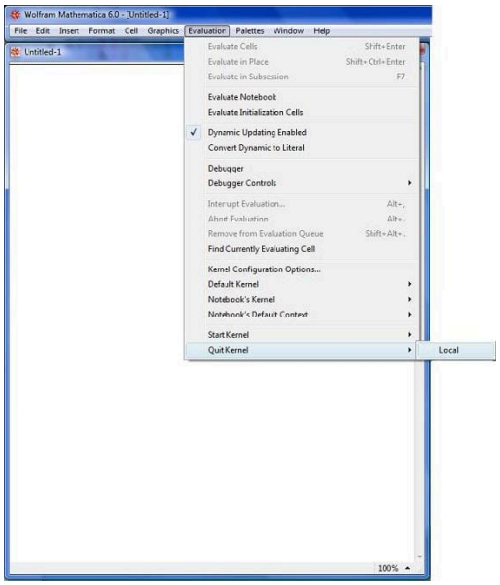
Mathematica features two functions for solving ODEs: **DSolve** and **NDSolve**. **DSolve** is used when the user wishes to find the general function or functions which solve the differential equation, and **NDSolve** is used when the user has an initial condition. The prompts for the functions are quite similar. **Note:** Two equal signs must be used to denote equality in an equation. Using one equal sign assigns a value to a variable.

Example:

f(*x*) = $5x^2$ + 7 This input creates a variable named *f(x)* that has a value of $5x^2$ + 7.

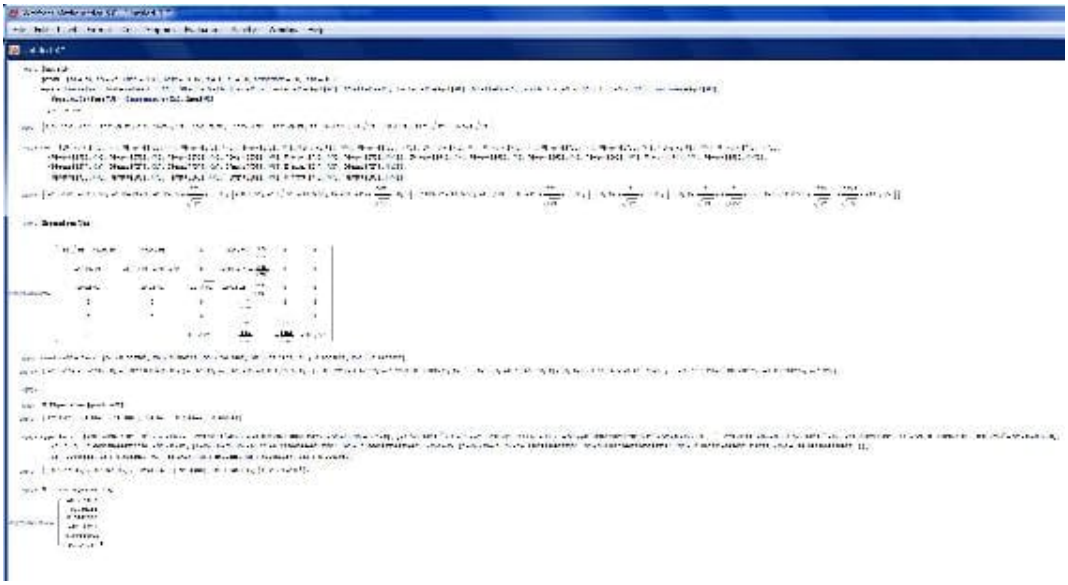
f(*x*) = = $5x^2$ + 7 This input creates a function *f*(*x*) which is defined by $5x^2$ + 7.

Mathematica will sometimes store user created functions or variables into its cache. This causes some (potentially) correct commands to output errors. It is recommended to quit the local kernel and reinitialize the commands after any changes to your code.

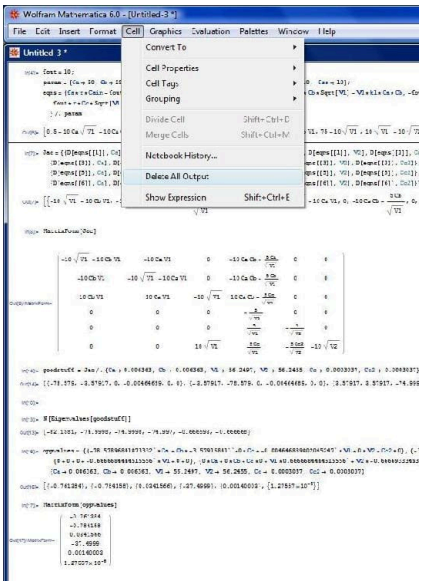


Also, check that every command in the code is being executed. Clearing all the outputs in the cell may be helpful to figure out which commands had not been executed. This can be done by going to the “Cell” option found on the top and choosing the “Delete All Output” Example:

OUTPUT IS NOT CLEARED



CLEARING THE OUTPUTS



CODE WITH CLEARED OUTPUTS

As seen in the examples, it is easier to troubleshoot and debug your program when it looks less confusing. Clearing the extra outputs helps you focus on just the code that you have written.

First Order ODEs

Notation: In the following examples, *eqn* represents the ODE, *y* represents the function being solved for, *i* represents the initial condition, and *x* and *t* are independent variables.

ODEs With Initial Conditions

NDSolve will not display a numerical value, but rather an “Interpolating Function” which can be displayed graphically. The easiest way to display this graph is to assign the solution to a variable (called *Solution* in the example) in the input line, and then use the “Plot” function to display it.

Input for One ODE: Solution **= NDSolve[{**eqn**,**y**[0] ==** i**},**y**,{**x**,**xmin**,**xmax**}]**

Input for Multiple ODEs: Solution **= NDSolve[{**eqn1**,**eqn2**,...,**y1**[0] ==** i1**,**y2**[0] ==** i2**,...},{**y1**,**y2**,...},{**x**,**xmin**,**xmax**}]**

Input for Partial Differential Equations: Solution **= NDSolve[{**eqn**,**y**[0] ==** i**},**y**,{**x**,**xmin**,**xmax**},{**t**,**tmin**,**tmax**}]**

After Mathematica solves the ODE, plot the solution by typing: **Plot[Evaluate[**y**[**x**] /.** Solution**],{**x**,**xmin**,**xmax**}]**, where “/.” is a notation used by Mathematica meaning “following the rule of”. It basically recalls the function stored as the variable in the previous input line. Depending on the needs of the user, the functions “ParametricPlot” or “Plot3D” may also be useful. The notation for these functions is the same as “Plot”.

Here is a simple example of what you would type into Mathematica:

- **Solution = NDSolve[{y'[x]==y[x]*Cos[x+y[x]],y[0]==1},y,{x,0,30}]**

Mathematica will output:

- Output[1]= {{y->InterpolatingFunction[{ { 0.,30.} },<>] } }

To plot this function you would type:

- **Plot[Evaluate[y[x]/.Solution],{x,0,30}]**

Note: Remember to type "Shift"+"Enter" to input the function

ODEs Without Initial Conditions

Input for One ODE: **DSolve[**eqn,y,x**]**

Input for Multiple ODEs: **DSolve[{**eqn1,eqn2**,...},{**y1,y2**,...},**x**]**

Input for a Partial Differential Equation: **DSolve[**eqn,y**,{**x1,x2**,...}]**

Here is a simple example of what you would type into Mathematica:

- **Solution = DSolve[y'[x] ==4*x-2*x*y[x],y[x],x]**

Mathematica will output:

- Output[1]=
  ParseError: invalid Primary (click for details)

  ```
  Callstack:
      at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Contro
  ```

Second Order and Higher ODEs

Input for Higher Order ODEs: Solution **= NDSolve[{**eqn,y**[0] ==** i1**,** y' **[0] ==** i2**,...},**y**,{** x,xmin,xmax**}]**

Plotting the solution: **Plot[Evaluate[{**y**[**x**],** y' **[**x**],...} /.** Solution**],{**x,xmin,xmax**}]**

Here is a simple example of what you would type into Mathematica:

- **Solution = NDSolve[{y**[x] +Sin[y[x]]+y[x]==0,y[0]==1,y'[0]==0,y,{x,0,30}]

Mathematica will output:

- Output[1]= {{y->InterpolatingFunction[{ {0.,30.} },<>] } }

To plot this function you type:

- **Plot[Evaluate[{y[x],y'[x],y**[x]}/.Solution],{x,0,30}]

Mathematica should output a graph.

Algorithms Used by Mathematica

Mathematica uses two main algorithms in order to determine the solution to a differential equation. These algorithms are the Adams method and the Gear method.

The Adams and Gear methods are forms of linear multistep methods. An example of these would be the following:

$$y_{n+1} = y_n + h(b_0 f(t_n, y_n) + b_p f(t_{n-p}, y_{n-p})$$

In the example above, *h* denotes the step size and the coefficients $b_0, \ldots, b_s$ are determined by the method used. Multistep methods are expansions of more familiar single-step methods used to solve differentials (i.e. Euler, Runge-Kutta, Taylor). Each of these methods requires an initial point in order to calculate the following point. Similarily, multistep methods also require initial points in order to solve the ODE. The number of initial points required depends on which method is used to solve the ODE. Multistep methods typically produce less error than the single-step methods because of multiple initial points.

In order to determine what method to use one must first find the stiffness of the function.

Adams Method

Euler, Taylor and Runge-Kutta methods used points close to the solution value to evaluate derivative functions. The Adams-Bashforth method looks at the derivative at old solution values and uses interpolation ideas along with the current solution and derivative to estimate the new solution [4]. In order to solve an ODE using this method, *f*(*t*,*y*) must be continuous and satisfy Lipschitz condition for the y-variable which states [5]:

$$|f(h) - f(0)| <= B|h|^{\beta}$$

for all | *h* | < ε where B and β are independent of *h*, β > 0 and α is an upper bound for all β for which a finite B exists

This is a basic form of the Adams-Bashforth method. Note that two initial points are required for this method.

$$y_{n+1} = y_n + h[\beta_1 f(y_n, t_n) + \beta_2 f(y_{n-1}, t_{n-1})]$$

There is another Adams method that requires three initial points. The method is solved the same way, however the equation varies a little bit and is referred to as the Adams-Moulton Method.

$$y_{n+1} = y_n + h[\beta_0 f(y_{n+1}, t_{n+1}) + \beta_1 f(y_n, t_n) + \beta_2 f(y_{n-1}, t_{n-1})]$$

The coefficients/constraints, β can be solved for using knowledge of ODE's and other math tools. It was stated earlier that $\frac{dy}{dt} = f(t, y)$. We can let *f*(*t*,*y*) = λ*y*, therefore $\frac{dy}{dt} = \lambda y$. We can also let $y_n = \sigma^n y_0$ if there is a constant step size and σ represents a polynomial. Through substitution we find [6]:

$$-\sigma^2 + (\beta_1 h\lambda + 1)\sigma + \beta_2 h\lambda = 0$$

We can expand the quadratic using another math identity and ultimately solve for constraints β_1 and β_2. Another method for solving for coefficients β_1,β_2 is mentioned below:

In order to find the coefficient β_j one must first use polynomial interpolation to find the polynomial *p* of degree *s* – 1 such that:

$$p(t_{n-i}) = f(t_{n-i}, y_{n-i}), \text{for } i = 1, \ldots, s.$$

From this the Lagrange formula for polynomial interpolation yields

$$p(t) = \sum_{j=1}^{s} \frac{(-1)^{s-j} f(t_{n-j}, y_{n-j})}{(j-1)!(s-j)!h^{s-1}} \prod_{\substack{i=1 \\ i \neq j}}^{s} (t - t_{n-i}).$$

Now the polynomial *p* is a locally good approximation of the right-hand side of the differential equation *y*′ = *f*(*t*,*y*) that is to be solved. Now we must consider the equation *y*′ = *p*(*t*) instead. This equation can be solved exactly by simply taking the integral of *p*.

$$y_n = y_{n-1} + \int_{t_{n-1}}^{t_n} p(t)\,dt.$$

The Adams–Bashforth method arises when the formula for *p* is substituted. The coefficients b_j turn out to be

$$beta_j = \frac{(-1)^{s-j}}{(j-1)!(s-j)!}\int_0^1 \prod_{\substack{i=1\\ i\neq j}}^{s}(u+i)\,du.$$

The Adams-Bashforth method is typically used for Linear and Non-liner ODE's with dense systems.

Gear Method

Taking

$$beta_i = 0$$

for i ≥ 1

The Gear method, also known as the **backward differentiation** formulae (BDF, a.k.a. Gear's formulae) is another multi-step method but is generally used for multiple equations. In order to use the gear method your function must have a stiffness greater than 500, meaning that the function is stiff.

$$y_{n+1} = h\beta_0 f(y_{k+1}, t_{k+1}) - \sum_{i=1}^{m} \alpha_i y_{k+1-i}$$

As a particular cases, taking $a_i = 0$ for i ≥ 2 and optimizing the remaining coefficients to maximize the accuracy of the resulting scheme recovers the Implicit Euler method. Taking $a_i = 0$ for i ≥ 3 gives

$$y_{n+1} = -\alpha_1 y_n - \alpha_2 y_{n-1} + h\beta_0 f(y_{n+1}, t_{n+1})$$

We now focus on this case in particular. Applying this method to the scalar model problem and assuming constant h and a solution of the form $y_n = \alpha^n y_0$, we find the following quadratic equation for σ

$$(1 - \beta_0 h\lambda)\sigma^2 + \alpha_1\sigma + \alpha_2$$

the two roots of which are given by

$$sigma = \frac{-\alpha_1 + \sqrt{\gamma(1+\epsilon)}}{2(1-\Sigma)} or \frac{-\alpha_1 - \sqrt{\gamma(1+\epsilon)}}{2(1-\Sigma)}$$

$$gamma = \alpha_1^2 - 4\alpha_2$$

$$\Sigma = (\beta_0)h\lambda$$

Applying the identities, we may expand both roots in terms of powers of h. By our assumed form of the solution, it follows that $y_n = \sigma_+^n y_0, + \sigma_-^n y_0$.The leading-order term in the expansion in h of σ_- (a “spurious root”) is proportional to h. For small h, σ_-^n quickly decays to zero, and thusmay be neglected. The leading-order terms in the expansion in h of σ_+ (the “physical root”) resemble the Taylor-series expansion of the exact solution over a single timestep.

$$\sigma_+ = (-\frac{\alpha_-}{2} + \frac{\sqrt{\gamma}}{2}) - \beta_t(-\frac{\alpha_1}{2} + \frac{\sqrt{\gamma}}{2}[1 + \frac{2\alpha_2}{\gamma}])\lambda h + \beta_0^2(-\frac{\alpha_1}{2} + \frac{\sqrt{\gamma}}{2}[1 + \frac{2\alpha_2}{\gamma} - \frac{2\alpha_2}{\gamma^2}])\lambda^2 h^2 + \ldots$$

Matching coefficients with the expansion of the exact solution $\sigma = e^{\lambda h} = 1 + \lambda h + \frac{\lambda^2 h^2}{2} + \ldots$ as indicated by underbraces in the above expression, and applying the definition $\gamma = \alpha_1^2 - 4\alpha_2$, we arrive at three equations for $\alpha_1, \alpha_2 and \beta_0$to achieve the highest order of accuracy possible with this form. It is easily verified that $\alpha_1 = -4/3, \alpha_2 = 1/3, and \beta_0 = 2/3$satisfy these three equations. The leading-order error term of this method is proportional to h^3. Thus, over a single timestep, the scheme is “locally third-order accurate”; more significantly, over a fixed time interval [0,T], the scheme is globally second-order accurate. The resulting method,

$$y_{n+1} = \frac{4}{3}y_n - \frac{1}{3}y_{n-1} + h\frac{2}{3}f(y_{n+1}, t_{n+1}),$$

is thus referred to as BDF2, and may be viewed as a implicit alternative to AM3 that, for the same number of steps into the past, p = max(m,n), has reduced order of accuracy but greatly improved domain of stability. Higher-order BDFs may be derived in an analogous fashion; BDF3, BDF4, BDF5, and BDF6 in particular are found to have excellent stability properties as compared with their AM counterparts with the same number of steps.

Worked out Example 1

You have been assigned a non-isothermal reactor with the following material and enthalpy balances.

$$fracdc_adt = c_a(t)^2 - c_b(t) - T(t)^2$$

Species a Material balance

$$fracdc_bdt = c_b(t)^3 - c_a(t) - T(t)^2$$

Species b Material balance

$$fracdTdt = c_a(t) - c_b(t) - T(t)^2$$

Enthalpy balance

When is the concentration of species a and species b equal? The initial conditions for the reactor are as follows:

T_0 = 1

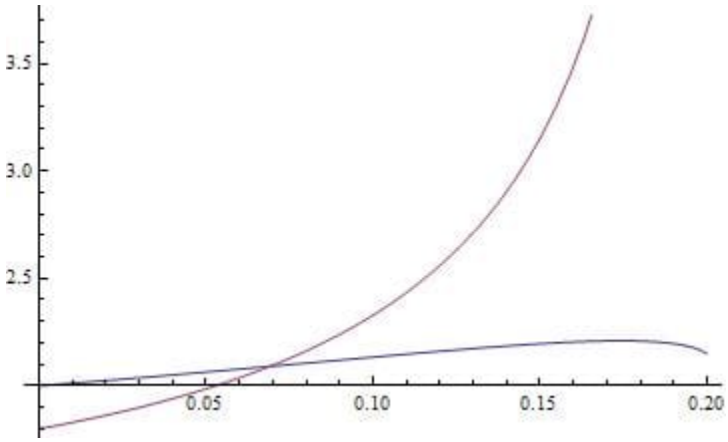
c_a0 = 2.0

c_b0 = 1.8

t_f = 0.2

s = NDSolve[{x'[t] == x[t]^2 - y[t] - z[t]^2, y'[t] == y[t]^3 - x[t] - z[t]^2, z'[t] == x[t] - y[t] - z[t]^2, x[0] == 2, y[0] == 1.8, z[0] == 1}, {x, y, z}, {t, 0.2}]

Plot[Evaluate[{x[t],y[t]} /. s], {t, 0, 0.2}]



Worked out Example 2

You have been asked to study a semi-batch reactor for the reaction $A \rightarrow B$. The rate law is $-r_a = \frac{kC_a}{1+C_b}$, where k = 2.7. Other parameters are: initial volume = 5, volumetric flow rate = 0.05, initial concentration of A = 2500. Use *Mathematica* to create a Conversion versus Time plot for 100 time units.

1. Mole Balance:

$$fracdXdt = -\frac{r_a V}{N_{A0}}$$

2. Rate Law:

$$-r_a = \frac{kC_a}{1+C_b}$$

3. Stoichiometry:

$$C_A = \frac{N_{A0}(1-X)}{V_0 + v_0 t}$$

$$C_B = \frac{N_{A0}X}{V_0 + v_0}$$

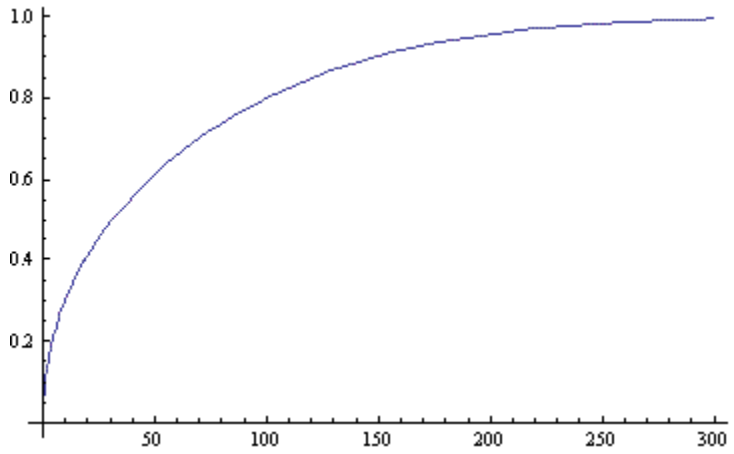
4. Combine:

$$\frac{dX}{dt} = \frac{k(1-X)(V_0 + v_0 t)}{N_{A0}X}$$

If the semi-batch reactor runs for 100 time units, the conversion is about 0.8. This can be seen in the plot below or by running the two commands that follow into *Mathematica*.

s = NDSolve[{y'[t] == 2.7 (1 - y[t]) (5 + .05 t)/(2500 y[t]), y[0] == 0.0001}, y, {t, 0, 100}];

Plot[Evaluate[y[t] /. s], {t, 0, 100}, PlotRange -> All]



This is the *Mathematica* notebook file for the example: Media:MathematicaEx2.nb

Note: The parameters used in this problem are fabricated and are intended to illustrate the use of *Mathematica* in solving ODEs.

Formatting Plots in Mathematica

Here are some useful tips to formatting plots in Mathematica:

Label Plot: e.g. Plot[Sin[x], {x,0,2Pi}, **PlotLabel->"Volume of Tank vs. Time"**]

Label Axes: e.g. Plot[Sin[x], {x,0,2Pi}, **AxesLabel->{"time(min)","volume(L)"}**]

Color Plot: e.g. Plot[Sin[x], {x,0,2Pi}, **PlotStyle->Red**]. You can also make the plot thick, dashed, etc. e.g. Plot[Sin[x], {x,0,2Pi}, **PlotStyle->{Red,Thick,Dashed}**]

Insert Legend: e.g. **Needs["PlotLegends`"]** <--MUST insert this BEFORE your plot command. Plot[{Sin[x],Cos[x]}, {x,0,2Pi}, **PlotLegend->{"sine", "cosine"}**]

Sage's Corner

Solving ODEs with Mathematica Video

video.google.com/googleplayer.swf?docId=5007050669435611182

slides for this talk

Additional Tips and Tricks for Troubleshooting in Mathematica

Mathematica is a powerful computing tool, however the syntax can be a bit difficult to understand. Here are some notes for troubleshooting in Mathematica.

1. Check to make sure that your variable names and signs are consistent.

- Ex) Make sure you use xI everywhere instead of xI and x1 or xl.

- Ex) Functions, including the ones you create, are usually followed by brackets such as Sin[x] or y[x]. However, brackets are not necessary when you are solving for a function in a set of differential equations such as NDSolve[eqns, {y}, {x, 0, 50}];
- Ex) Check to see if your parentheses are aligned such that you are actually entering the function you think you're entering. Recall order of operations and the distributive property. x*(a+b) is NOT equal (x*a) + b. This seems simple, but often gets overlooked when dealing with several lines of code.

2. You may find it easier to define all of your variables together at the beginning of your code. This makes it easy to go back and change the values assigned to each variable if trying to determine the impact one variable has on a system of equations. For instance, say you are trying to determine the effects the flow rates into two different tanks (F1, F2) will have on the tank volumes over ten time steps. The differential equations governing the situation are: $\frac{dV_1}{dt} = 4F_1 - V_1$ and $\frac{dV_2}{dt} = F_2 - \frac{1}{3}V_2$, where F1 = 2, F2 = 12, V1(0) = 0, V2(0) = 0.

If you write the ODEs in Mathematica by directly substituting in F1 = 2 and F2 = 12, you will have to change the ODEs each time you change the values of F1 and F2. Below is an example of what the Mathematica equations would look like.

s = NDSolve[{V1'[t] == 8 – V1[t], V2'[t] == 12 – (1/3)*V2[t], V1[0] == 0, V2[0] == 0},{V1,V2},{t,0,10}]

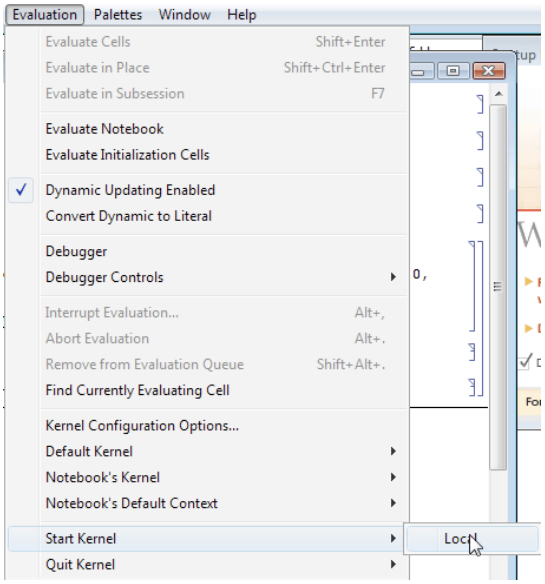
Another option is to define F1 and F2 before your equations, and then call them when solving for your ODEs. This allows you to easily substitute in new values for your variables without changing the differential equations, and reduces the odds of making simple computational errors in your ODE in the process. The Mathematica code would look as shown below.

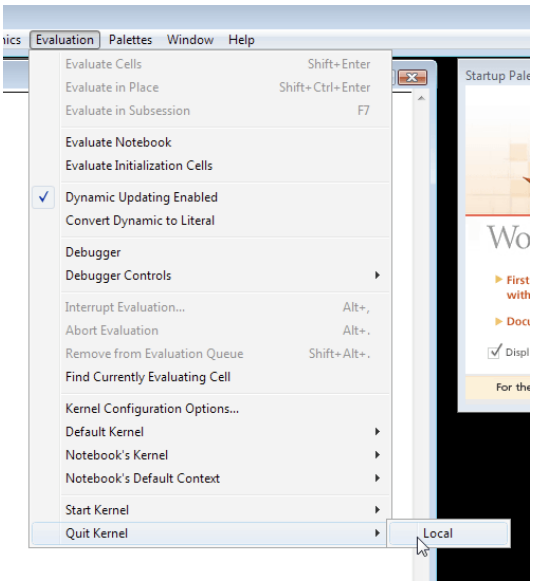
variables = {F1 -> 2, F2 -> 12};

s = NDSolve[{V1'[t] == 4*F1 – V1[t] , V2'[t] == F2 – (1/3)*V2[t], V1[0] == 0, V2[0] == 0} **/. variables**,{V1,V2},{t,0,10}]

3. Be Aware of the Kernel

The Mathematica Kernel stores all of the information about variable definitions. When you define a variable , the definition is stored there. The kernel is automatically started when a new Mathematica session is started. You may also start it manually. To start the kernel manually go to Evaluation -> Start Kernal -> Local. Once the kernel is started and you wish to go back and change a variable definition you must "Quit the Kernal" before you see the change occur. In order to "Quit the Kernal" you must go to Evaluation -> Quit Kernal -> Local. Terminating the Mathematica Kernal, erases all of the definitions previously entered. For this reason, after you "Quit the Kernal" and enter in new definitions for your variables you must re-enter all of your code. The images below show how to "Start Kernal" and how to "Quit the Kernel."





Example:

An example of when you would want to "Quit the Kernal" when using Mathematica for Controls is when you are finding a steady state point using a PID controller. We start by defining the variables for vset, taui, taud, and Kc. Mathematica solves this equation using the defined variables and shows that (x->3.9, V->10) is the steady state point. This is shown in the figure 1 below.

```
In[8]:= vars = {vset → 10, taui → .1, taud → 1, Kc → 10};
        eqnss =
         {(.5 + .5 * Kc * vset - .5 * Kc * V + (.5 / taui) * x - 2 * V) / (1 + .5 * taud) == 0,
          vset - V == 0}
        s = Solve[eqnss, {V, x}] /. vars
Out[7]= {(0.5 - 2 V - 0.5 Kc V + 0.5 Kc vset + 0.5 x / taui) / (1 + 0.5 taud) == 0, -V + vset == 0}
Out[8]= {{x → 3.9, V → 10.}}
```

Figure 1 when you defined Kc even though you have not defined it here.

```
In[2]= vset = 10;
In[3]= taui = .1;
In[4]= taud = 1;
In[5]= eqnss =
         {(.5 + .5 * Kc * vset - .5 * Kc * V + (.5 / taui) * x - 2 * V) / (1 + .5 * taud) == 0,
          vset - V == 0}
        s = Solve[eqnss, {V, x}]
Out[5]= {0.666667 (50.5 - 7. V + 5. x) == 0, 10 - V == 0}
Out[6]= {{V → 10., x → 3.9}}
```

Figure 3 shows the output once you have "Quit the Kernel" and re-entered the Mathematica code.

```
In[3]:= vset = 10;

In[4]:= taui = .1;

In[5]:= taud = 1;

In[6]:= eqnss =
         {(.5 + .5 * Kc * vset  .5 * Kc * V + (.5 / taui) * x  2 * V) / (1 + .5 * taud) == 0,
          vset - V == 0}
        s = Solve[eqnss, {V, x}]

Out[6]= {0.666667 (0.5 + 5. Kc - 2 V - 0.5 Kc V + 5. x) == 0, 10 - V == 0}

Out[7]= {{x → -0.1 - 10. (-0.4 - 0.1 Kc) - 1. Kc, V → 10.}}
```

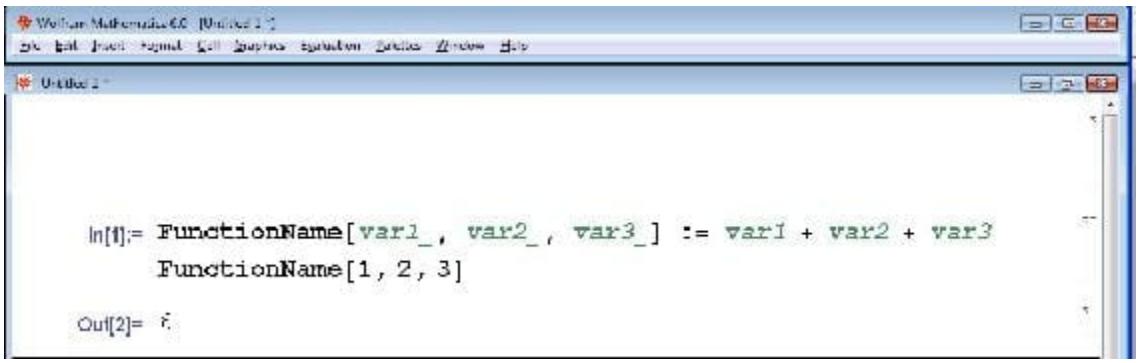
Now you see that the definition for Kc has been deleted because the steady state points are in terms of Kc. To find the impact of Kc you can use this solution.

4. Define functions or formulas that are used often

If you will use a function often and it is not already defined in Mathematica by default, define your own. This can be especially helpful if you intend repeat the same function multiple times while only changing variables. A scenario where defining your own function could save you time and error is when comparing P-Fisher's values while keeping the margins constant. Only the variables will change but the function will remain the same. The new function will only last as long as each session in Mathematica.

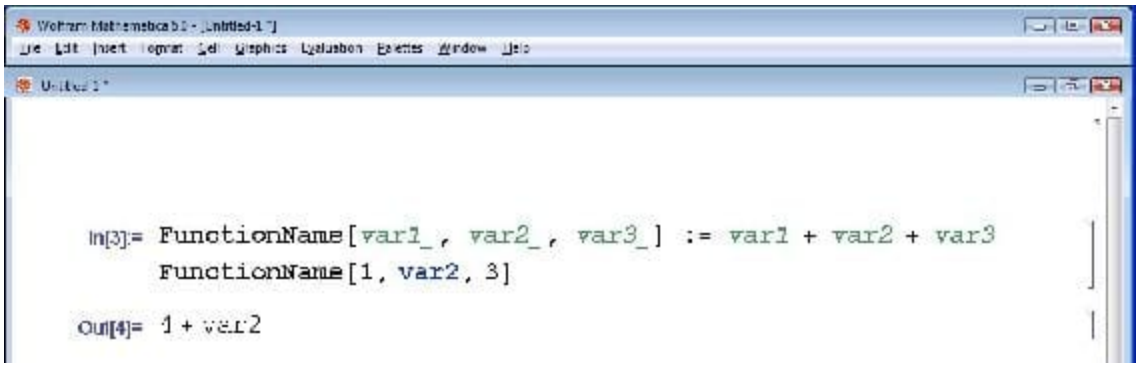
To define a function, begin by naming the function followed by an open bracket. List all the variables with an underscore after each name and separate each variable with a comma. Close the bracket when all the variables have been listed. To let Mathematica know that this is a newly defined function, after closing the bracket place a semi-colon and equal sign. Now define what the formula is using the variable names. When you have finished, click shift-enter simultaneously. To test the new formula, start with naming the function and open bracket, then list all the variables numerical values separated by comma's in the same order that was listed when defining the function. Close the bracket after all the variables have been listed and click shift-enter simultaneously. The output will be the next line in Mathematica with the answer.

Example:



It is also possible to use this tool to put equations in terms of unknown variables. To do this begin by defining the function the same way as before. When plugging in the numerical values of the variables, leave the unknown variables as their variable name. The Mathematica output will provide the answer in terms of the unknown variables.

Example:



5. "Make it pretty"

Oftentimes when people program in any language, they tend to crowd lines, almost as if they were trying to save paper. Well, since this does not apply when you are typing on a screen, don't worry about it.

- Ex) You could write:

vars = {V -> 1000, Cao -> 0.2, UA -> 20000, F -> 2000, DE1 -> 30000, k1 -> 1.0 10^13, DH1 -> 50000, rcp -> 60, R -> 1.987, To -> 500, Tf -> 500, Tjin -> 500, Fj -> 1000, Vj -> 100, Caf -> 0.2, DH2 -> -192000, DE2 -> 30000, k2 -> 1.0 10^13};

eqns = {Ca'[t] == (Caf - Ca[t]))/V - Ca[t]*k1*Exp[-DE1/(R*T[t])], Cb'[t] == (0 - Cb[t])/V +k1*Ca[t]*Exp[-DE1/(R*T[t])] - k2*Cb[t]*Exp[-DE2/(R*T[t])], T'[t] == (Tf - T[t])/V + (-DH1/(rcp))*k1*Ca[t]*Exp[-DE1/(R*T[t])] + (-DH2/rcp)*k2*Cb[t]*Exp[-DE2/(R*T[t])] - (UA (T[t] - Tj[t]))/(V *rcp), Tj[t] == (Fj (Tjin - Tj[t]))/Vj + (UA (T[t] - Tj[t]))/(Vj *rcp), Ca[0] == 0.2, Cb[0] == 0, T[0] == 500, Tj[0] == 500};

sol = NDSolve[eqns /.vars, {Ca, Cb, T, Tj}, {t, 0, 50}]; Plot[{Ca[t]} /. sol, {t, 0, 50}, AxesLabel -> {t, Ca}, PlotRange -> Full];Plot[{Cb[t]} /. sol, {t, 0, 50}, AxesLabel -> {t, Cb},PlotRange -> Full];Plot[{T[t]} /. sol, {t, 0, 50}, AxesLabel -> {t, T}, PlotRange -> Full];Plot[{Tj[t]} /. sol, {t, 0, 50}, AxesLabel -> {t, Tj}, PlotRange -> Full]

- but it looks much better if you do this:

```
vars = {V -> 1000, Cao -> 0.2, UA -> 20000, F -> 2000, DE1 -> 30000,
k1 -> 1.0 10^13, DH1 -> 50000, rcp -> 60, R -> 1.987, To -> 500,
Tf -> 500, Tjin -> 500, Fj -> 1000, Vj -> 100, Caf -> 0.2,
DH2 -> -192000, DE2 -> 30000, k2 -> 1.0 10^13};
```

```
eqns = {
Ca'[t] == (Caf - Ca[t]))/V - Ca[t]*k1*Exp[-DE1/(R*T[t])],
Cb'[t] == (0 - Cb[t])/V + k1*Ca[t]*Exp[-DE1/(R*T[t])] - k2*Cb[t]*Exp[-DE2/(R*T[t])],
T'[t] == (Tf - T[t])/V + (-DH1/(rcp))*k1*Ca[t]*Exp[-DE1/(R*T[t])] + (-DH2/rcp )*k2*Cb[t]*Exp[-DE2/(R*T[t])] - (UA (T[t] - Tj[t]))/(V *rcp),
Tj'[t] == (Fj (Tjin - Tj[t]))/Vj + (UA (T[t] - Tj[t]))/(Vj *rcp), Ca[0] == 0.2, Cb[0] == 0, T[0] == 500, Tj[0] == 500};
```

```
sol = NDSolve[eqns /.vars, {Ca, Cb, T, Tj}, {t, 0, 50}];
```

```
Plot[{Ca[t]} /. sol, {t, 0, 50}, AxesLabel -> {t, Ca}, PlotRange -> Full]
Plot[{Cb[t]} /. sol, {t, 0, 50}, AxesLabel -> {t, Cb}, PlotRange -> Full]
Plot[{T[t]} /. sol, {t, 0, 50}, AxesLabel -> {t, T}, PlotRange -> Full]
Plot[{Tj[t]} /. sol, {t, 0, 50}, AxesLabel -> {t, Tj}, PlotRange -> Full]
```

These thoughts will hopefully help you later on and make it easier to read your code.

6.) "Check the colors!"

Here is a list of font color that Mathmatica will output for a specific reason.

Local Variables in a certain Font Color

Local variables of Module and With in Green

1. Example:

```
Module[{x = 2, y = 3, exponent}, exponent = 2; {x, y}^(exponent+x)]
```

Function arguments and pattern names in *Green (Italics)*

1. Example:

```
func[x_, power_] := {x, x^2, x^power}
```

Variables made special by use in arguments in Turquoise

- Example:

```
Table[x^power, {x, 1, 5}, {power, 1, 5}]
```

Errors and Warnings in a certain Font Color

Syntax Errors in Purple

- Example:

```
Table[triangle[{x+ , y, z]]
```

Emphasized Syntax Errors in Dark Red with Yellow Background

- Example:

```
Table[triangle[{x+ , y, z]]
```

Missing arguments in Bright Red

1. Example:

```
{N[], Rule[expr,]}
```

Excess arguments in Bright Red

1. Example:

```
{Sqrt[5, 7], Log[base, a, wrong]}
```

Possible unwanted assignments in Bright Red

- Example:

```
If[x = 4, y = 9]
```

Unrecognized option names in Bright Red

- Example:

```
Plot[Sin[x], {x, 0, 2 π}, PlotArea → {-1, 1}]
```

Local Scope conflicts in Dark Red with Yellow Background

- Example:

```
f[x_, conflicted_] := Module[{ok, x, conflicted}, ok + x]
```

Variables that will go out of scope before being used in Bright Red

- Example:

```
Table[Dynamic[var], {var, 1, 10}]
```

Shadowing in multiple contexts in Bright Red

- Example:

```
{Global`Sin, Sin}
```

Other in a certain Font Color

Comments in Light Grey

- Example:
```
(* A comment explaining something *)
```

Strings in Dark Grey

- Example:
```
"A string of characters"
```

Global symbols that have no value assigned in Bright Blue

- Example:
```
{Table, undefinedsymbol, qq1}
```

Accessing Mathematica from your personal computer

If you prefer to use Mathematica from your home computer rather than one on campus, you can remotely log into a CAEN computer through U-M's Virtual Sites service. Follow the steps outlined below:

2. Visit http://virtualsites.umich.edu/ . The U-M WEBLOGIN page will appear. Log in using your unique name and Kerberos password.
3. The Virtual Sites page should appear on your screen. Click on the large, orange **Connect Now** button.
4. Under the heading **What software do you need?**, select **Engineering (CAEN) Legacy XP**. Then click on the **Request Connection** button toward the bottom of the page.
5. Follow the instructions under the heading **Launch Virtual Sites** by downloading and opening the file in the Internet Explorer information bar. A window will pop up on your screen titled **Remote Desktop Connection**. Click the **Connect** button.
6. You will then see the screen that appears when logging into a CAEN computer. Your unique name should already be included. Type your Kerberos password and click **OK** to log in to the CAEN computer.
7. Click **OK** on the **Notice** window that pops up and your personal settings will be applied. Also click **OK** if a **System Requirements Wizard** window appears.
8. If you scroll down using the scroll bar on the right-hand side of the screen, you will see the **Start** button. Click **Start -> All Programs -> Math and Numerical Methods -> Wolfram Mathematica -> Mathematica 6**. Once Mathematica opens on the computer screen, you may use it as though you were actually sitting in front of a CAEN computer on campus.
9. To log out, click **Start -> Log Off**. Under the **Log Off Windows** window, click **Log Off**.

References

- "Adams Method", Wolfram MathWorld, Online: August 5, 2007. Available http://mathworld.wolfram.com/AdamsMethod.html [1]
- "Gear Predictor-Corrector Method", Davidson University, Online August 5, 2007. Available webphysics.davidson.edu/Projects/SuFischer/node47.html [2]
- Fogler, H. Scott (2006), Elements of Chemical Reaction Engineering, New Jersey: Pretince Hall PTR. ISBN 0-13-047394-4 [3]
- "Adams-Bashforth", Pitt Math, Online: September 8, 2007. Available: www.math.pitt.edu/~sussmanm/2071Spring07/lab02/index.html#AdamsBashforthMethods [4]
- "Lipschitz Condition", Wolfram MathWorld, Online: September 7, 2007. Available http://mathworld.wolfram.com/LipschitzCondition.html [5]
- "Adams-Bashforth (AB) Methods", University of California-San Diego, Online: September 7, 2007. Available: renaissance.ucsd.edu/chapters/chap11.pdf [6]
- "Gears Method", University of California-San Diego, Online: September 7, 2007. Available: renaissance.ucsd.edu/chapters/chap11.pdf [7]
- Use the Harvard Referencing style for references in the document.
- For more information on when to reference see the following Wikipedia entry.

This page titled 2.7: Solving ODEs with Mathematica- How to find numerical and analytical solutions to ODEs with Mathematica is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

2.8: Fitting ODE parameters to data using Excel- Using regression to fit complex models in Excel

Section 8. Fitting ODE parameters to data using Excel: Using regression to fit complex models in Excel

Note: Video lecture available for this section!

Authors: Anthony Campbell, Victoria Cardine, David Hines, Stephen Kerns /Date Revised: 14th September 2007

Stewards: Jeff Byrd, Khek Ping Chia, John Cruz, Natalie Duchene, Samantha Lyu

Introduction

One of the most important abilities of an engineer is to model measured values of a system. Modeling of a system aids in optimizing objectives, as well as predicting future operations. This can be especially helpful when changes cannot be immediately implemented. In general modeling terms, a set of data can always fit a set of variables. However, an excessive number of degrees of freedom usually exist, creating multiple solutions. Therefore, the physical significance of the variables is an important consideration.

As time proceeds, various values may need to be changed to accommodate variations within the system. The following sections will discuss the use of Microsoft Excel© as a tool to adjust the values of variables given a set of experimental data. The models presented below take into account the design equations of the system being modeled, and the rate laws, where applicable.

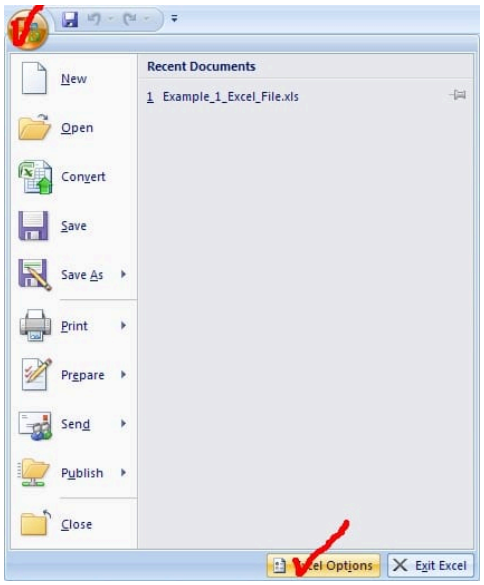
Adding in the Solver Application in Excel 2007

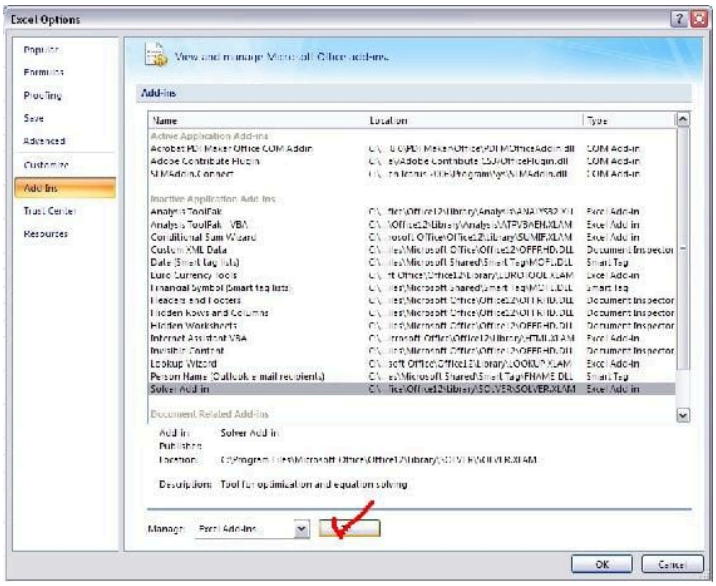
In order to use "Solver" in Excel 2007, one must first add the Excel tool-pack. To add the tool-pack, follow these instructions:

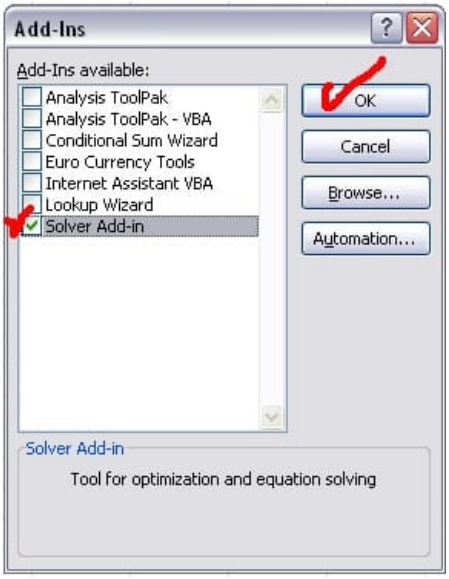
- Click on the Microsoft Office symbol at the top left corner of the screen
- Click on “Excel options” at the bottom right corner of the drop box
- Click on “Add-Ins” in the left hand column
- Click on the “Solver Add-In” option at the bottom of the list
- Click on “GO” at the bottom right corner
- Check the “Solver Add-in” box at the bottom of the list
- Click on “OK” at the top of the right column

To access Solver, click on the “Data” tab at the top of the screen. The Solver tool is located in the "Analysis" box on the right side of the screen.

The user interface of Excel 2007 differs from those of previous versions. The following images illustrate the process just described.


Recent Documents
1 Example_1_Excel_File.xls
New
Open
Convert
Save
Save As
Print
Prepare
Send
Publish
Close
Exit Excel


Excel Options
Add-Ins
Name
Location
Type
Solver Add-in
Manage:
OK



In order to use the Solver function, click on Data and then the Solver button in the Analysis box.

NOTE: If using Microsoft Office 2003, the Solver Application can be added by clicking on the "Tools" tab at the top of the screen. The same directions can be followed after this step.

NOTE: If you receive an error message while trying to access Solver, it may already be added in Excel. Uninstall Solver (by following the same steps, but unchecking "Solver Add-in") then reinstall it.

Adding in the Solver Application in Excel 2008

Excel 2008 for Macs does not come with the solver add-in. The software can be downloaded at Solver.com

This solver application requires Excel 12.1.2. If this is not the current version of Excel being used, it can be updated by running Microsoft AutoUpdate or by opening excel, going to help and then clicking on "Check for Updates". Make sure when updating Excel to close all Microsoft Applications.

This application allows the user to have a solver program that can be opened and used while running Excel. It works the same way as the solver add-in.

Using Excel Solver to Fit ODE Parameters

Excel Solver is a tool that can be used to fit function variables to given experimental data. The following process can be used to model data:

- Define the data set by entering values into an Excel spreadsheet
- Define the model that you want to fit to the data
- Define the sum of least squares in one of the spreadsheet blocks
- In Solver, have Excel minimize the sum of least squares by varying the parameters in the model

Accurate modeling is dependent on two factors; initial values and verifying results.

Initial values: As stated in the introduction, many data fitting problems can have multiple “solutions”. In numerical methods,

given a set of initial parameters the data will converge to a solution. Different initial values can result in differing solutions. Therefore, if the initial values are not set by the problem statement, multiple initial guesses should be entered to determine the “best” value for each variable.

Verification of curve fit: When fitting a curve to given data points, it is important to verify that the curve is appropriate. The best way to do this is to view the data graphically. It is fairly simple to use the Chart tool to graph the data points. Adding a trend line will show the mathematical relationship between the data points. Excel will generate a function for the trend line and both the function and the R-squared value can be shown on the plot. This can be done by clicking “Options” next to “Type” of trend and checking the empty boxes that are labeled “Display equation on chart” and “Display R-squared value on chart”.

There are various types of trendlines that can be chosen for your given data points. Depending on the trend that you may see, you might want to try a few of the options to obtain the best fit. Sometimes the squared residual value will be better for a certain type of trend, but the trend is not necessarily “correct”. When you have obtained sample data, be aware of the trend you should be seeing before fitting a certain trend type to the data points (if this is possible). The polynomial trend type gives you the option to change the order of the equation.

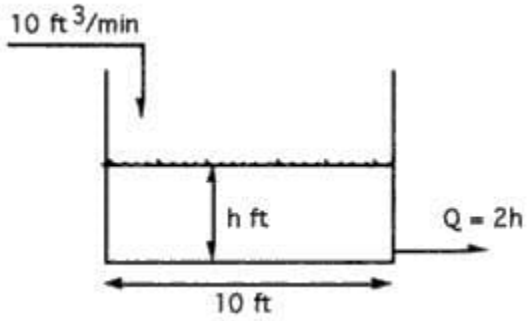
Typically, you will not acquire a squared residual value of zero. This value is a simple analysis of the error between the trend line, and the actual data. Once the function and squared residual values are generated, you can begin to evaluate the generated solution.

During your solution analysis, it is important to check for variables or parameters that may be unnecessarily influencing the data. Excel can generate a "best" solution, which is not correct due to background calculations or Excel's misinterpretation of the modeled system. To assess this possibility, you can graph data generated by using the solved for variables against the given data.

Additional information about the Excel Solver Tool can be found in Excel Modeling.

Worked out Example 1: Mass Balance on a Surge Tank

In this example we will walk through the scenario in which you want to model a non-heated surge tank in Microsoft Excel©. This model would be very critical in a process where smooth levels of process flow are required to maintain product specifications. Therefore, an engineer would like to know whether they will need to periodically refill or drain the tank to avoid accidents. Let's look at this simple example:



Problem statement:

Suppose water is being pumped into a 10-ft diameter tank at the rate of 10 ft3/min. If the tank were initially empty, and water were to leave the tank at a rate dependent on the liquid level according to the relationship Qout = 2h (where the units of the constant 2 are [ft2/min], of Q are [ft3/min], and of h are [ft]), find the height of liquid in the tank as a function of time. Note that the problem may be, in a way, over specified. The key is to determine the parameters that are important to fit to the experimental surge tank model.

Solution:

The solution with fitted parameters can be found here: Surge tank solution This Excel file contains this scenario, where given the actual data (the height of the liquid at different times), the goal is to model the height of the liquid using an ODE. The equations used to model h(t) are shown and captioned. In order to parameterize the model correctly, Solver was used to minimize the sum

of the least square differences at each time step between the actual and model-predicted heights by changing the parameters' values. The model predicts the actual height extremely well, with a sum of the least squared differences on the order of 10^-5. A graph showing the modeled and actual height of the tank as a function of time is included as well.

Worked out Example 2: Fitting a heated surge tank model(a heated CSTR) parameters to data in Excel

Problem Introduction: In this example we will analyze how we can fit parameters used for modeling a heated CSTR to data stored in Excel. The purpose of this problem is to learn how to vary ODE parameters in Excel in order to achieve a square difference between the calculated and experimental data of zero. Some questions to consider are, “How do I model this? Is this a first order ODE? Are there any coupled ODE's? How many variables need to be considered?" This example will use an Excel file to help you in this process.

Problem Statement: Download the heated CSTR problem file.CSTR Problem Take the experimental data produced from the heated CSTR model in the Excel spreadsheet*. You are modeling a heated CSTR. It would be wise and is highly recommended to review both the heated CSTR problem and a wiki on using Excel's solver.

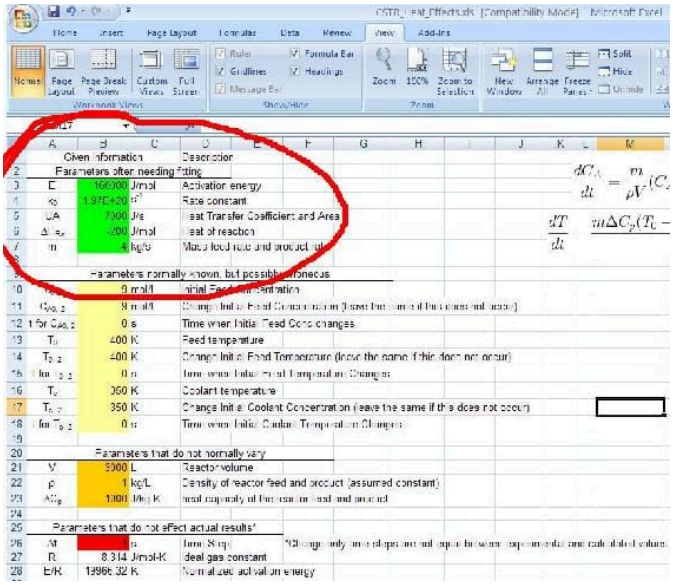
You know this is a heated CSTR...

2. Numerically solve the ODEs through modifying the given wrong set of parameters.
3. While changing the parameters, use Excel to calculate the error between the calculated model and the experimental data.
4. Use the Excel solver function to solve for some of the parameters to minimize error to less than 5%
5. How would your approach be different if our heat transfer area needed to be as small as possible due to capital constraints? Explain.
6. How would your approach be different if our mass feed rate has to be in the range of 3-3.5 kg/s? Explain.

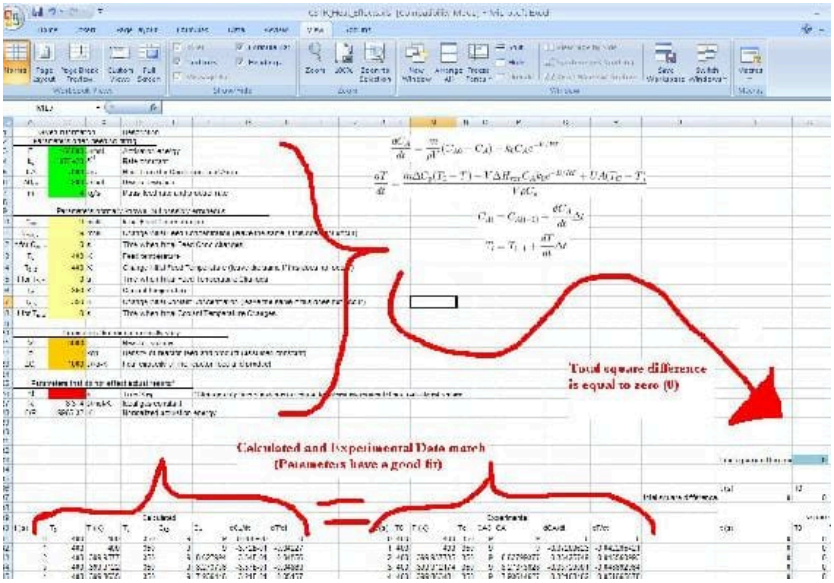
Problem Solution: The solution with fitted parameters can be found here.CSTR Solution

With the model and data given in the example the error must first be defined. Here there error is the squared residual or [(model observation at time i)-(data observation at time i)]^2 over (all i).

Chemical engineering intuition must be used to reasonably approximate some variables. In real life we would double check that "parameters normally known, but erroneous" are correct. (Due to problem constraints, we will assume these are accurate, but a true chemical engineer would check regardless). We must now modify "parameters often needing fitting". For part 3, we can modify the parameters as shown in the solution.



The original parameters have been modified to produce an error of **zero**.



If we have restrictions on heat exchange area then we should try to fit the "heat transfer coefficient and area" as low as possible. Likewise, if we have restrictions on feed rate, we must keep this above 4 while fitting other parameters as best as possible to achieve a small error. These have certain industrial applications as costing of heat exchangers and limitations on mass flow rates need to be modeled before a reactor design is implemented.

Side Note: The example above assumes a first order rate law. This is not always the case. To change the equations to fit a new rate law, one must derive a new dCa/dt equation and dT/dt using their knowledge of reactor engineering. Once the new equations are derived write the formula in an excel format in our dCa/dt and dT/dt column and then drag it to the end of the column. All other columns should stay the same.

Sage's Corner

A brief narration for a better understanding of Example 1 - Surge Tank Model and tips for effectively picking a model equation can be found here: video.google.com/googleplayer.swf?docId=7022226246958174l
A copy of the slides can be found here:

Unnarrated Slides

An example of fitting an ODE parameter to data using excel - based on a simple radioactive decay rate ODE: video.google.com/googleplayer.swf?docId=8229764325250355383
A copy of the slides can be found here: Unnarrated Slides
A copy of the excel file can be found here: Excel File

An example of modelling a reaction rate ODE with rate constant and rate order variable parameters:
video.google.com/googleplayer.swf?docId=-4220475052291185071
A copy of the slides can be found here: Unnarrated Slides
A copy of the excel file can be found here: Excel File

References

Holbert, K.E. "Radioactive Decay", 2006, ASU Department of Electrical Engineering, Online: October 1, 2007. Available www.eas.asu.edu/~holbert/eee460/RadioactiveDecay.pdf

---

This page titled 2.8: Fitting ODE parameters to data using Excel- Using regression to fit complex models in Excel is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

2.9: Helpful Mathematica Syntax- Hints on how to use Mathematica to model chemical processes

Section 9. Helpful Mathematica Syntax: Hints on how to use Mathematica to model chemical processes

Title: Useful functions in Mathematica as pertaining to Process controls

Introduction

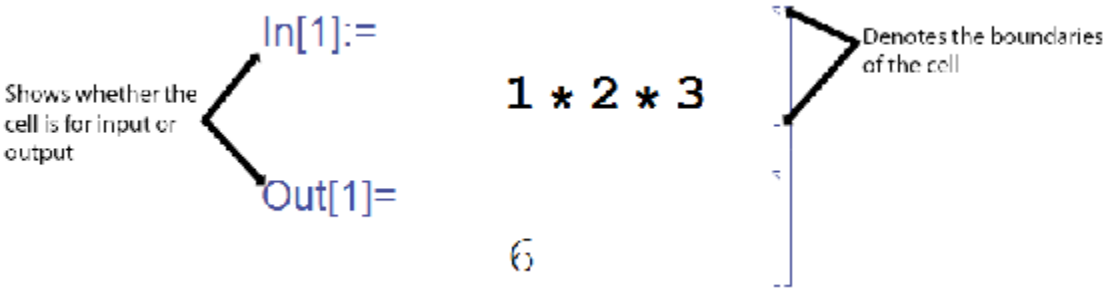
Mathematica has many different functions that are very useful for controlling chemical engineering processes. The Mathematica syntax of these functions is often very complicated and confusing. One important resource is the Mathematica Documentation center, which is accessed through the program help menu. The Documentation Center provides an extensive guide of every function in the program including syntax and sample code. A "First Five Minutes with Mathematica" tutorial is also present in the menu, which provides a quick guide to get started with Mathematica. This article will briefly outline several useful Mathematica functions and their syntax.

Mathematica Basics

Here are some basic rules for Mathematica.

Notebook, Cells and Evaluation

The standard file for Mathematica is called a notebook. The notebook consists of two kinds of cells: input and output. The inputs are typed into one cell and the outputs are displayed in another (See Below).



In order to have Mathematica calculate your inputs, you must evaluate the cells. Do this by pressing **Shift+Enter**. This will only evaluate the current cell that your cursor is located. To evaluate all the cells of a notebook press Ctrl+A, to select all, and then press Shift+Enter.

Parentheses vs. Brackets

Brackets, [], are used for all functions in Mathematica. The only thing parentheses, (), are used for is to indicate the order of operations.

Equality, = vs. ==

The single equal sign (=), is used to assign a value or a function to a variable. The double equal sign (==) is used to set two values equal to each other, such as solving for a value, or to test equalities. When testing equalities, Mathematica will output 'True' or 'False.'

If you recieve 'True' or 'False' as an output when you are expecting it, you likely used a == somewhere that you should have used a =.

Semicolon Use

Placing a semicolon after a line of code will evaluate the expression; however, there will be no output shown.

Mathematica Functions

All Mathematica functions start with a capital letter, eg. *Sin[x], Exp[x], Pi* or *Infinity*.

Assigning and Inserting Variables or Parameters

To assign values to variables, use -> rather than an equals sign. For example,

```
param = {Kc → 0.5, Ca0 → 1}
{Kc → 0.5, Ca0 → 1}
```

To insert a set of parameters to a function use /. This symbol applies a rule or list of rules in an attempt to transform each subpart of an expression.

For example, if you want to enter the above parameters into the expression $y = Kc * x + Ca0$, enter the following in Mathematica:

```
y = Kc * x + Ca0 /. param
1 + 0.5 x
```

Variables are case sensitive. For example, X is different than x. Also, if you define a variable, it will remain defined until you redefine it as something else or clear it. You can clear a variable, for example, x, by using the Clear[x] command. You can also quit the Kernel, which will clear all variables.

Forcing a Numerical Expression

To force Mathematica into giving a numerical solution, use the function *N[expression]*. For example,

```
In[7]:= (5 * 10 * 20 - 3) / (3! * 5)
Out[7]= 997/30

In[8]:= N[(5 * 10 * 20 - 3) / (3! * 5)]
Out[8]= 33.2333
```

Another method to do this is to place a decimal place in any of your numbers (i.e. "5." instead of just "5")

Integration

Mathematica can do some very complex integrations with fairly little coding. Integration can be very helpful for many engineering applications, but when the functions become very complex they become increasingly difficult to integrate. Mathematica will integrate any function that is possible to integrate.

To integrate a function *f(x)* the Mathematica, the command is *Integrate[]*.

For an indefinite integral the correct coding is *Integrate[f(x),x]* where x is the independent variable.

For a definite integral with limits of *y* and *z* the correct coding is *Integrate[f(x),{x,y,z}]*

For example:

We can integrate a function like the one below:

f(*x*) = *Sin*(5 * *x* / 8) * *x*

We can find the indefinite integral as seen here:

```
In[45]:= Integrate[Sin[5 x / 8] * (x), x]
```

$$\text{Out[45]= } -\frac{8}{5}\,x\,\text{Cos}\left[\frac{5\,x}{8}\right]+\frac{64}{25}\,\text{Sin}\left[\frac{5\,x}{8}\right]$$

We can find the definite integral from 0 to π as seen here:

```
In[46]:= Integrate[Sin[5 x / 8] * (x), {x, 0, Pi}]
```

$$\text{Out[46]= } \frac{8}{25}\left(8\,\text{Cos}\left[\frac{\pi}{8}\right]+5\,\pi\,\text{Sin}\left[\frac{\pi}{8}\right]\right)$$

Solver

Mathematica's *Solve* function attempts to solve an equation or set of equations for the specified variables. The general syntax is *Solve[eqns, vars]*

Equations are given in the form of *left hand side == right hand side*.

A single variable or a list of variables can be also be specified. When there are several variables, the solution is given in terms of a list of rules. For example,

```
Solve[{y == 5 * x + 3, y == 6 * x - 2}, {y, x}]

{{y → 28, x → 5}}
```

When there are several solutions, *Solve* gives a list of them. For example,

```
z = 0;

Solve[z == w^2 - 4 * w + 3, w]

{{w → 1}, {w → 3}}
```

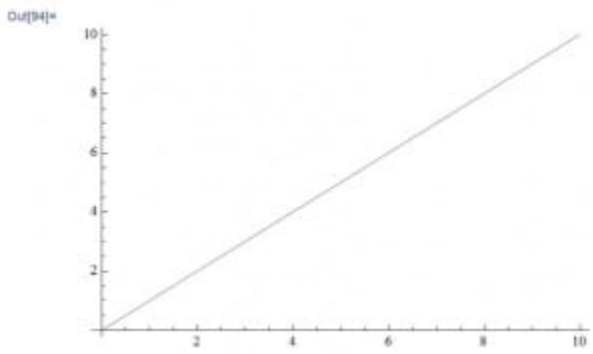
Solve will not always be able to get explicit solutions to equations. It will give the explicit solution if it can, then give a symbolic representation of the remaining solutions in terms of root objects. If there are few symbolic parameters, you can the use *NSolve* to get numerical approximations to the solutions.

Plotting

Mathematica contains a very powerful plotting interface and this overview will go over the basics of plotting. Using Mathematica's Plot[ ] function there is a basic syntax to be used:

Plot[function, {variable, variable_min, variable_max}]

For example, the command "Plot[y=x, {x,0,10}]" will output a graph of the function y=x for a range of x from 0 to 10.



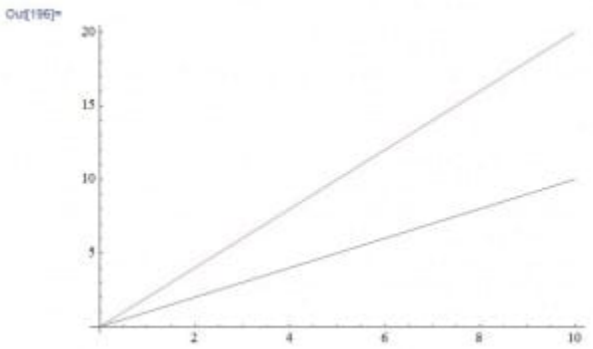
The function could also be previously defined and then referenced in the Plot [ ] command, this functionality provides a lot of simplicity when referencing multiple functions and helps prevent a typo from affecting your Plots.

y[x]==x

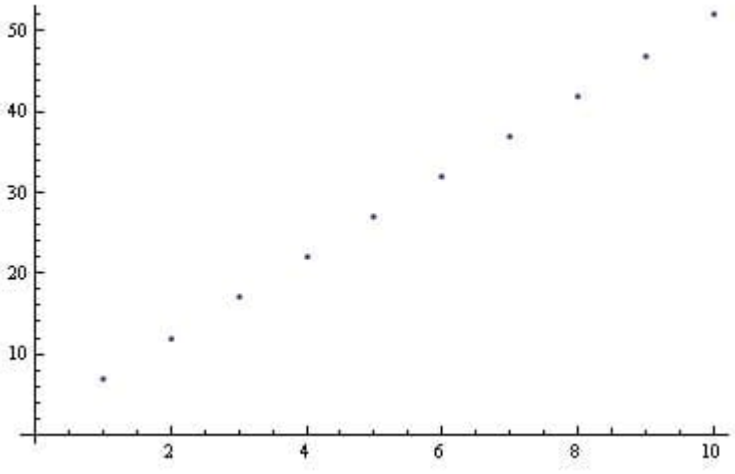
Plot[y[x],{x,0,10}]

To plot multiple functions on the same plot, use an array of functions using curly brackets { }, in place of the single one:

Plot[{y=x,y=2*x}, {x,0,10}]



Another useful plotting option in Mathematica is the ListPlot function. To use this function, the following formatting should be used: ListPlot[Table[f(x),{x,min,max}]



This is obviously only the basics, Mathematica provides many more options including colors, legends, line styles, etc. All of these extra features are well documented at Wolfram's Reference Site.

An additional piece of information: Type the following command on the line after the Plot command to determine the max value plotted. It will display the number associated with the absolute max of the function within the plotting range.

Max[Last/@Level[Cases[%,_Line,Infinity],{-2}]]

Solving ODEs

Within Mathematica, you have the option of solving ODEs numerically or by having Mathematica plot the solution.

Explicit solution:

The built-in function *DSolve* solves a differential equation or a list of differential equations for either one indepedent variable or a list of independent variables, which solves a partial differential equation. The most basic syntax for the function *DSolve* is *DSolve[eqn,y,x]*. For example,

```
DSolve[y'[x] + 3 y[x] + 4 == Sin[x], y[x], x]
```

$$\left\{\left\{y[x] \to e^{-3x} C[1] + \frac{1}{30}(-40 - 3\,\text{Cos}[x] + 9\,\text{Sin}[x])\right\}\right\}$$

The syntax for solving a list of differential equations is *DSolve[{eqn1,eqn2,...},{y1,y2,...},x]*, while the syntax for solving a partial differential equation is *DSolve[{eqn, y, {x1,x2,...}}]*. For example,

```
DSolve[{y'[x] == 2 * y[x], z'[x] == 3 * z[x] + 4 * Sin[x]}, {y[x], z[x]}, x]
```

$$\left\{\left\{y[x] \to e^{2x} C[1],\ z[x] \to e^{3x} C[2] - \frac{2}{5}(\text{Cos}[x] + 3\,\text{Sin}[x])\right\}\right\}$$

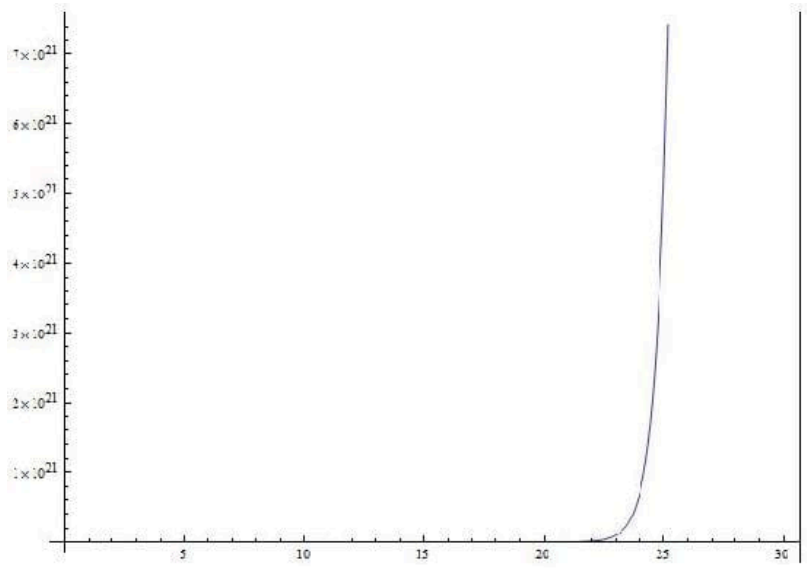
Note that differential equations must be stated in terms of derivatives, such as *y'[x]*. Boundary conditions can also be specified by giving them as equations, such as *y'[0]==b*.

Plotted Solution:

Solutions given by *DSolve* sometimes include integrals that cannot be carried out explicitly, thus *NDSolve* should be used. *NDSolve* finds a numerical solution to an ordinary differential equation within a specified range of the independent variable. The result will be given in terms of *InterpolatingFunction* objects, and thus the solution should be plotted. For example,

```
s = NDSolve[{y'[x] == 2 * y[x], y[0] == 1}, y, {x, 0, 30}]

{{y → InterpolatingFunction[{{0., 30.}}, <>]}}

Plot[Evaluate[y[x] /. s], {x, 0, 30}]
```

Note that a place holder was defined to store the *NDSolve* solution in order for the solution to be plotted. The general syntax for plotting is *Plot[Evaluate[y[x] /. s], {x, xmin, xmax}]*, where s is said placeholder. The plot below shows the solution to this example.



Another Example Solving ODEs

In this example the water level of a tank is controlled by adjusting the feed rate of water into the tank. The system is modeled by the following equation:

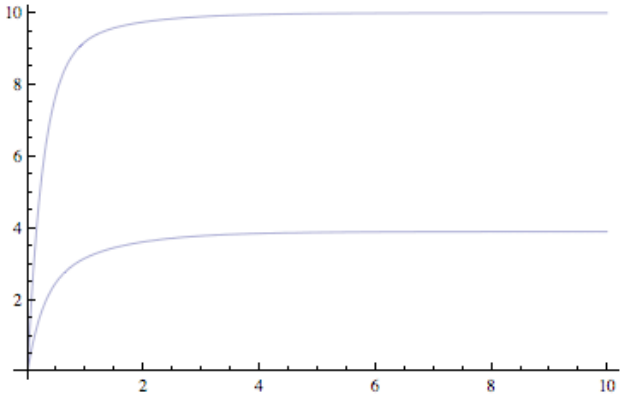
$$\frac{dV}{dt} = \frac{Fin}{2} - 2 * V$$

There is a PID controller on the valve but is not important in explaining the Mathematica syntax. Consider it just a different way to define Fin.

Parameters are assumed to be Vset = 10, Kc = 10, tauI = 0.1, and tauD = 1

For this problem the differential equation needs to be modeled with the given parameters in Mathematica. The problem asks what the steady state of the system is. This is solved for by substituting in a PID expression for Fin (not important in this context), setting the derivatives equal to zero, and solving for V and Xi(a made up variable for the PID controller). The following Mathematica code does all said things.

```
Fin = 1 + Kc (Vset - V[t]) + 1 / ti * xi[t] - td * V'[t];
eqns = {V'[t] == 0.5 * Fin - 2 * V[t], xi'[t] == Vset - V[t], V[0] == 0, xi[0] == 0};
param = {Kc → 10, ti → 0.1, td → 1, Vset → 10};
tmax = 10;
sol = NDSolve[eqns /. param, {V[t], xi[t]}, {t, 0, tmax}];
Plot[{V[t], xi[t]} /. sol, {t, 0, tmax}]
```



```
s = DSolve[eqns /. param, {V[t], xi[t]}, t];
Limit[s[[1]][[1]][[2]], t → Infinity]
Limit[s[[1]][[2]][[2]], t → Infinity]
(*OR*)
Solve[eqns /. param /. {V'[t] → 0, xi'[t] → 0}, {V[t], xi[t]}]
```

```
10.
```

```
3.9
```

```
{{xi[t] → 3.9, V[t] → 10.}}
```

The final line shows that the steady state values for the problem are V = Vset or 10 units of volume, and Xi (the made up variable for the PID controller) has a steady state value of 3.9

In this example there is syntax for defining parameters, applying those parameters to equations, setting up equations, and solving equations for steady state values. There is also syntax for plotting multiple equations on the same graph and how to evaluate multiple functions over time periods. Overall very useful for solving Controls Engineering problems in Mathematica.

Matrix

It is very simple to create a matrix in Mathematica. Simply pick a name for the matrix (for this example we will use a) then simply enter each row inside of curly brackets {} with the individual values separated by commas, and the the row separated by a comma. The function MatrixForm[] will display a matrix in its matrix form rather than as a list of numbers.

EX: a =

ParseError: EOF expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

MatrixForm[a] displays:

5 2 3

7 9 11

8 13 9

You can refer to any specific position in the matrix by using the system:

matrix[[row],[column]]

EX: For the matrix above a[[1],[2]] a value of 2 would be returned.

Jacobians and Eigenvalues

The Jacobian is a helpful intermediate tool used to linearize systems of ODEs. Eigenvalues can be used with the Jacobian to evaluate the stability of the system. Mathematica can be used to compute the Jacobian and Eigenvalues of a system by the following steps.

1. Enter the system as individual equations

$$\frac{dx}{dt} = 3x^2 + 2y$$

$$\frac{dy}{dt} = x^2 - 3y^2$$

entered as

eq1 = 3*x^2+2y

eq2 = x^2-3*y^2

2. Compute the Jacobian with Mathematica's Derivate function

$$Jac = \begin{pmatrix} 6x & 2 \\ 2x & -6y \end{pmatrix}$$

computed by

Jac =

ParseError: "]" expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

3. Use Mathematica's Eigenvalues function to evaluate the stability of the system

Begin by entering a desired value of {x,y} into the matrix (i.e. steady state values of x and y).

Jac = $\begin{pmatrix} -2.31 & 2 \\ -0.77 & 1.33 \end{pmatrix}$

Then call the Eigenvalue function.

Eigenvalues[Jac]

Result: λ={-1.82, 0.85}

The Jacobian can be computed as long as the matrix is square. Many other function types are acceptable when entering eq1 and eq2, including trigonometric and exponential functions.

User-Defined Functions

Mathematica also supports user-defined functions. Defining your own functions helps to simplify code, especially when lengthy expressions are used consistently throughout the code. The generic Mathematica syntax for defining a function is as follows:

f[a_,b_,c_]:=(a+b)/c

Here, the function name is "f" with three arguments: a, b, and c. The syntax ":=" assigns the expression on the right hand side to the function "f." Make sure that each of the arguments in the brackets on the left hand side are followed by underscores; this identifies the variables being called by the function as local variables (ones that are only used by the function itself). For this example, the function "f" will take three arguments (a, b, and c) and calculate the sum of a and b, and then divide by c.

After defining the function, you can use the function throughout your code using either real numbers, as the first example shows, or any variables. This is done by calling the function with the correct number of arguments. When referencing variables, if the variable already has an assigned value the value will be used, if it does not, Mathematica will display your function's results with the variables displayed. A few examples are shown below.

$$f[1,3,2] = \frac{1+3}{2} = 2$$

$$f[x,y,z] = \frac{x+y}{z}$$

Nonlinear Regression

Mathematica can be used to solve multiple variable, nonlinear regression problems. This form of regression can be useful to create models of how temperature or pressure sensors depend on different valves in a system. The NonLinearRegress function in Mathematica can be used to find a best-fit solution to a user defined functional form.

The NonLinearRegress function is not part of the standard package in Mathematica. For this reason, this function must be loaded into Mathematica before an attempt to use.

After this step, the NonlinearRegress is loaded and can be used to solve for a set of data. To use the this function, enter NonlinearRegress[data, expression, parameters, variables]. The best fit values for the parameters will be returned along with statistical information about how well this form fits the given data. An example of Nonlinear Regression in Mathematica is shown in the image below.

```
In[81]:= Needs["NonlinearRegression`"]
         NonlinearRegress[{{1, 2, 2}, {3, 2, 4}, {5, 4, 5}, {6.5, 4, 3}},
          a*x1 + b*x2 + c, {a, b, c}, {x1, x2}]

Out[82]=
{BestFitParameters → {a → 0.16, b → 0.2, c → 2.28},

                          |Estimate Asymptotic SE CI
                         ---------------------------------------
 ParameterCITable →      a|0.16     1.12          {-14.0709, 14.3909}
                         b|0.2      2.32164       {-29.2992, 29.6992}
                         c|2.28     3.69345       {-44.6498, 49.2098}

 EstimatedVariance → 3.92,

                                   DF  SumOfSq  MeanSq
              Model                3   50.08    16.6933
 ANOVATable → Error                1   3.92     3.92    ,
              Uncorrected Total    4   54.
              Corrected Total      3   5.

                                  (    1.     -0.904534  0.530669 )
 AsymptoticCorrelationMatrix →    (-0.904534    1.      -0.822871 ),
                                  ( 0.530669 -0.822871     1.      )

                                                  |Curvature
                                  ---------------------------
 FitCurvatureTable →              Max Intrinsic            |0
                                  Max Parameter-Effects    |0
                                  95. % Confidence Region  |0.0680875
```

Probability Density Function

Often in controls, it is necessary to calculate the probability that a certain event will occur, given a particular known distribution. In these scenarios, it is helpful to know how to use the built-in probability density function in Mathematica. The syntax for using the probability density function to calculate the probability of observing a value *x* for a generic distribution *dist* is shown below.

PDF[*dist*,*x*]

A variety of distributions can be used in conjunction with the probability density function. These distributions include the normal distribution and the binomial distribution.

Normal Distribution

For a normal distribution with a mean μ and a standard deviation σ, use the following syntax to specify the distribution.

NormalDistribution[μ,σ]

To calculate the probability of observing a value *x* given this distribution, use the probability density function as described previously.

PDF[NormalDistribution[*μ*,*σ*],*x*]

To calculate the probability of observing a range of values from *x1* to *x2* given this distribution, use the following syntax.

NIntegrate[PDF[NormalDistribution[*μ*,*σ*],*x*],{*x*,*x1*,*x2*]

For the theory behind the probability density function for a normal distribution, click here.

Binomial Distribution

For a binomial distribution with a number of trials *n* and a success probability *p*, use the following syntax to specify the distribution.

BinomialDistribution[*n*,*p*]

To calculate the probability of observing a successful trial exactly *k* times given this distribution, use the probability density function as described previously.

PDF[BinomialDistribution[*n*,*p*],*k*]

To calculate the probability of observing between a range of *k1* number of successes and *k2* number of successes given this distribution, use the following syntax.

NIntegrate[PDF[BinomialDistribution[*n*,*p*],*k*],{*k*,*k1*,*k2*]

Alternatively, binomial distribution can be found using a user defined function. Define the equations as follow: binom[nn_,kk_,pp_]:= nn!/(kk! (nn-kk)!) pp^kk (1-p)^(nn-kk)

where:

number of independent samples = nn

number of events = kk

probability of the event = pp

To find the odds of getting 5 heads out of 10 coin tosses,assuming the probability of head = 0.5, substitute the values for nn,kk and pp as following: binom[10,5,0.5]

Using the same scenario, but instead for the odds of getting 5 or more heads in the 10 tosses, the function can be used as: Sum[binom[10,i,0.5],{i,5,10}]

For the theory behind the probability density function for a binomial distribution, click here.

Manipulate Function

Often in controls, it is useful to manipulate certain variables dynamically and see the result in a plot or table. This is particularly useful in setting the constants in tuning PID controllers. The Manipulate function in Mathematica lets you do this. You can use the manipulate function in conjunction with the Plot command. The general syntax for using the manipulate function is as follows:

Manipulate[expr,{u,umin,umax}]

Where *expr* is the function that you want to add controls to and *u* is the variable you wish to manipulate.

Manipulate a function:

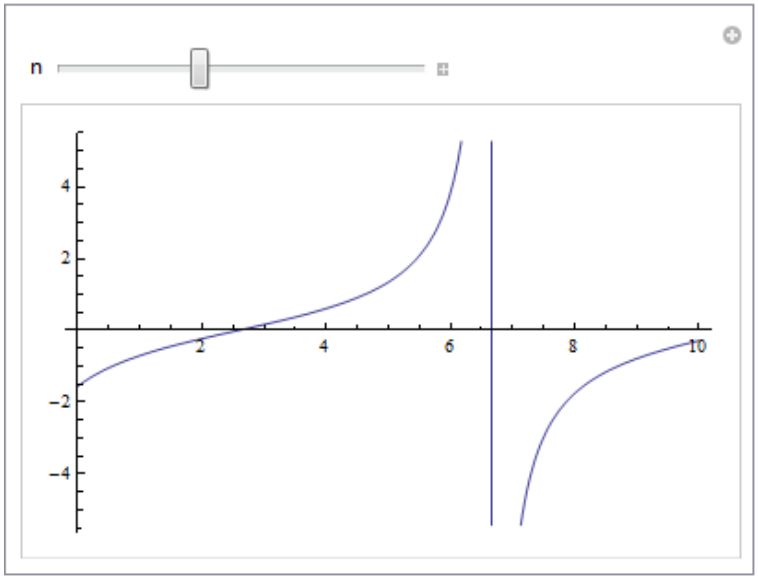
You can change the variable *x* in the function s. Moving the bar lets you change the variable *x*.

```
Manipulate[s = x^3 + 2 x^2 - 3 x + 6, {x, -10, 10, 1}]
```

Manipulate a plot:

You can also manipulate a plot. The slider bar lets change the variable *n* and the plot changes accordingly.

```
Manipulate[Plot[Tan[n * x - 1], {x, 0, 10}], {n, 0, 1}]
```



The Mathematica file used to create this example, can be downloaded below.

Manipulate Command Example

For more information regarding manipulate, please see the Wolfram's Reference Site.

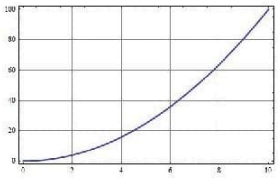
[Note - need to add manipulate for the Bode plots - R.Z.]

Adding Gridlines

```
"PLOTTING WITH GRIDLINES
Make your graphs more readable and attractive!

TO ADD STANDARD GRIDLINES TO YOUR PLOT:
        use GridLines->Automatic

Other gridline options:
IF YOU HAVE A SINUSOIDAL GRAPH:
        use Gridlines->{{Pi,2Pi},{-1,-.5,.5,1}}
IF YOU WANT TO DRAW GRIDLINES AT SPECIFIC PLACES:
        use GridLines->{{x1,x2},{y1,y2}}"

y1 = t^2
Plot[{y1}, {t, 0, 10}, PlotStyle → Thick, PlotRange → Full, Frame → True, GridLines → Automatic]
```

```
PLOTTING WITH GRIDLINES
Make your graphs more readable and attractive!

TO ADD STANDARD GRIDLINES TO YOUR PLOT:
        use GridLines->Automatic

Other gridline options:
IF YOU HAVE A SINUSOIDAL GRAPH:
        use GridLines->{{Pi,2Pi},{-1,-.5,.5,1}}
IF YOU WANT TO DRAW GRIDLINES AT SPECIFIC PLACES:
        use GridLines->{{x1,x2},{y1,y2}}
```

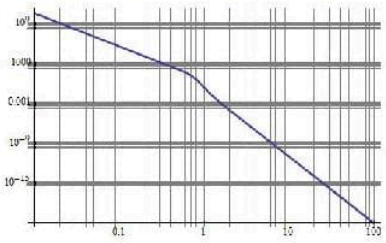
t^2



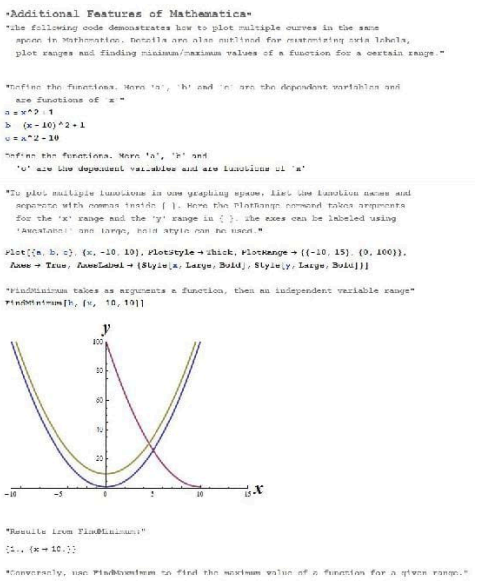
Plotting Log-Log Graphs

```
"CREATING LOG-LOG PLOTS

Sometimes when graphs are in the form of y = axa^b, it is more visually appealing for
  the graph to be plotted on a log-log plot. A log-log plot uses a logarithmic scale
  for both axes. It is so simple to create log log plots in Mathematica; use the simple
  steps below to create beautiful, functional graphs!"

omega = 4 * t^5;
LogLogPlot[AR, {t, 0.01, 100}, PlotStyle → {Thick}, GridLines → Automatic]
```

```
CREATING LOG-LOG PLOTS

Sometimes when graphs are in the form of y = axa^b, it is more visually
  appealing for the graph to be plotted on a log-log plot. A log-log plot uses
  a logarithmic scale for both axes. It is so simple to create log-log plots in
  Mathematica; use the simple steps below to create beautiful, functional graphs!
```



Other Useful Mathematica Tips



[It would be good to replace this images with actual text - R.Z.]

This page titled 2.9: Helpful Mathematica Syntax- Hints on how to use Mathematica to model chemical processes is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

3: SENSORS AND ACTUATORS

3.1: Control Systems - Measurement Devices

3.2: Temperature Sensors

3.3: Pressure Sensors

3.4: Level Sensors

3.5: Flow Sensors

3.6: Composition Sensors

3.7: pH and Viscosity Sensors

3.8: Miscellaneous Sensors

3.9: Valves - Types, Kinds, and Selection

3.10: Valves - Modeling Dynamics

This page titled 3: Sensors and Actuators is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

3.1: Control Systems - Measurement Devices

Control Systems: Industrial Applications

Many control systems are used today in a large number of industries consisting of applications from all kinds. The common factor of all control types is to sustain a desired outcome that may change during a chemical reaction or process. The most common control type used today in industry is a PID controller (proportional, integral, derivative), which allows the operator to apply different control techniques that can be used to achieve different settings in an experiment or process. A PID controller can be used in two main control mechanisms that include feed back and feed forward. The purpose of this article is to provide examples of common industrial control systems that apply different control architectures.

Temperature control: Thermocouple

A thermocouple is a device to measure and control temperature within a system. They are used in a wide variety of industrial applications (gas turbines, chemical reactors, exhaust, chemical manufacturing etc) due to their low cost and portability. The fundamental working principle for thermocouple operations is the Seebeck Effect. Mostly, thermocouples operate in a P or a PID control mode. In order to measure temperature between two points, the thermocouple employs two metallic ends (made from different alloys). When the two conductor ends are exposed to a thermal gradient, they generate a voltage between them. This voltage drop gives rise to the temperature measurement output that a thermocouple provides. Depending on the types of alloys in both conductor ends, and the magnitude of the thermal gradient, thermocouples can measure temperature differences between 1-23000C. They can operate in feedback loops or feed forward loops. Thermocouples are mostly digital control units. Thermocouple prices started around $100 and cost up to $2500 for more accurate, and self-calibrated models.

Pressure Control: Pressure Switch

A pressure switch is a device that controls systems against pressure drops or pressure spikes. The most basic types of pressure switches work on an ON-OFF basis, but can also be manufactured to work in PID mode. The fundamental method of operation is to set the “Set-Pressure” to a given quantity. This deactivates the pressure switch from the circuit connecting it to the control valve upstream. If, at any point during the process, the pressure rises past the set-point, the switch is activated and completes the circuit, thus shutting off the control valve. Pressure switches can be hydraulic, or pneumatic based (air-based pressure). One common application of a pressure switch in the industry is to protect PD (positive-displacement) pumps from over-pressurization. A PD pump can generate very high pressures if not controlled by a pressure switch; thus setting a pressure switch inline with a PD pump will prevent over pressurization since it shuts off the control valve. Pressure switches are common in any industry since all require pressurization of certain components during manufacturing, processing or refining stages. They sell between $200-$2000 depending on the magnitude of the set-point required for protection.

Composition Control: Ratio Control

A ratio controller is used to ensure that two or more process variables such as material flows are kept at the same ratio even if they are changing in value. Control modes can be operated in different types, but mostly feedback PI controller is used for ratio control. In industrial control processes, ratio control is used in the following processes: burner/air ratio, mixing and blending two liquids, injecting modifiers and pigments into resins before molding or extrusion, adjusting heat input in proportion to material flow. The prices for industrial ratio controllers start around $500 and increase depending on the sensitivity of the unit to different magnitudes of compositional changes and size of the equipment to be annexed to.

Level Control: Level Switches

Level controls are used to monitor and regulate the liquid level in industrial vessels. There are many various sensors used in level control including ultrasonic, lasers, and floatation sensors. They all work on the same general principle. A sensor measures the distance from the base of the vessel to the top of the liquid level, mainly by using sonar waves or a laser beam. Based on the time it takes for the wave or for the laser beam to return to the emitting source, the controller sends information to change or maintain the level. The magnetic float control reads on a sensor located on the wall of the vessel and sends that information to the controller. Some examples of level control in industry are: maintaining the liquid level in a distillation column, protecting from overflow, and measuring the amount of product in storage tanks. Level sensors/controls vary in price based on the type and the accuracy required. Below is a picture of a magnetic float level control.

Flow Control: Flow Meters

Flow controls are used to regulate the flow of a liquid or gas through a system. The main form of flow control is a valve. There are many different types of valves, but they all change flow rates by opening or closing based on what amount is needed. A flow sensor reads the flow rate, and a controller will operate the valve to increase or decrease the flow. Some basic types of flow sensors are rotameters and digital mass meters. These controls are used in all forms of industry to control flows including water treatment, product measurement, and fuel flow to furnaces. The low price end for flow sensors is typically around $500, and depending on the size of the process, and the process material, the price can range into the thousands. Below are two pictures. The first is a manual control valve that can be opened or shut to regulate flow. The second is a valve that can be used in an automatic control system to regulate flow.

Contributors and Attributions

- Christian Hatfield, Varun Kaushik, Alon Mandel

This page titled 3.1: Control Systems - Measurement Devices is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

3.2: Temperature Sensors

Introduction

Temperature sensors are vital to a variety of everyday products. For example, household ovens, refrigerators, and thermostats all rely on temperature maintenance and control in order to function properly. Temperature control also has applications in chemical engineering. Examples of this include maintaining the temperature of a chemical reactor at the ideal set-point, monitoring the temperature of a possible runaway reaction to ensure the safety of employees, and maintaining the temperature of streams released to the environment to minimize harmful environmental impact.

While temperature is generally sensed by humans as “hot”, “neutral”, or “cold”, chemical engineering requires precise, quantitative measurements of temperature in order to accurately control a process. This is achieved through the use of temperature sensors, and temperature regulators which process the signals they receive from sensors.

From a thermodynamics perspective, temperature changes as a function of the average energy of molecular movement. As heat is added to a system, molecular motion increases and the system experiences an increase in temperature. It is difficult, however, to directly measure the energy of molecular movement, so temperature sensors are generally designed to measure a property which changes in response to temperature. The devices are then calibrated to traditional temperature scales using a standard (i.e. the boiling point of water at known pressure). The following sections discuss the various types of sensors and regulators.

Temperature Sensors

Temperature sensors are devices used to measure the temperature of a medium. There are 2 kinds on temperature sensors: 1) contact sensors and 2) noncontact sensors. However, the 3 main types are thermometers, resistance temperature detectors, and thermocouples. All three of these sensors measure a physical property (i.e. volume of a liquid, current through a wire), which changes as a function of temperature. In addition to the 3 main types of temperature sensors, there are numerous other temperature sensors available for use.

Contact Sensors

Contact temperature sensors measure the temperature of the object to which the sensor is in contact by assuming or knowing that the two (sensor and the object) are in thermal equilibrium, in other words, there is no heat flow between them.

Examples (further description of each example provide below)

- Thermocouples
- Resistance Temperature Detectors (RTDs)
- Full System Thermometers
- Bimetallic Thermometers

Noncontact Sensors

Most commercial and scientific noncontact temperature sensors measure the thermal radiant power of the Infrared or Optical radiation received from a known or calculated area on its surface or volume within it.

An example of noncontact temperature sensors is a pyrometer, which is described into further detail at the bottom of this section.

Thermometers

Thermometers are the most common temperature sensors encountered in simple, everyday measurements of temperature. Two examples of thermometers are the Filled System and Bimetal thermometers.

Filled System Thermometer

The familiar liquid thermometer consistsof a liquid enclosed in a tube. The volume of the fluid changes as a function of temperature. Increased molecular movement with increasing temperature causes the fluid to expand and move along calibrated markings on the side of the tube. The fluid should have a relatively large thermal expansion coefficient so that small changes in temperature will result in detectable changes in volume. A common tube material is glass and a common fluid is alcohol. Mercury used to be a more common fluid until its toxicity was realized. Although the filled-system thermometer is the simplest and cheapest way to measure temperature, its accuracy is limited by the calibration marks along the tube length. Because filled system thermometers are read visually and don't produce electrical signals, it is difficult to implement them in process controls that rely heavily on electrical and computerized control.

Bimetal Thermometer

In the bimetal thermometer, two metals (commonly steel and copper) with different thermal expansion coefficients are fixed to one another with rivets or by welding. As the temperature of the strip increases, the metal with the higher thermal expansion coefficients expands to a greater degree, causing stress in the materials and a deflection in the strip. The amount of this deflection is a function of temperature. The temperature ranges for which these thermometers can be used is limited by the range over which the metals have significantly different thermal expansion coefficients. Bimetallic strips are often wound into coils and placed in thermostats. The moving end of the strip is an electrical contact, which transmits the temperature thermostat.

Resistance Temperature Detectors

A second commonly used temperature sensor is the resistance temperature detector (RTD, also known as resistance thermometer). Unlike filled system thermometers, the RTD provides an electrical means of temperature measurement, thus making it more convenient for use with a computerized system. An RTD utilizes the relationship between electrical resistance and temperature, which may either be linear or nonlinear. RTDs are traditionally used for their high accuracy and precision. However, at high temperatures (above 700°C) they become very inaccurate due to degradation of the outer sheath, which contains the thermometer. Therefore, RTD usage is preferred at lower temperature ranges, where they are the most accurate.

There are two main types of RTDs, the traditional RTD and the thermistor. Traditional RTDs use metallic sensing elements that result in a linear relationship between temperature and resistance. As the temperature of the metal increases, increased random molecular movement impedes the flow of electrons. The increased resistance is measured as a reduced current through the metal for a fixed voltage applied. The thermistor uses a semiconductor sensor, which gives a power function relationship between temperature and resistance.

RTD Structure

A schematic diagram of a typical RTD is shown in Figure 1.

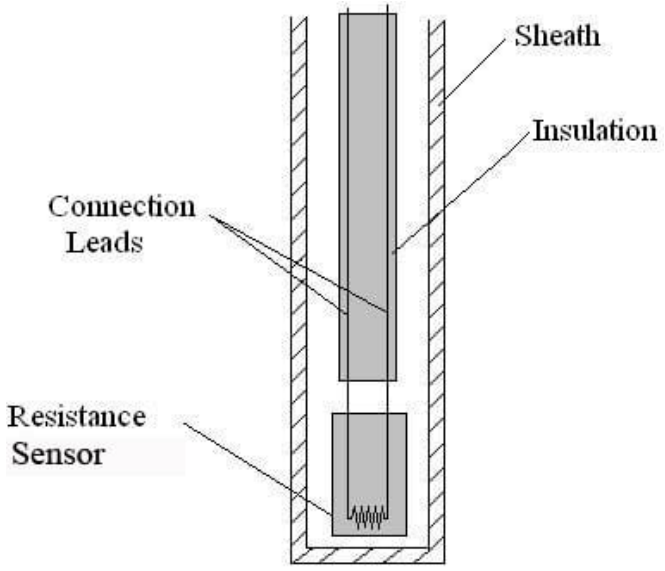


Figure 1. Schematic Diagram of Resistance Temperature Structure

As shown in Figure 1, the RTD contains an outer sheath to prevent contamination from the surrounding medium. Ideally, this sheath is composed of material that efficiently conducts heat to the resistor, but resists degradation from heat or the surrounding medium.

The resistance sensor itself is responsible for the temperature measurement, as shown in the diagram. Sensors are most commonly composed of metals, such as platinum, nickel, or copper. The material chosen for the sensor determines the range of temperatures in which the RTD could be used. For example, platinum sensors, the most common type of resistor, have a range of approximately -200°C – 800°C. (A sample of the temperature ranges and resistances for the most common resistor metals is shown in Table 1). Connected to the sensor are two insulated connection leads. These leads continue to complete the resistor circuit.

Table 1. Common Metal Temperature and Resistance Ranges

Element Metal	Temperature Range	Base Resistance	TCR(Ω/Ω/°C)
Copper	-100 – 260 °C	10 Ω at 0 °C	0.00427
Nickel	-100 – 260 °C	120 Ω at 0 °C	0.00672
Platinum	-260 – 800 °C	100 Ω at 0 °C	0.003916

There are 4 major categories of RTD sensors. There are carbon resistors, film thermometers, wire-wound thermometers and coil elements.

- Carbon resisters are the most commonly used. They are inexpensive and are accurate for low temperatures. They also are not affected by hysteresis or strain gauge effects. They are commonly used by researchers.
- Film thermometers have a very thin layer of metal, often platinum, on a plate. This layer is very small, on the micrometer scale. These thermometers have different strain gauge effects based on what the metal and plate are composed of. There are also stability problems that are dependent on the components used.
- In wire-wound thermometers the coil gives stability to the measurement. A larger diameter of the coil adds stability, but it also increases the amount the wire can expand which increases strain and drift. They have very good accuracy over a large temperature range.
- Coil elements are similar to wire-wound thermometers and have generally replaced them in all industrial applications. The coil is allowed to expand over large temperature ranges while still giving support. This allows for a large temperature range while decreasing the drift.

RTD Operation

Most traditional RTD operation is based upon a linear relationship between resistance and temperature, where the resistance increases with temperature. For this reason, most RTDs are made of platinum, which is linear over a greater range of temperatures and is resistant to corrosion. However, when determining a resistor material, factors such as temperature range, temperature sensitivity, response time, and durability should all be taken into consideration. Different materials have different ranges for each of these characteristics.

The principle behind RTDs is based upon the **Callendar – Van Dusen equation** shown in Equation 3.2.1, which relates the electrical resistance to the temperature in °C. This equation is merely a generic polynomial that takes form based upon experimental data from the specific RTD. This equation usually takes on a linear form since the coefficients of the higher-order variables (a_2, a_3, etc.) are relatively small.

$$R_T = R_0 \left(1 + a_1 T + a_2 T^2 + a_3 T^3 + a_4 T^4 + ... + a_n T^n\right) \tag{3.2.1}$$

R_T: Resistance at temperature T, in ohms
R_0: Resistance at temperature = 0°C, in ohms
a_n: Material's resistance constant, in $°C^{n-1}$

Another type of RTD is the thermistor, which operates based upon an exponential relationship between electrical resistance and temperature. Thermistors are primarily composed of semiconductors, and are usually used as fuses, or current-limiting devices. Thermistors have high thermal sensitivity but low temperature measuring ranges and are extremely non-linear. Instead of the Callendar - Van Dusen equation, the thermistor operates based upon the nonlinear equation, equation (2), shown in degrees K.

$$R_T = R_0 \exp\left(b\left(\frac{1}{T} - \frac{1}{T_0}\right)\right) \tag{3.2.2}$$

T_0: Initial temperature, usually set at 298K
b: Material's temperature coefficient of resistance, in K

Errors associated with resistance thermometers will occur due to the individual or collective efforts of: defective insulation, contamination of the resistor, or insecure lead wire connections.

Thermocouples

Another temperature sensor often used in industry is the thermocouple. Among the various temperature sensors available, the thermocouple is the most widely used sensor. Similar to the RTD, the thermocouple provides an electrical measurement of temperature.

Thermocouple Structure

The thermocouple has a long, slender, rod-like shape, which allows it to be conveniently placed in small, tight places that would otherwise be difficult to reach. A schematic diagram of a typical thermocouple is shown in Figure 2.

Figure 2. Schematic Diagram of Thermocouple Structure

As illustrated in Figure 2, the thermocouple contains an outer sheath, or thermowell. The thermowell protects the contents of the thermocouple from mechanical and chemical damage.

Within the thermowell lies two metal wires each consisting of different metals. Various combinations of materials are possible for these metal wires. Three common thermocouple material combinations used for moderate temperature measurements are the Platinum-Rhodium, Iron-Constantan, and Chromel-Alumel metal alloys. The metal alloys chosen for a thermocouple is based upon the emf value of the alloy pair at a given temperature. Sample emf values for the most common materials at various temperatures are shown in Table 2. For a given pair of materials, the two wires are connected at one end to form a junction. At the other end, the two wires are connected to a voltage measuring device. These ends of the wires are held at a different reference temperature.

Table 2. Common Metal Temperature and Emf Values

Alloy Type	Emf Value at 20 °C	Emf Value at 50 °C	Emf Value at 100 °C
Platinum-Rhodium	0.113 mV	0.299 mV	0.646 mV
Iron-Constantan	1.019 mV	2.585 mV	5.269 mV
Chromel-Alumel	0.798 mV	2.023 mV	4.096 mV

Various methods are used to maintain the reference temperature at a known, constant temperature. One method consists of placement of the reference junction within either an ice bath or oven maintained at a constant temperature. More commonly, the reference temperature is maintained electronically. Though not as stable as an ice bath, electronically controlled reference temperatures are more convenient for use. Reference temperatures could also be maintained through temperature compensation and zone boxes, which are regions of uniform temperature. The voltage difference across the reference junction is measured and sent to a computer, which then calculates the temperature with this data.

Thermocouple Operation

The main principle upon which the thermocouple function is based on is the difference in the conductivities of the two wire materials that the thermocouple is made of, at a given temperature. This conductivity difference increases at higher temperatures and conversely, the conductivity difference decreases at lower temperatures. This disparity results in the thermocouples being more efficient and useful at higher temperatures. Since the conductivity difference is small at lower temperatures and thus more difficult to detect, they are inefficient and highly unreliable at low temperatures.

The conductivity difference between the two wires, along with a temperature difference between the two junctions, creates an electrical current that flows through the thermocouple. The first junction point, which is the point at which the two wires are connected, is placed within the medium whose temperature is being measured. The second junction point is constantly held at a known reference temperature. When the temperature of the medium differs from the reference temperature, a current flows through the circuit. The strength of this current is based upon the temperature of the medium, the reference temperature, and the materials of the metal wires. Since the reference temperature and materials are known, the temperature of the medium can be determined from the current strength.

Error associated with the thermocouple occurs at lower temperatures due to the difficulty in detecting a difference in conductivities. Therefore, thermocouples are more commonly used at higher temperatures (above -125°C) because it is easier to detect differences in conductivities. Thermocouples are operable over a wide range of temperatures, from -200°C to 2320°C, which indicates its robustness and vast applications. Thermocouples operate over this wide range of temperatures, without needing a battery as a power source. It should be noted that, the wire insulation might wear out over time by heavy use, thus requiring periodical checks and maintenance to preserve the accuracy of the thermocouple.

To determine the temperature of the medium from the current strength, the emf or voltage values of the current and of the wire materials at the reference temperatures must be known. Often, the measured temperature can be found by using standard thermocouple tables. However, these tables are often referenced at 0°C. To correct for this different reference temperature, equation (3) can be used to calculate the temperature from a given current.

$$\xi_{T_1,T_3} = \xi_{T_1,T_2} + \xi_{T_2,T_3} \tag{3.2.3}$$

ξ : emf of an alloy combination generated at two different temperatures
T_1: temperature of the medium whose temperature is to be determined
T_2: reference temperature of the thermocouple
T_3: reference temperature of the standard thermocouple table, which in this case is 0°C

Once the emf between two alloys is calculated relative to a reference temperature when T_3 is 0°C, the standard thermocouple table can be used to determine the temperature T1 of the medium. This temperature is usually automatically displayed on the thermocouple.

Apart from the common occurrence of the thermocouples being placed in the fluid to measure temperature change, thermocouples can be also embedded in solids with excellent results. This is highly effective while establishing the different thermal properties for a solid. The heat transfer to the thermocouple will now be in the form of conductive heat transfer. As a result, this setup would be very similar to heat conduction in series, since the thermocouple is almost always made from a different material then the actual solid. Such discrepancies depend on the manner in which the thermocouple is embedded in the solid and should be taken into account when the thermal properties are being calculated and analyzed. One example is shown in the photo below.

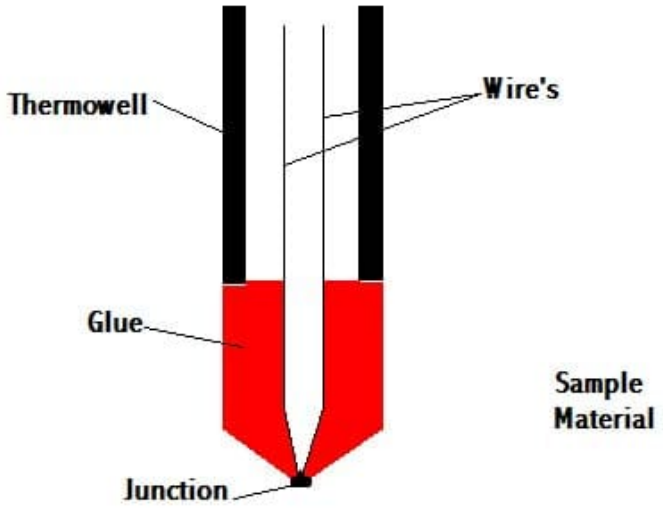


Figure 3. Schematic diagram of how the thermocouple function

Laws for thermocouples

- Law of homogenous material: If all the wires and the thermocouple are made of the same material, temperature changes in the wiring do not affect the output voltage. Thus, need different materials to adequately reflect the temperature.
- Law of intermediate materials: The sum of all the thermoelectric forces in a circuit with a number of dissimilar materials at a uniform temperature is zero. This implies that if a third material is added at the same temperature, no net voltage is generated by the new material.
- Law of successive or intermediate temperatures: If two dissimilar homogeneous materials produce thermal emf1 when the junctions are at T1 and T2 and produce thermal emf2 when the junctions are at T2 and T3 , the emf generated when the junctions are at T1 and T3 will be emf1 + emf2 .

Application

- Steel industry: Monitor temperature and chemistry throughout the steel making process
- Heating appliance safety: Thermocouples in fail-safe mode are used in ovens and water heaters to detect if pilot flame is burning to prevent fire and health hazard
- Manufacturing: Used for testing prototype electrical and mechanical apparatus
- Process plants: Chemical production plants and refineries use computer programs to view the temperature at various locations. For this situation, a number of thermocouple leads are brought to a common reference block.

Pyrometers

Unlike the thermometer, RTD and the thermocouple, pyrometers (non-contact temperature sensors) measures the amount of heat radiated, rather than the amount of heat conducted and convected to the sensor. Various types of pyrometers, such as total radiation and photoelectric pyrometers, exist. Below is a schematic of an optical pyrometer in Figure 4.

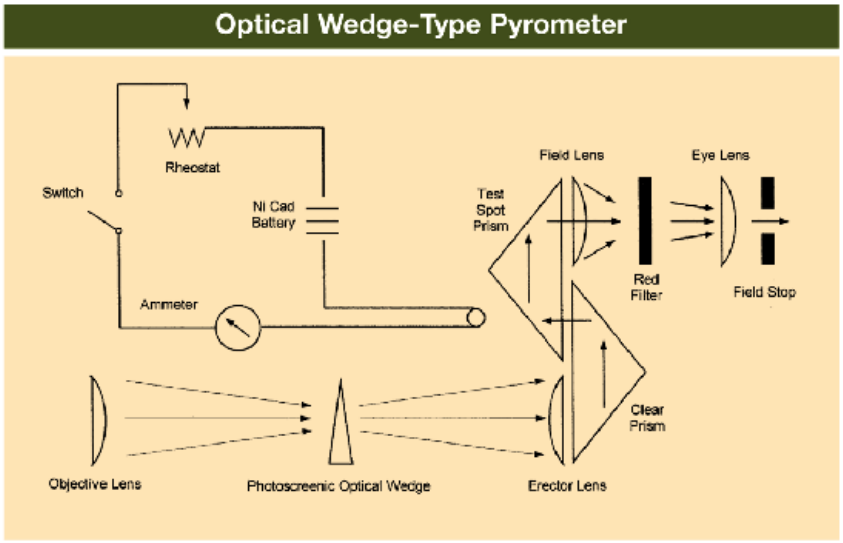


Figure 4. Schematic diagram of an optical pyrometer

These pyrometers differ in the type of radiation they measure. There are many factors that influence the amount of radiated heat detected, thus there are many assumptions that must be made regarding the emissivity, or the measure of the manner in which heat is radiated, of the object. These assumptions are based upon the manner in which heat is radiated as well as the geometry of the object. Because temperature is dependent on the emissivity of a body, these assumptions regarding the emissivity introduce uncertainties and inaccuracies in the temperature readings. Therefore, because of the error associated with them, pyrometers are not often used in industry.

Table 3. Summary of Temperature Sensors

	Type	Function	Operating Range (C)	Advantage	Disadvantage
Thermometer	Liquid-filled	Calibrated with reference to a standard temperature	-200 C to 250 C	Cheap, good for everyday usage	Limited range, not connected to temperature regulation system, visual readings introduce human error
	Filled			Good for checks on temperatures	
RTD	Platinum (also called resistance sensing element);	Provides very precise measurements, making it applicable to the semiconductor system	-200 C to 7,000 C	Very precise readings, provides readings in both digital and analog, stable for long periods of time.	Small bias error because the thermocouples are used to measure small voltages, sensitive to environmental disturbances
	Thermistor (also called semiconductor sensor)				
Thermocouple		Semiconductor materials are used to fabricate diodes and transistors	-200 C to 2320 C depending on the metal combination the junction is made of	Relatively rugged, can be used in many temperature ranges	Can require accuracy and must be used for a long time due to thermoelectric inhomogeneities

There are a few different types of pyrometers. There are optical and radiation pyrometers.

How Optical Pyrometers Work:

- Compares the color of visible light given off by the object with that of a electrically heated wire
- The wire can be preset to a certain temperature
- The wire can be manually adjusted to compare the two objects

How Radiation Pyrometers Work:

- This sensor works by measuring the radiation (infrared or visible light) that an object gives off
- The radiation heats a thermocouple in the pyrometer which in turn induces a current
- The larger the current induced, the higher the temperature is

Pyrometers are usually used at very high temperatures, but can be used at colder temperatures as well. There are lots of industrial applications to pyrometers. Plant operators can use pyrometers to get a sense of what temperature certain processes are running at. The downside to pyrometers is that they are not very accurate as thermocouples or RTD sensors are. This is because they rely on quantifying colors of light.

Temperature Regulators

Temperature regulators, also known as temperature control valves (TCVs), physically control, as well as measure, temperature. Temperature regulators are not capable of directly maintaining a set value; instead, they relate the load (in this case the valve opening) with the control (temperature measurement). These regulators are most useful when temperature is correlated to a flow of a substance. For example, a TCV may be used to control the temperature of an exothermic reaction that requires constant cooling. The TCV measures the temperature of the reaction and, based upon this temperature, either increases or decreases the flowrate of cooling fluid to adjust the temperature of the reaction. Similarly, the regulator could be used to adjust the flow amount of steam, which is typically used to heat a substance. Therefore, by adjusting flowrate, the regulator can indirectly adjust temperature of a given medium.

Regulator Structure

The structure of a typical thermal regulator consists of four main parts, as shown in Figure 3. The temperature detecting element, which in most cases is a temperature sensor, as described above, sends either an electrical or mechanical signal through the connector to the actuator. The actuator then uses this signal to act upon the power source, which determines the position of the valve. (This will be further described in the next section.)

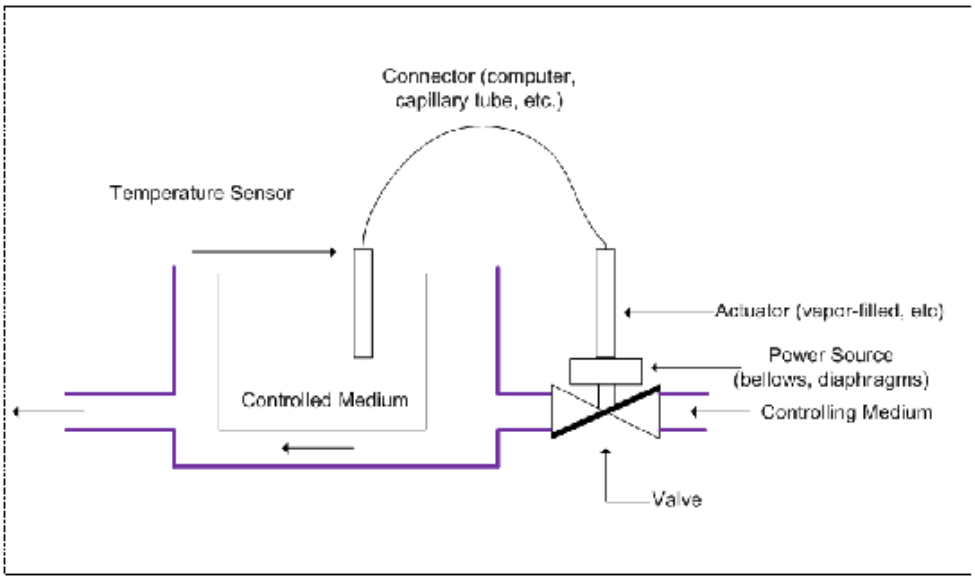


Figure 3. Schematic Diagram of Temperature Regulator Structure. Note: The actuator consists of the capillary tubing, the power source, and the adjustment.

Regulator Operation

The temperature regulator operates based upon a mechanical means of temperature control. As previously mentioned, the bulb of the regulator is typically filled with a heat conducting substance. Due to the thermal expansion properties of this substance, the substance expands as the temperature increases. This expansion causes a change in the pressure of the actuator, which correlates to the temperature of the medium. This pressure change repositions a valve on the regulator, which controls the flowrate of a coolant. The temperature of the medium is then altered by the change in the flowrate of this coolant.

Types of Temperature Regulators

Though all regulators have the same basic build and purpose, they exist in a variety of forms. In particular, these regulators vary

in four primary ways: temperature detecting elements, temperature detector placement, actuator type, and valve type.

Temperature Detecting Elements

Most temperature regulation systems use thermocouples or RTDs as temperature sensing devices. (Described above) For these systems, the connector is a computer. The sensors send an electric signal to the computer, which calculates the temperature. The computer then compares the temperature measured by the sensor to a programmed set-point temperature, thus determining the required pressure in the actuator. The pressure in the actuator changes position of the power source (diaphragm or bellows), which consequently changes the flowrate through the valve.

Some temperature regulation systems use a filled bulb as a temperature sensor. Based on the thermal expansion properties of the material within the bulb, the material expands as the temperature increases. This expansion causes a change in the pressure of the actuator. The pressure change actuator then repositions power source. Again, the change in the power source changes the flowrate through the valve.

Temperature regulation systems using thermocouples or RTDs as temperature sensing devices are much more common than regulation systems using filled bulbs.

Temperature Detector Placement: Internal and Remote Detection

Temperature detection can be done with internal or remote elements. For internal temperature detectors, the thermal actuator and temperature detector are located entirely within the valve. For remote temperature detectors, the primary temperature detecting element is separate from the actuator and valve, and is connected to the actuator with either electrical wiring or capillary tubing, depending on the mechanism of the temperature sensor. Remote temperature detectors are more common, as internal temperature detectors are limited in use. Internal temperature detectors can only measure the temperature of the fluid flowing through the valve and not the temperature of the process.

Actuator Type: Thermal Systems

There are four main categories of thermal actuators used in temperature regulators. Thermal actuators produce power and work, proportional to the measured temperature of the process, on the power source. Actuator types include the vapor-filled system, the liquid-filled system, the hot chamber system, and the fusion-type or wax-filled system. Of all the thermal systems mentioned, liquid-filled systems are the most common, because they relate temperature and pressure change in a linear fashion.

Vapor-Filled Systems

In the vapor-filled system, the thermal actuator is partially filled with a volatile liquid. As the temperature of the sensor increases, the vapor pressure of the liquid also increases. This increases the pressure on the power source, and adjusts the flowrate through the valve.

Liquid-Filled Systems

In liquid-filled systems, the thermal actuator is filled with a chemically stable liquid, such as a hydrocarbon. As the temperature increases, the liquid expands, which produces a force on the power source.

Hot Chamber Systems

In hot chamber systems, the thermal actuator is partially filled with a volatile fluid. An increase in temperature of the system forces some of this fluid into the power unit, where the heat of the unit causes this liquid to turn into a superheated vapor. The pressure increase produces a force on the power source.

Fusion-Type (Wax-Filled) Systems

Of all the systems mentioned, the fusion-type system is the least common. In the fusion-type system, the thermal actuator is filled with special waxes such as hydrocarbons, silicones, and natural waxes. The wax contains large amounts of copper, which increases the heat-transfer quality of the wax. As temperature increases, the wax expands, producing a force that repositions the power source.

Valve Type: Direct and Pilot Actuated

The two main types of valves used in thermal regulators are the direct and pilot actuated valves. In all such thermal regulators, there is a power source (such as bellows and diaphragms) that provides the force required to reposition the valve to control the temperature. These power sources rely on a change in the pressure of the actuator in order to properly regulate temperature. In direct-actuated TCVs, this power unit is directly connected to the valve, which provides the force required to open and close the valve. In pilot-actuated TCVs, the thermal actuator moves a pilot valve, which then transfers energy in the form of pressure to a piston, which then provides the work necessary to reposition the main valve.

Direct-actuated TCVs are often much simpler in structure than pilot-actuated TCVs, and therefore they are also much cheaper. In addition, they respond better to smaller changes in temperature and more accurately reflect the temperature of the medium. Thus, if the exact temperature of the system is essential to ensure correct operation, a direct-actuated TCV should be used. Pilot-actuated TCVs usually have much smaller temperature sensing devices, a faster response time, and the ability withstand much higher pressures through the regulating valve. Therefore, at high pressures or rapid temperature changes, a pilot-actuated TCV should be used.

Resistance Temperature Detector Example

Note: This example problem has been made up for demonstration purposes.

A newly hired chemical engineer at Hypothetical Industries is responsible for monitoring and maintaining temperatures for one of the company's exothermic reactions. The process uses a platinum resistor thermometer to measure the temperature of the process, the properties of which are given below. The reaction's ideal range is between 250°C – 350°C. Below 250°C, the catalyst ceases to function, and above 350 °C, it can be classified as a runaway reaction. The engineer can control steam and cooling water to regulate the process temperature.

Describe what changes the engineer should make to either the flowrate of the coolant or the steam to keep the system functioning optimally.

Case I: $R_T = 25 ohms$

Case II: $R_T = 13.9 ohms$

Case III: $R_T = 19.4 ohms$

GIVEN DATA:

$R_T = R_0(1 + a_1T + a_2T^2)$

$R_0 = 10$ ohms
$a_1 = 3.91x10^{-3}(°C)$
$a_2 = -6.72x10^{-8}(°C^{-2})$

Solution:

Substitute the given value of R_T for each of the 3 cases into the mathematical calculations below to get the answers, we have worked out Case I with the corresponding numbers below.

Case I:

1. Factor R_0 throughout the right side of the equation.
2. Subtract R_T from both sides of the equation.
3. Solve for T using the quadratic equation.

$$25 = 10(1 + \ 3.19 \times 10^{-3}T + \ -6.72 \times 10^{-8}T^2)$$
$$25 = 10 + \ 3.19 \times 10^{-2}T + \ -6.72 \times 10^{-7}T^2$$

$$0 = -15 + \ 3.19 \times 10^{-2}T + \ -6.72 * 10^{-7}T^2$$

$$T = \frac{-3.19 \times 10^{-2} \pm \sqrt{(3.19 \times 10^{-2})^2 - 4 \times (-6.72 \times 10^{-7}) \times (-15)}}{(2 \times -6.72 \times 10^{-7})}$$

$$T = 474.97^{\circ}C$$

ALTERNATE Solution:

1. Since the constant a_2 is so small ($x10^{-7}$) we can just neglect that term.
2. Now the problem is just a simple linear equation, which can be solved by isolating T on one side of the equation.

$$25 = 10(1 + \ 3.19 \times 10^{-3}T)$$

$$25 = 10 + \ 3.19 \times 10^{-2}T$$

$$15 = 3.19 \times 10^{-2}T$$

$$T = \frac{15}{3.19 \times 10^{-2}}$$

$$T = 470.21^{\circ}C$$

AnswerS:

Case I. The engineer should start flowing cooling water into the reaction process because the temperature is ~ 500°C which is above the appropriate range.
Case II. The engineer should increase steam feed to the reaction process because the temperature is ~ 125°C which is below the appropriate range.
Case III. The engineer does not have to do anything because the temperature is in the appropriate range ~ 300°C.

Example 3.2.1: Temperature Sensor

You are a chemical engineer at Hypothetical Industries and you are responsible for monitoring and regulating the temperature for one of the company's reactions. Determine which temperature sensor should be used to optimally measure and regulate the reaction's temperature in each of the following situations.

- Case I. T=900°C
- Case II. T=500°C but sensor location in a large continuous reactor makes it difficult to repair
- Case III. T=50°C and you are estimating the current temperature of a lab scale reaction

Solution

Case I. We know that the reaction occurs at 900°C. Taking this optimal temperature as the only necessary parameters, the temperature sensor would be a thermocouple. The temperature is outside of the range for resistance thermometers. Regular thermometers do not send electric signals; therefore they can not be used for this process.
Case II. Although this temperature is within the operating ranges of both thermocouples and RTDs, since the sensor may not be readily accesible for repair, we must use an RTD because thermocouples lose accuracy after extended usage.
Case III. Since the temperature is within the range of all three sensors, and we only need a rough estimate of temperature, we can use a thermometer which will be much less costly than the alternative sensors.

References

- Dougherty, E.P. ,*Temperature Control Principles for Process Engineers*, Hanser Publishers, 1993 ISBN 0344615980
- Liptak,B.G. ,*Instrument Engineer's Handbook*, 4th Edition,Volume 2. Taylor and Francis Group,2005 ISBN 0849310814
- Perry,R.H. ,*Perry's Chemical Engineer's Handbook*, 7th Edition, McGraw-Hill,1997 ISBN 0070498415
- Richardson,J.F. and Peacock,D.G. ,*Chemical Engineering*, 3rd Edition,Volume 3. Butterworth-Heinemann,1994 ISBN 0080410030
- Moffat,R.J. ,Notes on Using Thermocouples, 1997
- **Error! Hyperlink reference not valid.**
- Pyrosales
- Cole parmer Suppliers
- Temperature.com

Contributors and Attributions

- Authors: (14 September 2006) Ardemis Boghossian, James Brown, Sara Zak
- Stewards: (11 September 2007) Arthur L. Edge III, Kathryn Meintel, Renu Rao, Kaveh Saba

This page titled 3.2: Temperature Sensors is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

3.3: Pressure Sensors

Introduction

Pressure must be considered when designing many chemical processes. Pressure is defined as force per unit area and is measured in English units of psi and SI units of Pa. There are three types of pressure measurements:

1. Absolute pressure - atomospheric pressure plus gauge pressure.
2. Gauge Pressure - absolute pressure minus atmospheric pressure.
3. Differential Pressure - pressure difference between two locations.

There are various types of pressure sensors that are available in the market today for use in industry. Each functions best in a certain type of situation.

Sensor Selection Criteria

In order for a pressure controlled system to function properly and cost-effectively, it is important that the pressure sensor used be able to give accurate and precise readings as needed for a long period of time without need for maintenance or replacement while enduring the conditions of the system. Several factors influence the suitability of a particular pressure sensor for a given process: the characteristics of the substances being used or formed during the process, the environmental conditions of the system, the pressure range of the process, and the level of precision and sensitivity required in measurements made.

Process

The pressure sensing element (elastic element) will be exposed to the materials used in the process, therefore materials which might react with the process substances or degrade in corrosive media are unsuitable for use in the sensor. Diaphragms are optimal for very harsh environments.

Environment

The environment (as in the system -- pipes, vibrations, temperature, etc.) in which the process is carried out also needs to be considered when choosing a suitable pressure sensor. Corrosive environments, heavy vibrations in the piping units, or extreme temperatures would mean that the sensors would have to have an added level of protection. Sealed, strong casing materials with interior liquid containing glycerine or silicone are often used to encase the internal components of the sensor (not including the sensing element), protecting them from very harsh, corrosive environments or frequent vibrations.

Pressure Range

Most processes operate within a certain pressure range. Because different pressure sensors work optimally in different pressure ranges, there is a need to choose pressure gauges which are able to function well in the range dictated by the process.

Sensitivity

Different processes require different levels of precision and accuracy. In general, the more precise the sensor, the more expensive it is, thus it is economically viable to choose sensors that are able to satisfy the precision desired. There is also a compromise between precision and the ability to detect pressure changes quickly in sensors, hence in processes in which the pressure is highly variable over short periods of time, it is unadvisable to use sensors which take a lot of time to give accurate pressure readings, although they might be able to give precise pressure values.

Pressure Measuring Methods

Several pressure measuring methods have been developed and utilized; these methods include visual inspection of the height of liquid in a column, elastic distortion, and electrical methods.

Height of Liquid in Column

The height of a liquid with known density is used to measure pressure. Using the equation $P = \rho gh$, the gauge pressure can be easily calculated. These types of pressure measuring devices are usually called manometers. Units of length may be used to measure the height of the liquid in the column as well as calibrated pressure units. Typically water or mercury is used as the liquid within these columns. Water is used when you desire greater sensitivity (its density is much less than liquid mercury, so its height will vary more with a pressure change). Mercury is used when you desire higher pressure measurements and not as great sensitivity.

Elastic Distortion

This pressure measuring method is based on the idea that deformation of an elastic material is directly proportional to the pressure being measured. There are mainly three sensor types that are used in this method of measuring pressure: Bourdon-tubes, diaphragms and bellows. (refer to “Types of Sensors” Section)

Electrical Methods

Electrical methods used for measuring pressure utilize the idea that dimensional changes in a wire affect the electrical resistance to the conductor. These devices that use the change in resistance of the wire are called strain gauges. Other electrical sensors include capacitive sensors, inductive pressure transducers, reluctive pressure transducers, potentiometric transducers, piezoresistive transducers and piezoelectric sensors. (refer to “Types of Sensors” Section)

Types of Sensors

There are many different pressure sensors to choose from when considering which is most suitable for a given process, but they can generally be placed into a few categories, namely elastic sensors, electrical transducers, differential pressure cells and vacuum pressure sensors. Listed below each general category are specific internal components, each functioning best in a certain situation.

Elastic Sensors

Most fluid pressure sensors are of the elastic type, where the fluid is enclosed in a small compartment with at least one elastic wall. The pressure reading is thereby determined by measuring the deflection of this elastic wall, resulting in either a direct readout through suitable linkages, or a transduced electrical signal. Elastic pressure sensors are sensitive; they are commonly fragile and susceptible to vibration, however. In addition, they tend to be much more expensive than manometers, and are therefore preferentially used for transmitting measured data and measuring pressure differences. A wide variety of flexible elements could conceivably be used for elastic pressure sensors; the majority of devices use some form of a Bourdon tube, bellows, or diaphragm.

Bourdon Tube Gauges

The principle behind all Bourdon tubes is that an increase in pressure on the inside of the tube in comparison to the outside pressure causes the oval or flat shaped cross-section of the tube to try to achieve a circular shape. This phenomenon causes the tube to either straighten itself out in the c-type or spiral cases or to unwind itself for the twisted and helical varieties. This change can then be measured with an analog or digital meter connected to the tube. Tube materials can be changed accordingly to suit the required process conditions. Bourdon tubes can operate under a pressure range from 0.1-700 MPa. They are also portable and require little maintenance; however,they can only be used for static measurements and have low accuracy.

Types of Bourdon tubes include C-type, spiral (a more coiled C-type tube), helical and straight tube Bourdon tubes. C-type gauges

can be used in pressures approaching 700MPa; they do have a minimum recommended pressure range, though -- 30 kPa (i.e., it is not sensitive enough for pressure differences less than 30 kPa).

Bellows

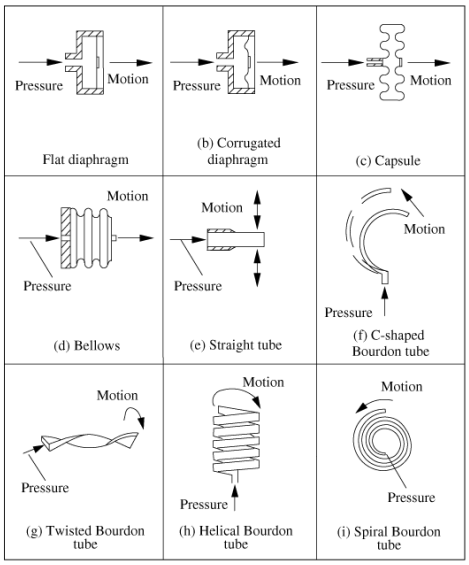
Bellows elements are cylindrical in shape and contain many folds. They deform in the axial direction (compression or expansion) with changes in pressure. The pressure that needs to be measured is applied to one side of the bellows (either inside or outside) while atmospheric pressure is on the opposite side. Absolute pressure can be measured by evacuating either the exterior or interior space of the bellows and then measuring the pressure at the opposite side. Bellows can only be connected to an on/off switch or potientiomenter and are used at low pressures, <0.2 MPa with a sensitivity of 0.0012 MPa.

Diaphragms

Diaphragm elements are made of circular metal discs or flexible elements such as rubber, plastic or leather. The material from which the diaphragm is made depends on whether it takes advantage of the elastic nature of the material, or is opposed by another element (such as a spring). Diaphragms made of metal discs utilize elastic characteristics, while those made of flexible elements are opposed by another elastic element. These diaphragm sensors are very sensitive to rapid pressure changes. The metal type can measure a maximum pressure of approximately 7 MPa, while the elastic type is used for measuring extremely low pressures (.1 kPa - 2.2 MPa) when connected to capacitative transducers or differential pressure sensors. Examples of diaphragms include flat, corrugated and capsule diaphragms. As previously noted, diaphragms are very sensitive (0.01 MPa) . They can measure fractional pressure differences over a very minute range (say, inches of water) (elastic type) or large pressure differences (approaching a maximum range of 207 kPa) (metal type).

Diaphragm elements are very versatile -- they are commonly used in very corrosive environments or with extreme over-pressure situations.

Examples of these elastic element pressure sensors are shown here.



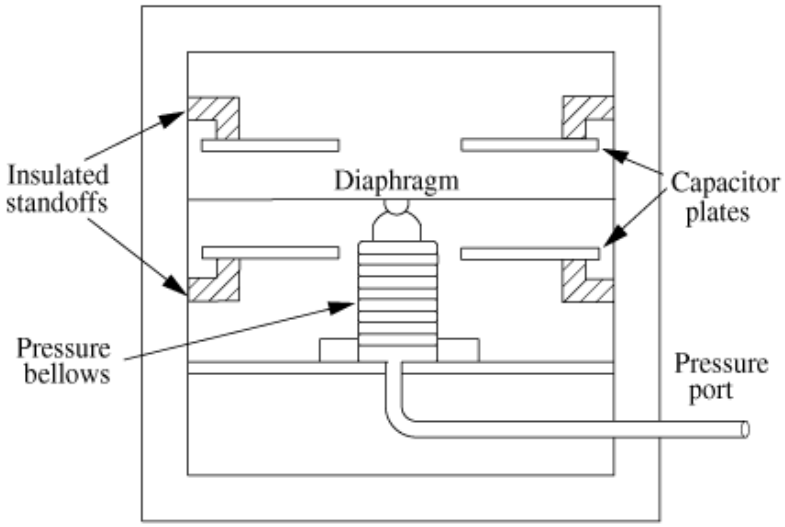
Electric Sensors

Sensors today are not neccessarily only connected to a gauge meter needle pointer to indicate pressure, but may also serve to convert the process pressure into an electrical or pneumatic signal, which can be transmitted to a control room from which the pressure reading is determined. Electric sensors take the given mechanics of an elastic sensor and incorporate an electrical component, thus heightening the sensitivity and increasing the amount of instances in which you could utilize the sensor. The types of pressure transducers are capacitive, inductive, reluctive, piezoelectric, strain gauge, vibrating element, and potentiometric.

Capacitive

A capacitive sensor consists of a parallel plate capacitors coupled with a diaphragm that is usually metal and exposed to the process pressure on one side and the reference pressure on the other side. Electrodes are attached to the diaphragm and are charged by a high frequency oscillator. The electrodes sense any movement of the diaphragm and this changes the capacitance. The change of the capacitance is detected by an attached circuit which then outputs a voltage according to the pressure change. This type of sensor can be operated in the range of 2.5 Pa - 70MPa with a sensitivity of 0.07 MPa.

An example of a capacitive pressure sensor is shown on the right.

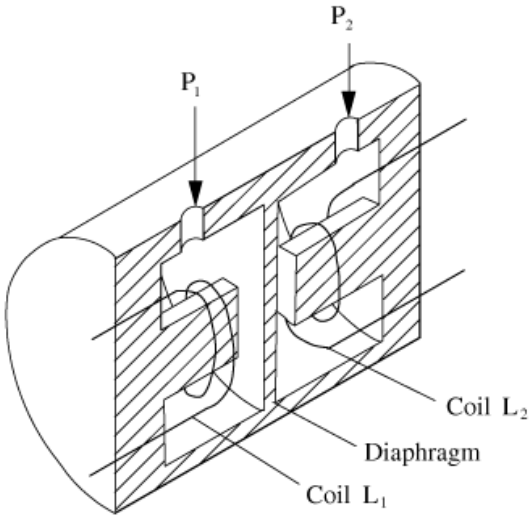


Inductive

Inductive pressure sensors are coupled with a diaphragm or a Bourdon tube. A ferromagnetic core is attached to the elastic element and has a primary and 2 secondary windings. A current is charged to the primary winding. When the core is centered then the same voltage will be induced to the two secondary windings. When the core moves with a pressure change, the voltage ratio between the two secondary windings changes. The difference between the voltages is proportional to the change in pressure.

An example of an inductive pressure sensor utilizing a diaphragm is shown below. For this kind of pressure sensor, taking Chamber 1 as the reference chamber with a reference pressure P_1 coming into the chamber and the coil being charged with a reference current. When the pressure in the other chamber changes, the diagphragm moves and induces a current in the other coil, which is measured and gives a measure of the change in pressure.

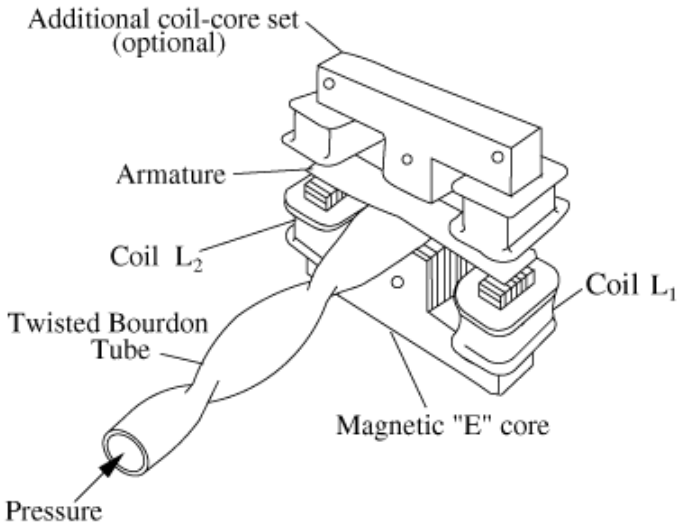
These may be used with any elastic element (though, it is typically coupled with a diaphragm or a bourdon tube). The pressure reading generated will be determined by voltage calibration. Thus, the range of pressure in which this sensor may be used is determined by an associated elastic element but falls in the range of 250 Pa - 70 MPa .



Reluctive

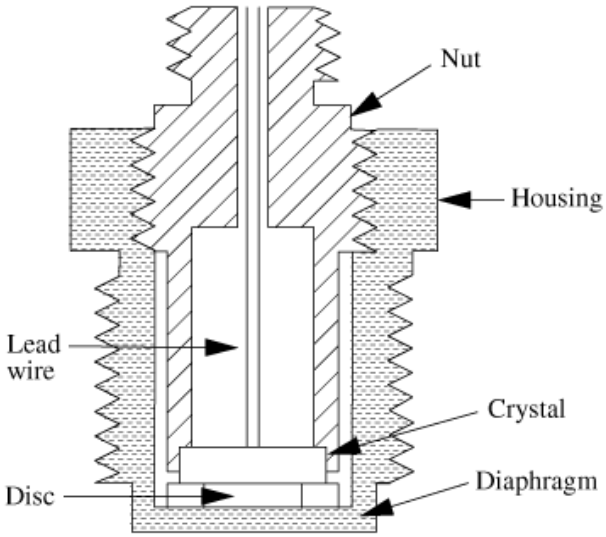
Reluctive pressure sensors also charge a ferromagnetic core. When the pressure changes, the flexible element moves a ferromagnetic plate, leading to a change in the magnetic flux of the circuit which can be measured. The situations in which one would use a reluctive electric element is one in which the inductive sensor does not generate a precise enough measurement. The pressure range is 250 Pa - 70 MPa with a sensitivity of 0.35 MPa.

An example of a reluctive pressure sensor can be seen on the right.



Piezoelectric

Piezoelectric sensors use a crystal sensor. When pressure is applied to the crystal, it deforms and a small electric charge is created. The measurement of the electric charge corresponds to the change in pressure. This type of sensor has a very rapid response time to constant pressure changes. Similar to reluctive electric element, the piezoelectric element is very sensitive, but responds much, much faster. Thus, if time is of the essence, a piezoelectric sensor would be desired. The pressure range is .021 - 100 MPa with a sensitivity of 0.1 MPa.

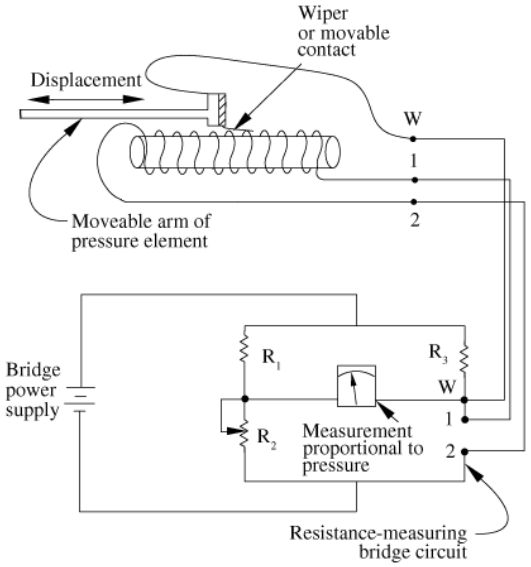


On the right is an example of a piezoelectric pressure sensor.

Potentiometric

Potentiometric sensors have an arm mechanically attached to the elastic pressure sensing element. When pressure changes, the elastic element deforms, causing the arm to move backwards or forwards across a potentiometer and a resistance measurement is taken. These sensing elements do posess an optimum working range, but are seemingly limited in their resolution by many factors. As such, they are low end sensors that aren't used for much. With a low sensitivity and working range, they may be best suited as a cheap detector evaluating a coarse process.The pressure range is 0.035 - 70 MPa with a sensitivity of 0.07 -0.35 MPa.

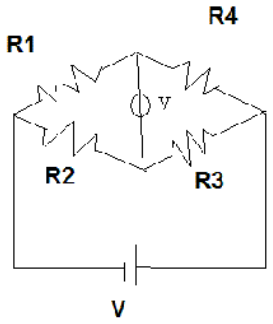
An example of a potentiometric pressure sensor is shown on the right.



Strain Gauge

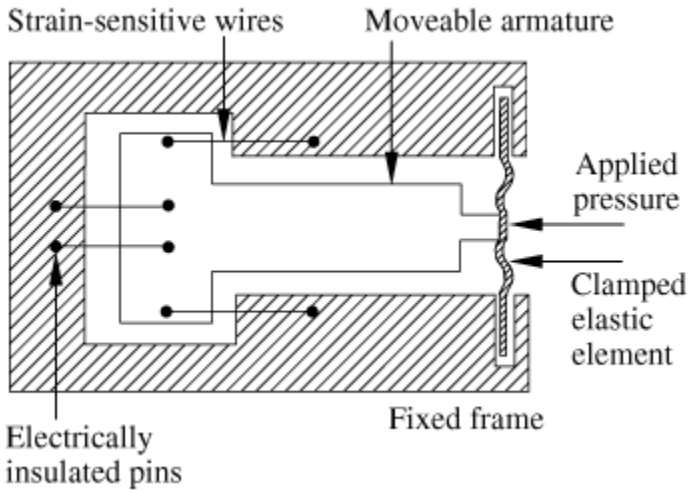
The strain gauge detects changes in pressure by measuring the change in resistance of a Wheatstone bridge circuit. In general, this circuit is used to determine an unknown electrical resistance by balancing two sections of a bridge circuit such that the ratio of resistances in one section ($\frac{R_3}{R_2}$) is the same as that in the other section($\frac{R_4}{R_1}$), resulting in a zero reading in the galvanometer

in the center branch. One of the sections contains the unknown component of which the resistance is to be determined, while the other section contains a resistor of known resistance that can be varied. The Wheatstone bridge circuit is shown below:

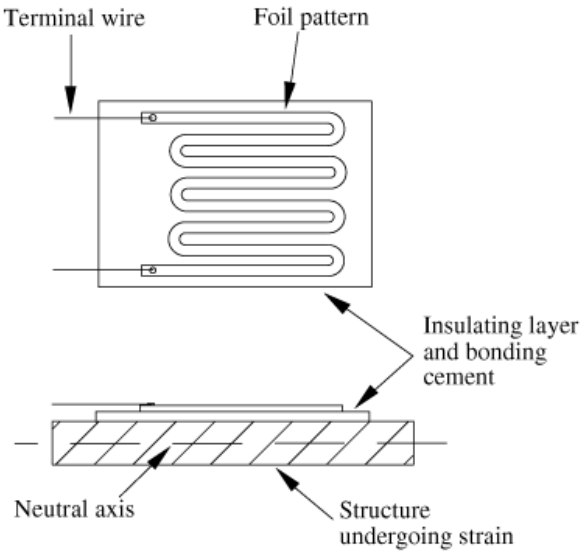


The strain gauge places sensors on each of the resistors and measures the change in resistance of each individual resistor due to a change in pressure. Resistance is governed by the equation $R = \rho \frac{L}{A}$ where ρ = resistivity of the wire, L = length of the wire, and A = cross-sectional area of the wire. A pressure change would either elongate or compress the wire, hence a compression sensor is needed on one resistor and an elongation sensor on the other. To control the effects of temperature (a wire would also either elongate or compress with a change in temperature), a blank sensor would be placed on the remaining two resistors. These gauges are frequently a type of semiconductor (N-type or P-type). Thus, their sensitivity is much greater than their metal counterparts; however, with greater sensitivity comes a more narrow functional range: the temperature must remain constant to obtain a valid reading. These gauges are affected greatly by variations in temperature (unlike the other types of electrical components). The pressure range is 0 - 1400 MPa with a sensitivity of 1.4 - 3.5MPa.

An example of an unbonded strain gauge is shown below. This makes use of strain-sensitive wires one end fixed to an immobile frame and the other end attached to a movable element, which moves with a change in pressure.

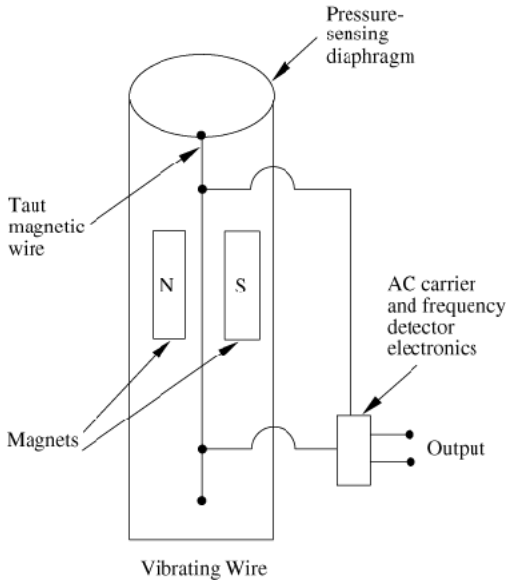


An example of a bonded strain gauge can be seen below. This is placed on top of a diaphragm, which deforms with change in pressure, straining the wires attached to the diaphragm.

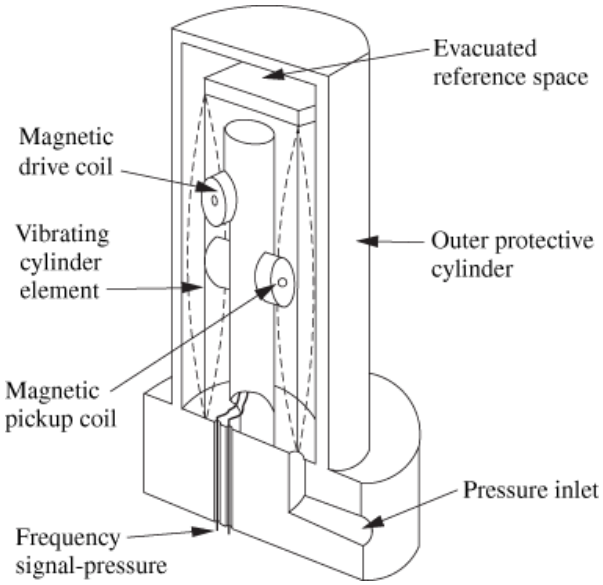


Vibrating Element

Vibrating element pressure sensors function by measuring a change in resonant frequency of a vibrating element. A current is passed through a wire which induces an electromotive force within the wire. The force is then amplified and causes oscillation of the wire. Pressure affects this mechanism by affecting the wire itself: an increase in pressure decreases the tension within the wire and thus lowers the angular frequency of oscillation of the wire. The sensor is housed in a cylinder under vacuum when measuring absolute pressures. These absolute pressure measuring sensors are very efficient: they produce repeatable results and are not affected by temperature greatly. They lack sensitivity in meausurement, though, so they would not be ideal for a process in which minute pressures need monitoring. The pressure range is 0.0035 - 0.3 MPa with a sensitivity of 1E-5 MPa.



A vibrating wire pressure sensor is shown below.
A vibrating cylinder pressure sensor (for absolute pressures) is shown below.



Differential Pressure Cells

Differential pressure cells are used with various kinds of sensors in which measurement of pressure is a result of a pressure differential such as orifice plates, flow nozzles, or venturi meters. The differential pressure cell converts the pressure differential into a transmittable signal. Where the differential pressure (DP) cell is placed depends on the nature of the fluid stream that is being measured. A typical DP cell is minimally invasive (an external component attached across the point of measurement); it is commonly employed with a capacitive element paired with a diaphragm that allows the capacitive body to separate or move together, generating a signal (via change in capacitance) that can be interpreted to a pressure drop. They are often used to detect small differences in large pressure drops. Its placement is similar to connecting a voltmeter in parallel to a resistor to measure its voltage "drop" (analgous to the pressure drop).

An example of a differential pressure cell using bellows can be found here: http://www.tpub.com/fluid/ch2s.h49.gif

The range of pressure measured and sensitivity of a differential cell depends upon the electric and elastic components used in the cell itself. It is a great sensor to use when measuring a pressure drop; however, for all other applications, it is fairly useless.

Vacuum Sensors

Such sensors are able to measure extremely low pressures or vacuum, referring to pressures below atmospheric pressure. Besides diaphragm and electric sensors designed to measure low pressures, there are also thermal conductivity gauges and ionization sensors.

Thermal Conductivity Gauges

The principle involved here is the change in gas thermal conductivity with pressure. However, due to deviation from ideal gas behavior in which the relationship between these two properties is linear, these kind of gauges, which are also called Pirani gauges, can only be used at low pressures, in the range of (0.4E-3 to 1.3E-3) MPa. They are amazingly sensitive elements as well (can detect changes of 6E-13 MPa).

In these gauges, a coiled wire filament has a current flowing through it, which heats up the coil. A change in pressure changes the rate of heat conduction away from the filament, thereby causing its temperature to vary. These changes in temperature can be detected by thermocouples in the gauge, which are also connected to reference filaments in the gauge as part of a Wheatstone bridge circuit.

An example of a Pirani gauge can be found here: www.bama.ua.edu/~surfspec/vacbasics_files/image046.jpg

Ionization Gauges

There are two categories for these types of gauges: hot cathode and cold cathode. For hot cathode gauges, electrons are emitted by heated filaments, while for the cold cathode gauges electrons are released from the cathode due to collision of ions. Electrons hit the gas molecules entering the gauge forming positive ions, which are collected and cause an ion current to flow. The amount of cation formation is related to the gas density and consequently the pressure to be measured, as well as the constant electron current used, hence the ion current flow is a direct measure of the gas pressure. These both are highly sensitive instruments and thus most suited for fractional pressures. The hot cathode gauges are even more sensitive than cold cathode gauges and are able to measure pressures near 10^{-8}Pa. Their sensitivity ranges from (1E-16 to 1E-13) MPa

An example of an ionization gauge can be found here: www.bama.ua.edu/~surfspec/vacbasics_files/image049.jpg

Summary

Types of elements	Press. Range	Sensitivity	Advantages	Disadvantages
Bourdon tube	0.1–700 MPa	0.03 MPa	Portable Low Maintenance	Static measurements Low accuracy
Bellows	0–2 MPa	0.0002 MPa	Can be used at low press.	Can only be connected to on/off [illegible]
Diaphragms	0.1–2.2 MPa	0.01 MPa	Quick response time Pressure units Generally good linearity Can be exposed to corrosives	Very expensive
Capacitive	0.5 Pa–70 MPa	0.07 MPa	Used for low pressure measurement ranges Robust High accuracy	Electrically complex Capacitor plates can stick together Limited by electronic circuits Pick up noise due to sensitive
Resistive	250Pa – 70 MPa	0.25 MPa	High sensitivity	Requires external AC source
Piezoelectric	0.001–100 MPa	0.1 MPa	Prompt response time	Affected by high temp. and static [illegible]
Potentiometric	0.035–70 MPa	0.07 – 0.35 MPa	Can be made very small	Low sensitivity and working range
Strain Gauge	0–1400 MPa	1.4–6.9 MPa	Very high sensitivity Can be used on mobile part.	Extremely slow response time Very small signal output
Differential	Depends on other elements	Depends on other [illegible]	Used for pressure drop	Can only be used for press. drop
Thermal Conductivity Gauges Viscosity	0.1 Pa – 1.3 MPa 1.3E-10 – 1.3E-2 MPa 0.005 MPa – 0.2 MPa	5E-10 MPa 1E-10 – 1.3E-6 MPa 1E-3 MPa	Measures vacuum ranges Highly sensitive Can measure high and ultra-high vacuum ranges Pressure [illegible] by temp.	Flat linearity only at low pressures Limited by piezoelectric effect Can't be used at high pressures

Example 3.3.1: Semi-batch Reactor Problem

Suppose you have a semi-batch reactor (1,000 L) with 50 kg of zinc within the reactor at 1 atm pressure and 25 C. 6M hydrochloric acid is flowing in at 1 L / min to react with the zinc to produce zinc chloride for use in another process.

a. What factors should be considered?
b. Say the valve fails at an operating pressure of 4 atm (i.e., it will not close and the reactor will be flooded with HCl) At what pressure would you (safely) set the shutdown point?
c. What type of sensor should be used?

Solution

3.5.1 The factors that should be considered are:

1. Process

1. Hydrochloric acid is very, very caustic (especially at such a high molarity), and thus whatever sensor you choose, it should be able to withstand the corrosive nature of hte process.

2. Range of pressures

1. Initially the reactor is at 1 atm of pressure. Considering the reaction $2\ HCl_{(aq)} + Zn_{(s)} \rightarrow H_{2(g)} + ZnCl_{2(aq)}$, you're producing one mole of hydrogen gas in addition to the existing air pressure in the vessel. As the reaction proceeds, the pressure would increase substantially within the vessel. Modeling the pressure of $H_{2(g)}$ ideally, $P = nRT / V$

2. $P - \frac{(6MHCl)(1L/min)(1molH_{2(g)}/2molHCl_{(aq)})(0.08206Latm/molK)((25+273)K)}{1000L}$ = 0.073 atm / min.
3. After about 1 hour, the pressure of $H_{2(g)}$ would increase to 4.38 atm, putting the total pressure in the vessel at 5.38 atm (past failing pressure for the valve).

3. Environment

1. There is no danger of high temperatures or excessive vibration due to high flow / reaction.

4. Sensitivity

1. Since this is a moderately dangerous process, we need to have the sensor output to a computer so an engineer can monitor the process as a fail-safe. We assume the sensor will signal the HCl valve to close once the operating pressure becomes 3 atm, however devices do fail from time to time. We also need high sensitivity, so electric components will be preferrable (i.e., we don't want the process to runaway as it potentially could if the sensor isn't highly sensitive to incremental changes).

The shutdown point

Considering the rapid increase in pressure, as evaluated in (2), and the failure of the valve at 4 atm, the shutdown point should be approximately 3 atm.

The type of sensor:

1. Considering the types of sensors discussed, we can automatically eliminate vacuum sensors as they operate at very low pressures (almost vacuum, hence the name). We can also eliminate differential pressure sensors as we are not looking for a pressure drop across the vessel.
2. Since we desire high sensitivity, we would like to use electric components. Considering the range of pressures (3 atm max ~ .3 MPa) a capacitive element would be optimal because it is robust and functions well in low pressure system.
3. Noting the corrosiveness of the HCl in the system, a diaphragm is likely to be used as the elastic element. Diaphragms are also fairly robust and offer quick response times.
4. This combination is likely to be filled in a strong, sealed, glycerine/silicone filled case so as to protect the sensor from degradation.

So, overall, the sensor we choose will utilize a diaphragm as the elastic component, a capacitive element as the electrical component, and an anti-corrosive body.

Example 3.3.2

Your supervisor told you to add a pressure sensor to a very expensive and important piece of equipment. You know that the piece of equipment is operating at 1MPa and at a very high temperature. What sensor would you choose?

Solution

Because the piece of equipment that you are dealing with is very expensive, you need a sensor that has the greatest sensitivity. An electric sensor would be desired because you could have it connected to a computer for quick and easy readings. Also, you need to pick a sensor that will operate at 1 MPa and a sensor that will be able to withstand high temperatures. From the information given in this article you know that there are many sensors that will operate at 1 MPa, so you need to decide on the other factors. One of the most sensitive electric sensors is the capacitive sensor. It has a sensitivity of 0.07MPa. The capacitive sensor usually has a diaphragm as the elastic element. Diaphragms have a rapid response time, are very accurate and work at 1 MPa.

References

- 2005/2006 Cole-Parmer General Catalog, 2005

- Liptak, Bela G.(2005). *Instrument Engineers Handbook*, 4th Edition, Volume 2. Taylor & Francis CRC Press. ISBN 0849310814
- Perry, Robert H.(1997), *Perry's Chemical Engineers' Handbook*, 7th Edition. McGraw-Hill. ISBN 0070498415
- Richardson, J.F. and Peacock, D.G. (1994), *Chemical Engineering*, 3rd Edition, Volume 3. Butterworth-Heinemann. ISBN 0080410030
- Freeman, Lawrence A., Carpenter, Michael C., Rosenberry, Donald O., Rousseau, Joseph P., Unger, Randy and McLean, John S. *Use of submersible pressure transducers in water-resources investigations*. http://pubs.usgs.gov/twri/twri8a3/#N10423
- "Pressure Sensors," www.data-acquisition.us/industrial_electronics/input_devices_sensors_transducers_transmitters_measurement/sensors/pressure-sensors.html
- Wikipedia, en.Wikipedia.org/wiki/Ion_gauge

Contributors and Attributions

- Authors: (09/14/06 / Date Revised: 09/20/06) Wen Chung Lim, James Bennett, Jamila Grant, Ajay Bhasin

This page titled 3.3: Pressure Sensors is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

3.4: Level Sensors

Introduction

Level sensors allow for the level control of fluid in a vessel. Examples of where these sensors are installed include reactors, distillation columns, evaporators, mixing tanks, etc. Level sensors provide operators with three important data for control: (1) the amount of materials available for processing, (2) the amount of products in storage, (3) the operating condition. Installing the correct level sensor ensures the safety of the operator and the surrounding environment by preventing materials in vessels from overflowing or running dry.

There are several different types of level monitors, including:

- Visual
- Float
- Valve Controlled
- Electronic
- Radiation

These different types of sensors can also be grouped into categories of process contact and non-process contact. As the name suggests, process contact sensors are within the tank, in physical contact with the material. Non-process contact sensors transmit various types of signals to reflect off of the material and thus measure the level. This sensor design can maintain its integrity within a potentially corrosive material and/or be positioned such that it can monitor the level changes from above the tank.

Visual Level Sensors

Visual level controls were the earliest developed level sensor. These types of monitoring devices can be something as simple as looking into an open container or inserting a marked object such as a dipstick. This type of sensor is the simplest and possibly the most reliable. These devices do not provide a way of connecting to a control device. They require human input with no way of automation; however, electronic issues will not be a possible problem.

Common Uses

Visual indicators can be found in many places and are not limited to chemical engineering applications. Examples of these are: dipsticks found monitoring the oil levels in a car and measuring cups with markings indicating different volumes. In chemical plants, visual indicators are used to measure changes in level as well as for high and low level alarms.

Benefits of Visual Level Sensors

Visual sensors are generally less expensive than other types of sensors. They are more reliable because of the simplicity of design.

Restrictions

Visual indicators do not always accurately measure how much volume is in the tank. They do not allow for any digital connections to process control systems. In the case of the sight tubes, they must be affixed directly to the tank and cannot be read remotely without a camera or other transmission tool. Tanks placed in hard-to-reach places would also cause problems contributing to inaccuracies in measurement, such as not reading markers on eye level or interpolating between marker intervals. Also, the connections between the tube and the tank are subject to residue buildup which could prevent the readings from being accurate. Weather is also a concern with sight tubes. External temperature changes could affect the fluid in the tube and thus measurement inaccuracy. For example, the fluid in the sensor could freeze, or the sensor could become clogged. This type of indicator requires modification to the vessel wall and would therefore be optimally installed at the time of initial construction. If considered at the time of the initial capital installation of the tank a sight tube will not add a large cost to the project. Yet, later modification to include this in the design and associated tank changes are potentially very costly. Despite these shortfalls, a sight tube is a

reliable and appropriate choice in many common level reading applications.

Sight Tube Indicators

Sight tube indicators allow operators to monitor levels with precision while keeping the tank sealed. This type of monitoring device is comprised of a vertical tube equal in height to the actual vessel. This tube is connected in at least two places directly to the vessel so that its contents can flow into the monitoring tube. This ensures that the height of liquid in the tank will be equal to the height of liquid in the sight tube. Markings can be affixed to the indicator for both calibration and volume readings. There are many styles of this indicator. Some models have a float ball contained in the tube so that it will float on top of the liquid. Other models of the indicator have many paddles on fixed posts through the length of the tube. These paddles float horizontally when submerged and are vertical when not submerged. This results in a color change allowing for an easy identification of the tank volume.

A sight tube. More pictures of this type of gauge can be found on Seetru Limited's website.

Float Type Level Sensors

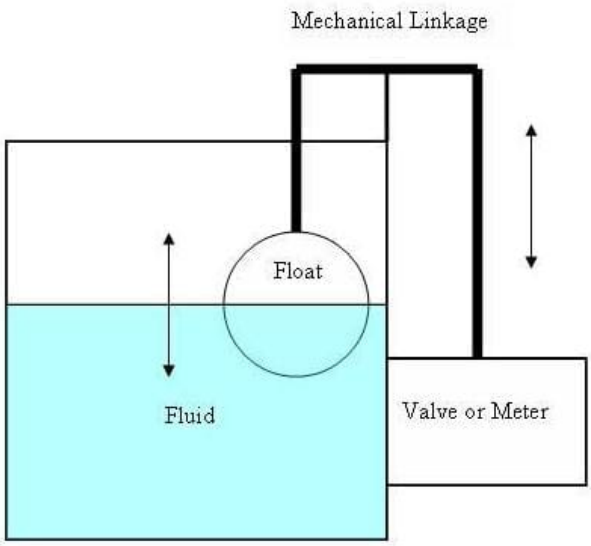
Float Types of level sensors are based on the principle of buoyancy which is the upward force produced on a submerged object by the displaced fluid. This force is equal to the weight of the displaced fluid. Float sensors take their measurements at the interfaces of materials, where the movement of the float and/or the force on the float are caused by the differing densities of the float and the fluid. There are two broad categories of Float Type level sensors: Buoyancy and Static.

Buoyancy Types

Buoyancy level sensors are less dense than the fluid and thus change position along with the fluid level. The movement of the float transmits the level information through some mechanical linkage to an output such as a valve or operator observation. There are basic three types of mechanical linkage - chain / tape sensors, lever / shaft mechanisms and magnetically coupled devices.

- Chain / Tape sensors – The linkage is by a flexible chain or tape.
- Lever / Shaft mechanisms – The linkage is a rigid shaft.
- Magnetically coupled devices – These devices are similar to the Chain/Tape sensors, except a magnet is attached to the float and another is acted upon by the floating magnet moving a tape like the chain/tape type devices. The moving magnet can be sequestered from the float attached magnet for use in corrosive media.

Schematic of a float type level sensor is shown below. As the fluid level rises or falls, buoyant force is transferred through mechanical linkages to your output device.



Static Types

Static level sensors are denser than the media being measured and thus do not move. As the level changes, the buoyant force acting on the "float", which is actually a weight, changes. The change in weight is measured by a scale. The level of the tank is calculated by measurements of the weight change in the float, not its actual change in position.

Formulas

Changes in the volumetric rate of material flow, resulting from changes in such things as the pressure of your pump will affect the operation of a float type level sensor. The same movement of the float will have the same effect on the valve, but as valve is moderating a different flow it has a different proportional sensitivity K_c.

$$K_c = \frac{q}{k}$$

where q is the flow rate and k is the height the float moves to completely traverse the valves operating range.

This is important because the given change in fluid height used to previously design the system will still change the valve's position in the same manner, but won't affect the same change in flow rate. If your pressure has dropped, your valve will have to be opened wider to match the same flow rate. The amount wider that it has to be opened is called an offset. Manipulating the valve response to the float movement, also known as the gain:

$$\text{Gain} = \frac{(\text{valve}_{\text{response}})}{\text{float}_{\text{movement}}}$$

can be used to alleviate this problem, but it won't work exactly the same way it did before the change in load.

Operation power is the amount of power the float has available to perform actions to control the system. Changes in the specific gravity of the fluid (or the float) will affect buoyancy of the float and thus will affect the amount of force exert on your system. A temperature is a very common parameter that can change your fluid's specific gravity. Warmer fluids tend to be less dense, and so your float will generate less buoyant force for a given change in fluid level.

Zero buoyancy force level is the level at which the weight of the float is exactly as buoyant as the fluid it replaces. It can be approximated by the following formula:

$$\text{Zero Buoyant Force Level} = D\frac{(SG_{\text{float}})}{SG_{\text{fluid}}}$$

where D is the diameter of the float and SG is the specific gravity.

Common Uses

Float Type Level Sensors regulate how much water is in the reservoir of a flush toilet; the float is attached to a lever which has a rotating axis that stops when the flow of water reaches a certain level.

- Chain / Tape sensors - regulating the level on storage tanks at atmospheric pressures.
- Lever / Shaft Mechanisms - regulating the level on vessels under pressure.

Benefits

Float Type level sensors do not require external energy sources to operate. Since they are simple robust machines it is easy to repair. The cost of these units makes them on the economical side, ranging from approximately $20 (homedepot.com) for the apparatus in a standard toilet to a few hundred dollars for a 3/4" cast iron valve with float (Liptak 2005). Prices would increase with higher quality materials and valve size. Floats could be made out of plastics or metals (steel, stainless steel, etc.). Material selection will depend on the application; more chemically resistant materials would be used for corrosive mediums.

Restrictions

Float type level sensors should only be used in clean fluids. Fluids that are a suspension of solids or slurries could foul the operation of the machine. Anything that could increase friction on the mechanical linkages could increase the dead band, which is the delay the sensor experiences due to the excess of force required to overcome the static friction of a non-moving system. An excess of force is required to be built up, so a float type level sensor will not respond immediately to changes in level.

Valve-based Level Sensors

Valve-based level sensors not only measure the fluid level, but also cause the fluid level to change accordingly. Two basic types of valves that will be discussed are altitude valves and diverter valves.

Altitude Valves

A simple altitude valve uses a spring that opens and closes different ports and lines when pressure changes due to the changing fluid levels. When fluid levels exceed the setting of the spring, a diaphragm connected to the spring lowers, closes the drain port, and opens the main line pressure. This turns off the main valve and stops fluid from flowing into the tank. When fluid levels decrease, the diaphragm raises, opens the drain port, and closes the main line pressure. This turns on the main valve and more fluid is supplied to the tank.

An example of an altitude valve is shown in the bottom left corner of the following website: http://www.gaindustries.com/html/01a_SAVV.htm

Common Uses

Altitude valves are on/off controls and can be found in supply lines connected to basins, tanks, and reservoirs. The main job of these valves is to prevent overflow of the fluid and to hold fluid level constant.

Benefits

There is no external power source needed when operating altitude valves because they are controlled by the pressure of the process fluid. More complicated altitude valves can also be used for other purposes other than an on/off control. These features include the following:

- Open the valve when pressure drops to a predetermined point
- Delayed opening so the valve only opens when fluid level drops a certain amount below the set point
- Close the valve slowly to eliminate pressure build-up
- Two-way flow to allow fluid to return when the level has dropped below the set point
- Open the valve while maintaining constant inlet pressure and plant distribution pressure

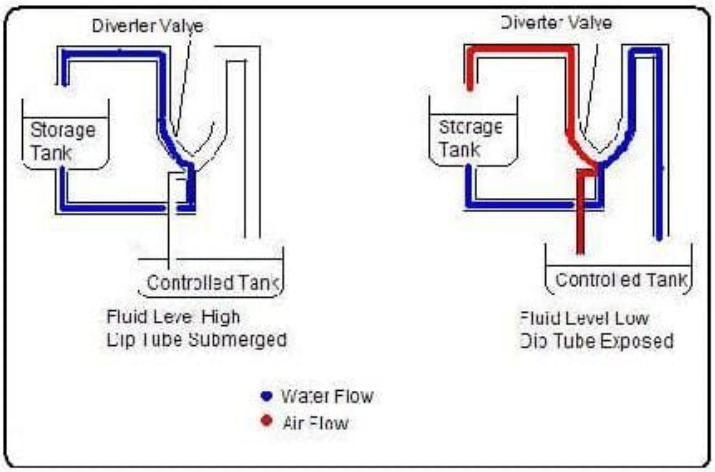
- Reduce pressure for outlet streams. This may be needed when working with equipment that has high supply pressures, such as aerator basins. Aerator basins are typically used in wastewater treatment because they have the capability of holding large amounts of water.
- Check valves can be installed to allow the valve to close if the pressure reaches a predetermined low point (which could be the cause of equipment failure elsewhere in the plant)

Restrictions

Altitude valves should not be used in pressurized vessels. These valves are constructed from a limited number of materials, such as cast iron. This restricts the number of fluids that can be used with the valves. Therefore, altitude valves are usually used for water service. Since these valves are usually used to operate a large amount of fluid, they are subsequently large with multiple functions that make them very expensive. A 30 inch altitude valve can cost over $50,000. In addition, frequent maintenance is required because of the many components and moving parts. These valves are operated at ambient temperatures and should be checked for freezing of stationary sensing lines.

Diverter Valves

Diverter valves are connected to a dip tube that is submerged in a controlled tank. When the fluid level of the tank drops, the dip tube is exposed to atmospheric pressure. Once air enters the dip tube, it utilizes the Coanda effect. Coanda effects occur when a curved surface exists. Gases tend to follow the nearest curved surface while pushing other fluids in a different direction. In the case of the diverter valve, air from the dip tube follows the curved surface of the control port back to the storage tank and pushes fluid flowing out of the storage tank to the other wall of the valve. This fluid then flows into the controlled tank to adjust the level.



Common Uses

Diverter valves control fluid level by acting as an on/off switch. They can also be modified to be used for throttling. Diverter valves are used as low level indicators so they should only been used when low levels of a tank are a concern.

Benefits

Diverter valves need no external power source for operation. They have no moving parts, plugs, or packing, so they require minimal maintenance. Diverter valves can be operated at any temperature and are not affected by vibration. They can be constructed from many materials such as metals, plastics, and ceramics. Therefore, there is limited restriction to the type of liquid used with the valves. However, when using fluidized solids and slurries, the dip tube must remain open, so build-up does not occur. Any build-up near or in the tube should be removed.

Restrictions

Diverter valves should not be used when operating under conditions other than atmospheric pressure. They are also not for use when dealing with hard to handle processes. Outlet pressure cannot be greater than atmospheric pressure or back pressure will occur and the valve will not operate correctly.

Electrical Level Sensors

Conductive Level Sensing

Conductive level sensors work by applying a low-voltage across two electrodes at different levels in a vessel. When both electrodes are immersed in a conductive liquid, a current flows. This type of electrical conductivity setup is best applicable for point level detection (level detection at a specific point in the material.) They are typically made of titanium, Hastelloy B, or stainless steel.

Example of a conductive level sensor.

Common Uses
Conductive level sensors are commonly used to measure both conductive and corrosive liquids. A common conductive liquid is water, whereas some common corrosive liquids are nitric acid, ferric chloride, and hydrochloric acid.

Benefits
This method is considered extremely safe due to the low voltages and currents used. Conductive level sensors are also known for their easy installation and use.

Restrictions
The greatest concern with conductive sensors is maintenance. The probe needs to be monitored for buildup on the sensor. Residues from wet or sticky fluids that cause build up can be prevented by coating the sensors with Teflon or polyethylene-based materials.

Capacitance Level Sensing

A capacitor is made up of two conductors (electrodes/plates) that are electrically separated by a nonconductor (dielectric). In the case of level sensing, one of the electrodes is typically a vertically positioned rod while the other is the metallic vessel wall. The dielectric between them is the material being measured in the vessel. The principle of capacitance level sensing is based on the formula below:

Capacitance = Dielectric constant x (Area of the plates ÷ Distance between plates)

If the dielectric is a liquid, the capacitance probe can measure the combined capacitance of both the liquid and gas. When the liquid level rises or falls, the total capacitance value will change. Since the dielectric constant and distance between the plates are constant (the rod and the vessel are stable), the only value changing the capacitance is the area of the plates immersed in the liquid. Total capacitance changes approximately proportionally to the liquid rise or fall in the column. Consequently, the liquid level can be calculated by the change in capacitance.

Common Uses

The list below shows just a few applications of capacitance level sensors found in industry and the materials they sense.

- Chemical/Petrochemical - Oil, clay, soda ash
- Food - Flour, Powered Milk, Sugar
- Charcoal - Wood, Charred sawdust
- Pharmaceuticals - Various powders and liquids
- Mining - Various minerals, metals, stone,

Benefits

Capacitance level sensing is useful in its ability to sense a wide variety of materials such as solids, organic and aqueous solutions, and slurries. For example, materials with dielectric constants as low as 2.1 (petroleum oil) and as high as 88 (water) or more can be detected with capacitance level sensing. The equipment, typically made of stainless steel, is simple to use, clean and can be specifically designed for certain applications. For example, they can be made to withstand high temperatures and pressures or have built-in splashguards or stilling wells for environments prone to turbulence.

Restrictions

There are limitations to using capacitance level sensors. One major limitation for capacitance probes is found when using tall bins that store bulk solids. Probes are required to extend for the length of the vessel, so in a lengthy vessel the long cable probes can be subject to mechanical tensions and breakage. Another limitation is from build up and discharge of a high-voltage static charge that can result from the rubbing and movement of low dielectric materials, but this danger can be eliminated with proper design and grounding. Also, abrasion, corrosion, and build up of material on the probe can cause variations in the dielectric constant of the material being measured. To reduce this issue, capacitance probes can be coated with Teflon, Kynar, polyethylene or other materials.

Radiation-based Level Sensors

Radiation-based level sensors are based on the principle of a material's ability to absorb or reflect radiation. The common types of radiation used in continuous level gages are ultrasonic, radar / microwave and nuclear.

Ultrasonic (Sonic) Level Sensors

Ultrasonic level sensor transmitters emit high frequency ultrasonic acoustic waves which are reflected back by the medium to the receivers. By measuring the time it takes for the reflected echo to be received, the sensor can calculate the actual distance between the receiver and the fluid level. These sensors can be accurate from a distance of 5mm to 30m.

Common Uses

Ultrasonic level sensors are commonly used for point level detection. They are best used for viscous liquids, slurries and bulk solids.

Benefits

The combination of ultrasonic level sensors' high functionality for relatively low prices makes them a popular choice for non-contacting level sensing.

Restrictions

Since the speed of sound in air can fluctuate in different conditions, ultrasonic sensors are not suitable for use in all applications. Environments that have varying moisture and temperatures will influence the sensors readings. Turbulence, steam, and foam prevent waves from reflecting properly and distort readings. In addition, the level sensor must be mounted properly so that it correctly senses the distance between the transmitter and the fluid level.

Microwave / Radar Level Sensors

Microwave / radar level sensors are similar to ultrasonic level sensors in that they require a transmitter and receiver. In addition to these materials, radar sensors also need an antenna and operator interface to use electromagnetic waves to calculate level distance.

Common Uses

These sensors are frequently used for non-contacting situations that require level sensing in varying temperature and pressure environments.

Benefits

Microwave / radar sensors have an advantage over ultrasonic sensors in that they are able to operate in high pressure and high temperature environments. They can also sense solids with grain sizes larger than 20mm.

Restrictions

When fluid level is under a thick layer of foam or dust, these sensors may not detect the fluid level, but instead detect the level of the dust or foam. Microwave / radar level sensors are also more costly than ultrasonic level sensors.

Nuclear Level Sensors

Nuclear level sensors rely on gamma rays for detection. Although these gamma rays can penetrate even the most solid of mediums, the intensity of the rays will reduce in passage. If the gamma ray emitter and the detector are placed on the top and bottom of a vessel, the thickness of the medium (level) can be calculated by the change in intensity.

Common Uses

These sensors are usually used when the material being measured presents a risk to human life or the environment. These include, for example, materials that are toxic, carcinogenic, or explosive.

Benefits

Nuclear radiation has the ability to detect level even through solid tank walls. Since they appear to “see” through walls, the nuclear gage may be modified and/or installed while the process is running and avoid expensive down time.

Restrictions

Nuclear level sensors are typically the last resort when choosing a level sensor. They not only require a Nuclear Regulatory Commission (NRC) license to install, but are extremely expensive in comparison to other level sensors.

Summary of Benefits and Restrictions

	Benefits	Restrictions
Visual	▪ Economical, less expensive ▪ Not subject to much technical/electronical issues	▪ Require human operator ▪ No connection to process control systems ▪ Cannot be read remotely ▪ Connections between tube and tank subject to clogging, freezing, etc ▪ Require modification to vessel wall
Float	▪ Do not require external energy source to operate ▪ Economical ▪ Easy to repair	▪ Should be used in clean fluids
Valve Controlled	▪ Do not require external energy source to operate ▪ Can be hooked up to process control systems ▪ Allow for two-way flow ▪ Eliminate pressure issues	▪ Should not be used in pressurized vessels ▪ Require frequent maintenance ▪ Costly
Electronic	**Conductive Level Sensing** ▪ Considerably safe: low voltage and current ▪ Easy installation **Capacitance Level Sensing** ▪ Able to sense a wide variety of materials ▪ Flexible in design	▪ Maintenance ▪ Probes must extend for full length of vessel, prone to mechanical damage

Radiation	**Ultrasonic Level Sensors** ▪ Non-contact level sensing ▪ Low cost	▪ Speed of sound fluctuates depending on conditions ▪ Readings can be influenced by environment
	Microwave/Radar Level Sensors ▪ Able to operate in high pressure, high temperature, and vacuum environments ▪ Able to sense solids with grain size >20 mm	▪ More costly ▪ Cannot work if liquid level is under thick layer of dirty foam or dust
	Nuclear Level Sensors ▪ Able to detect even through solid tank walls ▪ Able to be modified while process is running to avoid expensive down time	▪ Require licensing ▪ Extremely expensive

Example 3.4.1: Plant Storage Vessel

Imagine that you are responsible for maintaining a certain amount of chemical in a storage vessel for use later on downstream in your process. Your chemical has a specific gravity of 1.2. Assume that it is effectively independent of temperature. At maximum flow rate, a valve is feeding 80 gallons per minute into a storage vessel. 2 inches of valve travel is required to completely shut off the flow through the valve. The spherical float you are given to work with is 8 inches across, chemically compatible with your chemical, and its specific gravity can be altered. Your process dictates that you have to maintain about 400 gallons in this tank for use downstream, and that corresponds to a height of twenty inches from the bottom of the tank. At this height, the ball is 4 inches submerged. The ball has to travel 5 inches to completely close your valve. Calculate:

a. the specific gravity the float should be
b. the gain of the regulator
c. the proportional sensitivity of your system
d. assume the specific gravity of your chemical increases dramatically. Explain qualitatively what that should do to the gain, zero buoyant force level and dead band of the sensor
e. was a float type valve appropriate to be used in this problem?

Solutions

Using the formula for Zero Buoyant Force level $ZeroBuoyantForceLevel = D\frac{(SG_f)}{SG_p}$ **and plugging in values from the problem statement, we come up with the following** $4in = 8in\frac{(SG_f)}{1.2}$ **We used 4 inches, and not 20 inches because the zero buoyant force level indicates the level the float is submerged, not the level of the fluid. Isolating the value for SGfwe arrive at a specific gravity of 0.6 for the float.**

b) Using the formula for the gain of the regulator $Gain = \frac{(valve_{response})}{float_{response}}$ and values from the problem statement, we come up with the following $Gain = \frac{(2in)}{5in}$ proceeding with the calculation we arrive at a value of 0.4 for the gain.

c) Using the formula for the proportional sensitivity of the system $K_c = \frac{(q)}{k}$ and values from the problem statement, we come up with the following $K_c = \frac{(80gpm)}{2in}$ proceeding with the calculation we arrive at a value of $K_c = \frac{(40gpm)}{in}$

d) Float type level sensors work on the principal of buoyancy. If the density of the liquid being displaced increases, the buoyant force on the float increases for the same change in liquid height. The gain of the sensor refers to the amount of travel in the valve versus the amount of travel in the float, so for a given change in fluid height, the valve will still change the same amount as it did. However, since the buoyant force is increased due to the increased process fluid density, the dead band should decrease because the force of an incremental increase in height should enable the sensor to overcome friction more easily.

e) A float type valve was appropriate to be used in this problem, because the requirements were that "about 400 gallons" be maintained and the fluid was chemically compatible with the float. Due to unavoidable dead band, or the time it takes the float to respond to a given change in fluid, it won't maintain exactly 400 gallons, but it will stay close to that.

Example 3.4.2: Sensor Selection

Please answer the questions for the four hypothetical scenarios regarding selection of an appropriate level sensor. Give reasons for your selections.

a. A commonly used tank contains Dichloromethane which requires constant monitoring to ensure that no leaks exist. The precise volume will be regulated to monitor any decrease in level when a known transfer is not taking place. This tank will be filled to ~5,000 gallons and not be refilled until the level is less than 1,000 gallons of dichloromethane. What type of sensor should you install?
b. Under what circumstances would you select a radiation sensor?
c. Toilets have a Float Level Sensor to indicate when a flush has occurred and the bowl needs to be refilled. Which type of Float Sensor does the toilet have?
d. When would be a good time to use a visual sensor?

Solutions

a. The most appropriate selection of level sensor for this application is the capacitance sensor or the ultrasonic sensor. These are both electrical sensors that allow for measurements of changes in height, whereas other types offer a binary result: either the level is above or below a set point. The capacitance and ultrasonic sensors differ only by the way they are installed. Capacitance sensors require contact with the liquid while ultrasonic sensors do not. In this specific case for dichloromethane, which is non-corrosive, it is not necessary to monitor the volume without contacting the liquid, and therefore the capacitance sensor is also an acceptable selection.
b. A radiation-based level sensor is most appropriate when sensors cannot be brought into direct contact with the measured substance. This is only a possible option and often an expensive one.
c. Toilets have a lever/shaft mechanism. This float is less dense than the water which fills the tank. It will drop quickly when the toilet is flushed and will slowly rise as the tank fills with water. When the tank is full the float will rotate the shaft upwards enough to close off the flow of water.
d. Visual sensors should only be used when an electrical sensor is not required or as a backup. This is because they cannot be monitored automatically as visual sensors do not provide a digital response. It is also a good idea to provide a secondary way of obtaining level information when installing a new tank, and manual checks are often a good idea.

References

- [Image Available] Seetru, www.seetru.com/index_main.html, 12 Sept 2006.

- Liptak, Bela G., Instrument Engineers Handbook, 4th Edition, Volume 2. Taylor & Francis CRC Press, 2005.
- Perry, Robert H., Perry's Chemical Engineers' Handbook, 7th Edition. McGraw-Hill, 1997.
- Richardson, J.F. and Peacock, D.G., Chemical Engineering, 3rd Edition, Volume 3. Butterworth-Heinemann, 1994.
- Transactions in Measurement & Control: Volume 4, 2006. Omega Engineering, Inc. Retrieved September 10, 2006, from http://www.omega.com/literature/transactions/volume4/
- Wikipedia. Retrieved September 9, 2007, from en.Wikipedia.org/wiki/Level_sensor
- Considine, Douglas M., Process/Industrial Instruments & Controls Handbook, 4th Edition, McGRAW-HILL, INC., 1993.

Contributors and Attributions

- Authors: (September 14, 2006) Andrew MacMillan, David Preston, Jessica Wolfe, Sandy Yu
- Stewards: (September 10, 2007) Yoo Na Choi, Yuan Ma, Larry Mo, Julie Wesely

This page titled 3.4: Level Sensors is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

3.5: Flow Sensors

Introduction

Flow is defined as the rate (volume or area per unit time) at which a substance travels through a given cross section and is characterized at specific temperatures and pressures. The instruments used to measure flow are termed flow meters. The main components of a flow meter include the sensor, signal processor and transmitter. Flow sensors use acoustic waves and electromagnetic fields to measure the flow through a given area via physical quantities, such as acceleration, frequency, pressure and volume. As a result, many flow meters are named with respect to the physical property that helps to measure the flow.

Flow measurement proves crucial in various industries including petroleum and chemical industries. Consequently, flow measurement becomes a major component in the overall economic success or failure of any given process. Most importantly, accurate flow measurements ensure the safety of the process and for those involved in its success.

Before reading about the intricate details of various flow meters it's a good idea to think about aspects other than the design, governing equations, and the mechanism a flow meter uses in identifying a flow profile. To gain more from this section consider the following questions when learning about each flow meter because when choosing a flow meter the two main things to think about are cost and application:

1. **What mechanism does this flow meter use to measure flow?**
2. **How expensive is it?**
3. **What impact does it have on the system?** (How much power does it withdraw from the system and does the disturbance of flow cause significant problems up or downstream?
4. **What are the accuracy limits of the tool?** (Increased accuracy and precision results in higher cost especially for those with automated noise filtering and signal amplifying capabilities.

Sometimes you may be given an assignment to purchase a tool for your company's system and given with it a list of qualities to look for in order of importance, sometimes you're going to have to make this list yourself. The following table is a component of the TRIZ method for developing a list of qualities in descending order of importance on the left, perpendicular to the factor it has an impact on, to the right. For a given system you figured out that the cheapest solution for accurately measuring the flow rates of various pipe lines is by placing multiple flow meters in a series of positions. From a few calculations using propagation of error you find a moderate range of accuracy limits necessary for your system to be considered well monitored. You know the accuracy limits are less important than the total cost of all the flow meters because your boss told you not to waste any money. On the table you can see that **Cost** is at the top of the left column so **Cost** is the main concern. **Accuracy** is in the farthest left column on the top row so **Cost** most dramatically exacerbates the **Accuracy** of the flow meter when it's minimized. In the intersecting box you see that the lowest price flow meters are the ones with the least amount of accuracy, generally speaking. The next most important quality of your assignment is to get flow meters with the appropriate accuracy, thus **Accuracy** is below **Cost** in the left column. Looking to the top row you see that the **Accuracy** of the flow meter most greatly effects the **Impact on the system**. If you have a low quality flow meter it may be due to the side of the mechanism used to measure flow or the power the meter draws from the system (through mixing, frictional losses, increase in the turbulence of the flow, or buoyant effects caused from heat transfer). Completing the rest of the table you can decide if there are contradictions to what you thought were the most important qualities based on the inputs in the intersecting cells.

	Accuracy	Impact on system	Cost	Maintenance
Cost	Low accuracy/precision meters costs the least	Lower costs will result in larger meters that increase it effects on flow		Lowest cost means the flow meter will need frequent Maint.
Accuracy		Higher accuracy low cost flow meters have more surface area	What is the right price? How much error can I afford?	How long will it be reliable before it needs to be cleaned and recalibrated?
Maintenance	Application dependent, more Maint. Increases reliability	Down time, accumulation of deposits, disturbing flow	Down time, buildup on meter increases drag	
Impact on system	Lower impact may mean lower accuracy		Are there other costs of impact from installing the flow meter	How intricate is the meter, can it be quickly cleaned and recalibrated

An example where this model is important is for a system containing suspended particles in a fluid. If you want to measure the flow rate it might be cheapest to use a pitot tube yet the increased cost of maintaining the flow meter extends into down time of the system, more work from the technicians for a relatively small aspect of the process and lower profits as a result. For this system, maintenance would be the most important factor followed by accuracy, impact on the system and cost. If cost were least important factor you could afford to install a couple doppler meters, gather accurate measurements and the impact of the measurements on the system would be relatively low.

Common Types of Flow Meters

The flow rate as determined by the flow sensor is derived from other physical properties. The relationship between the physical properties and the flow rate is derived from fundamental fluid flow principles, such as Bernoulli's equation.

DIFFERENTIAL PRESSURE

These sensors work according to Bernoulli's principle which states that the pressure drop across the meter is proportional to the square of the flow rate.

$$-\Delta p \propto V^2$$

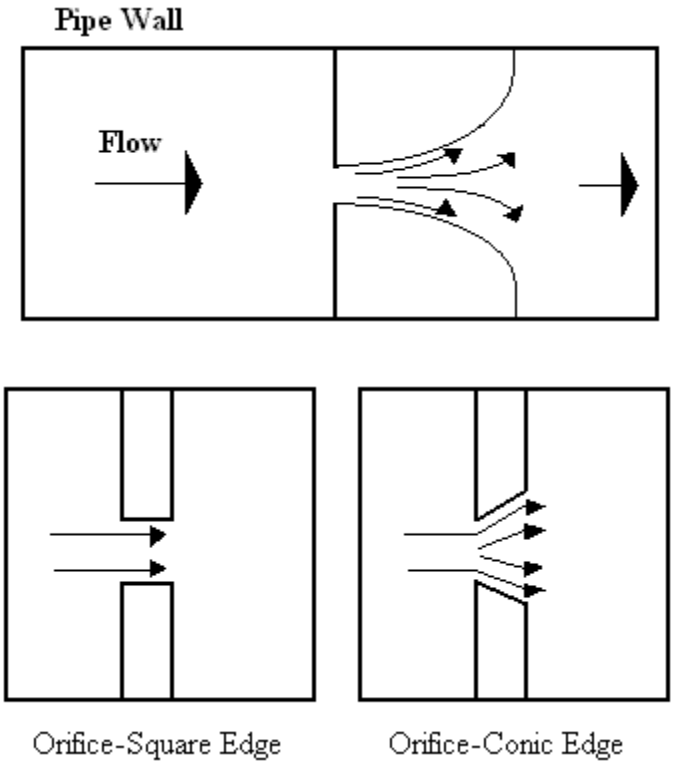
Using the pressure drop across a pipe's cross section is one of the most common manners to determine a flow measurement. As a result, this property has heavy implications for industrial applications. Flow meters characterized by differential pressure come in several different varieties and can be divided into two categories, laminar and turbulent. Differential pressure sensors operate with respect to Bernoulli's principle. Bernoulli's principle states that the pressure drop across the meter is proportional to the square of the flow rate.

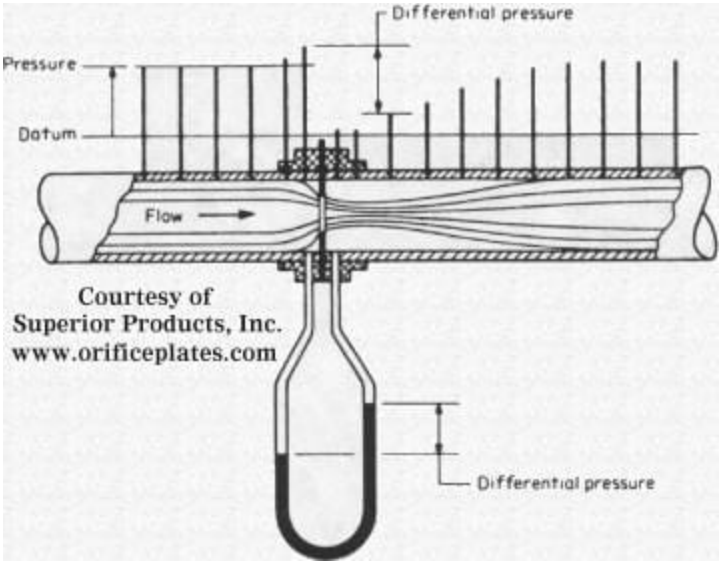
Orifice Meter

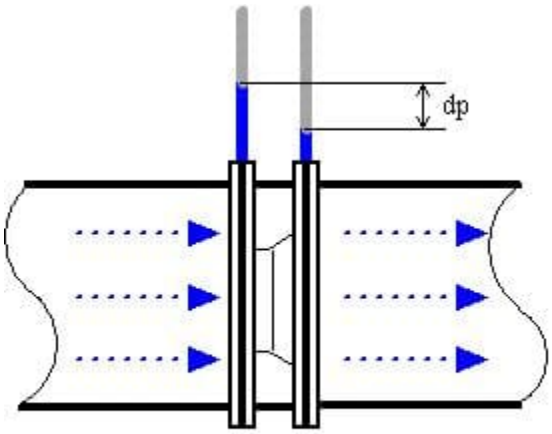
Orifice plates are installed in flow meters in order to calculate the material balances that will ultimately result in a fluid flow measurement on the sensor. An orifice plate is placed in a pipe containing a fluid flow, which constricts the smooth flow of the fluid inside the pipe. By restricting the flow, the orifice meter causes a pressure drop across the plate. By measuring the difference between the two pressures across the plate, the orifice meter determines the flow rate through the pipe. The larger the pressure drop, the faster the flow rate would be. There are two types of orifice meters that are usually used in industry, they are the orifice-square edge and the orifice-conic edge. The orifice-square edge has insignificant friction at the interface between the fluid and the orifice plate. These types of orifice plates are recommended for smooth fluid flows, particularly clean liquids and gases such as steam. Generally, drain holes are incorporated in the design so that liquids and gases do not accumulate inside the pipes. Multi-phase fluids are not recommended for the orifice-squared edge because clogging becomes a significant problem as time progresses. The orifice-conic edge is similar to the orifice-square edge, the primary difference being that the orifice-conic

edge has a gradually expanding cross-sectional area, and this cross-sectional area is circular in shape. A conic-edge design is often a better choice for low velocity, high viscosity flows. Both types operates best under comparable temperature and pressure conditions, pipe sizes and provide similar accuracies.

Orifice meters used in conjunction with DP (Differential Pressure) cells are one of the most common forms of flow measurement. In addition, an orifice meter can be used to measure flows when there is a significant difference in pressure in the pipe, like between the upstream and downstream sides of a partially obstructed pipe, which is exactly what the orifice meter does on its own. The plate offers a precisely measured obstruction that essentially shrinks the pipe and forces the flowing substance to constrict. A DP cell allows the comparison of pressure on the upstream (unobstructed) side and the downstream (constricted) side. A greater rate of fluid flow would usually result in a larger pressure drop, since the size of the orifice remains constant and the fluid is held longer building potential energy on the upstream side of the orifice. Some of the other types of orifice plates include concentric, eccentric and segmental plates, each having different shapes and placements for measuring different processes. These plates are available in varied shapes so that the meter has the optimum structure for different applications. Moreover, the density and viscosity of the fluid, and the the shape and width of the pipe also influences the choice of plate shape to be used.







www.EngineeringToolBox.com

Such a pressure drop across the plate is then related to the flow rate using basic fluid mechanics principles that incorporate parameters such as density of the fluid and size of the pipe. The flow rate Q, given by the orifice meter, is usually modeled with the following equation:

Where $p_1 - p_2$ is the pressure drop across the plate, ρ is the fluid density, A_1 is the pipe cross-sectional area, A_2 is the orifice cross-sectional area, and C_d is the discharge coefficient (usually about 0.6). C_d is used to account for energy losses within the system.

The orifice meter is one of the most commonly used flow meters, since it is inexpensive to install and operate, it is uncomplicated and easy to construct, and it is highly robust and long lasting. Orifice meters are not only simple and cheap, they can also be delivered for almost any application and be made of any material. This simplicity of its design and function is one of its paramount advantages, with the meter essentially consisting of just a modified plate. This not only reduces its initial price but also shrinks its operating costs, maintenance expenses, and spare parts expenditure and availability. Lower flow rates reduces their accuracy, whereas higher flow rates combined with high quality, unworn orifice plates increases it. The orifice plate is best when a sharp edge is present towards the upstream side of the meter. Wear reduces the accuracy of orifice plates. The turndown rate of orifice plates are generally less than 5:1. More information about turndown rates is mentioned in a later section.

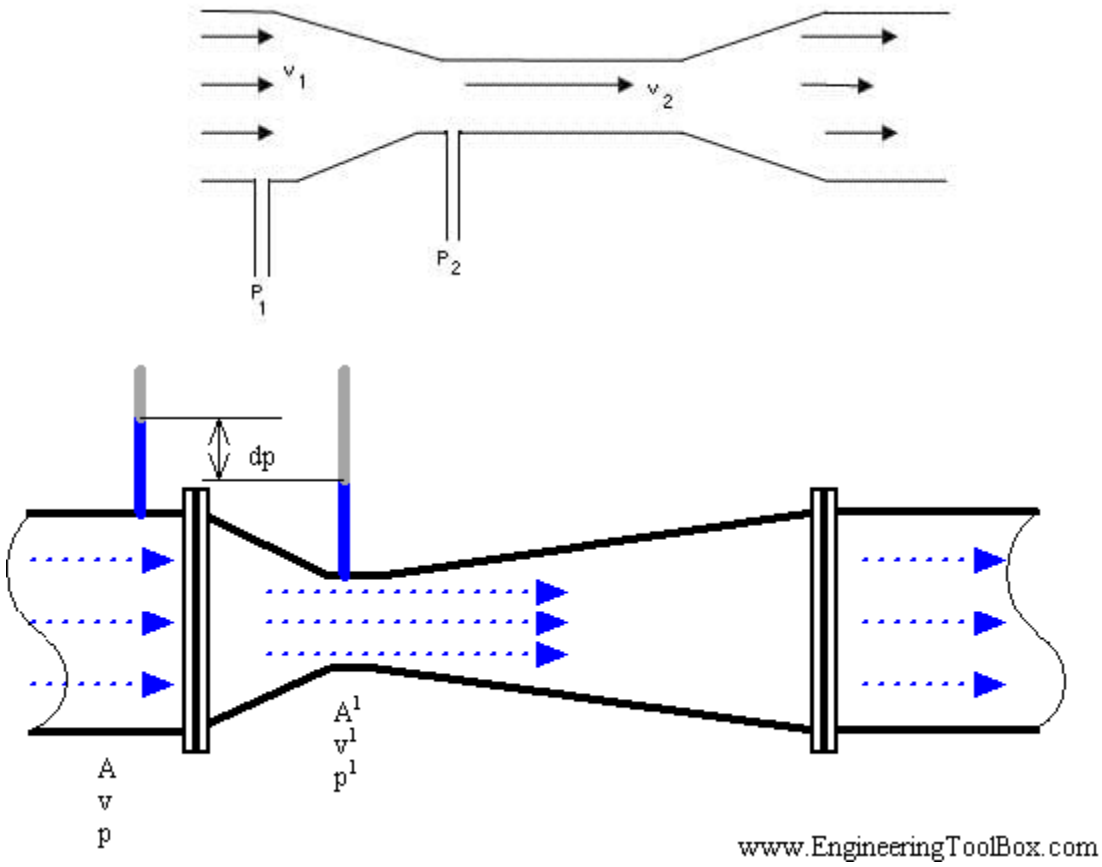
Venturi Meter

Venturi meters can pass 25 – 50% more flow than an orifice meter. In a Venturi meter setup, a short, smaller diameter pipe is substituted into an existing flow line. Because the Venturi meter is insensitive to changes in the velocity profile of the fluid flow, the pipe design does not need to be straight like the orifice meter. Though initially expensive, the Venturi meter has relatively low maintenance and operation costs.

In the Venturi Tube the fluid flowrate is measured by reducing the cross sectional flow area in the flow path, generating a pressure difference. After the constricted area, the fluid is passes through a pressure recovery exit section, where up to 80% of the differential pressure generated at the constricted area, is recovered.

There are two main types of Venturi meters. The first one, known as the classical Herschel Venturi meter, is a very long meter characterized below. Pressure readings at different points in the meter are combined to provide an average pressure reading. Cleaning the classical Venturi meter is limited. The second type of Venturi meter is known as the short form Venturi meter. This differs from its longer counterpart by reduced size and weight.

By Bernoulli's principle the smaller cross-sectional area results in faster flow and therefore lower pressure. The Venturi meter measures the pressure drop between this constricted section of pipe and the non-constricted section.



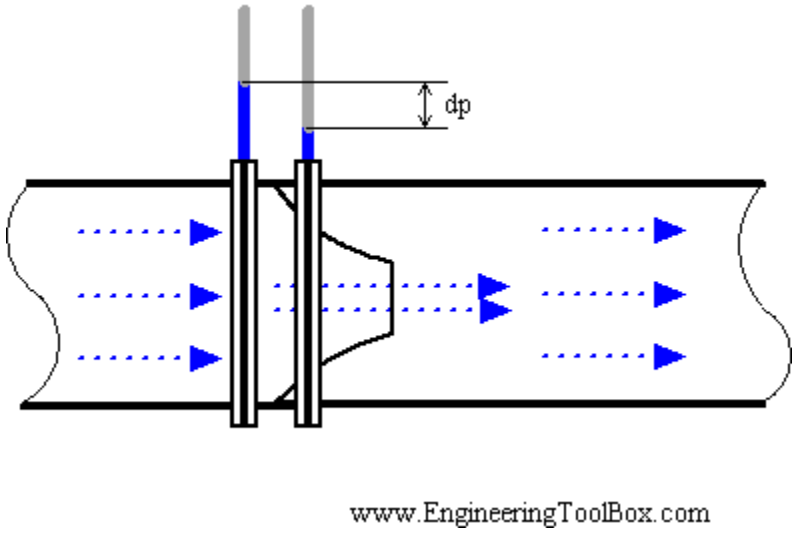
The discharge coefficient for the Venturi meter is generally higher than that used for the orifice, usually ranging from .94 to .99.

The Venturi meter is most commonly used for measuring very large flow rates where power losses could become significant. The Venturi flow meter has a higher start up cost than an orifice, but is balanced by the reduced operating costs.

Due to simplicity and dependability, the Venturi tube flowmeter is often used in applications where higher turndown ratiosor lower pressure drops than orifice plates can provide are necessary. With proper instrumentation and flow calibrating the venturi meter flowrate can be reduced to about 10% of its full scale range with proper accuracy. This provies a turndown ratio of around 10:1.

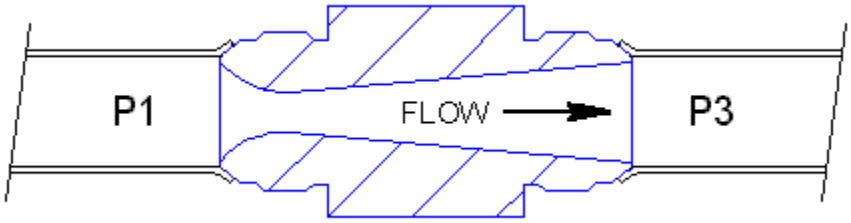
Flow Nozzle

Another type of differential pressure flowmeter is the flow nozzle. Flow nozzles are often used as measuring elements for air and gas flow in industrial applications. At high velocities, Flow Nozzles can handle approximately 60 percent greater liquid flow than orifice plates having the same pressure drop. For measurements where high temperatures and velocities are present, the flow nozzle may provide a better solution than an orifice plate. Its construction makes it substantially more rigid in adverse conditions and the flow coefficient data at high Reynolds numbers is better documented than for orifice plates. Liquids with suspended solids can also be metered with flow nozzles. However, the use of the flow nozzles is not recommended for highly viscous liquids or those containing large amounts of sticky solids. The turndown rate of flow nozzles is similar to that of the orifice plate. The flow nozzle is relatively simple and cheap, and available for many applications in many materials.



The Sonic Nozzle- Critical(Choked)Flow Nozzle

One type of flow nozzle is the sonic nozzle. The Sonic Nozzle is a converging-diverging flowmeter. It consists of a smooth rounded inlet section converging to a minimum throat area and diverging along a pressure recovery section or exit cone.



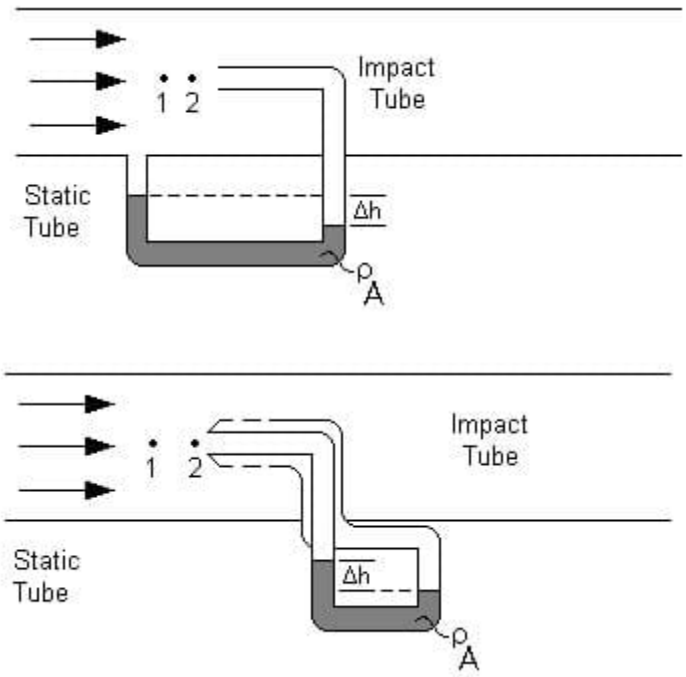
The Sonic Nozzle is operated by either pressurizing the inlet (P1) or evacuating the exit (P3), to achieve a pressure ratio of 1.2 to 1 or greater, inlet to outlet. When a gas accelerates through a nozzle, the velocity increase and the pressure and gas density decrease. The maximum velocity is achieved at the throat, the minimum area, where it breaks Mach 1 or sonic. At this point it's not possible to increase the flow by lowering the downstream pressure. The flow is choked. Pressure differences within a piping system travel at the speed of sound and generate flow. Downstream differences or disturbances in pressure, traveling at the speed of sound, cannot move upstream past the throat of the Nozzle because the throat velocity is higher and in the opposite direction.

Sonic Nozzles are used in many control systems to maintain fixed, accurate, repeatable gas flow rates unaffected by the downstream pressure. If you have a system with changing or varying gas consumption downstream and you want to feed it a constant or locked flowrate, a Sonic Nozzle is an excellent way to achieve this.

Pitot Tubes

Pitot tubes measure the local velocity due to the pressure difference between points 1 and 2 in the diagrams below. Unlike the

other differential flow meters, the pitot tubes only detect fluid flow at one point rather than an overall calculation. The first diagram shows a simple pitot tube configuration while the second shows a compact pitot tube configuration.



Both tubes work in a similar manner. Each pitot tube has two openings, one perpendicular to the flow and one parallel to the flow. The impact tube has its opening perpendicular to the fluid flow, allowing the fluid to enter the tube at point 2, and build up pressure until the pressure remains constant. This point is known as the stagnation point. The static tube, with openings parallel to the fluid flow gives the static pressure and causes a sealed fluid of known density to shift in the base of the tube. Pressure drop can be calculated using the height change along with the fluid densities and the equation below.

$$\Delta p = \Delta h \left(\rho_A - \rho\right) g$$

with Δp as the pressure drop, ρ_A as the known fluid density, ρ as flowing fluid's density, and g as the acceleration due to gravity.

This pressure drop can be related to the velocity after accounting for the losses throughout the piping in the system, given by C_p. This dimensionless coefficient is found through accurate calibration of the pitot tube. The equation below describes this relationship.

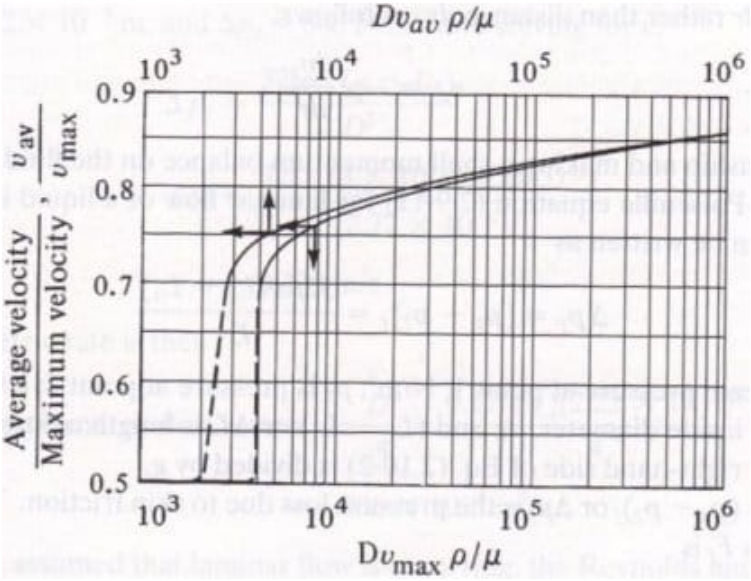
$$v = C_p \sqrt{\frac{2\left(p_1 - p_2\right)}{\rho}}$$

with v as the fluid velocity, C_p as the loss coefficient, p_1 as the pressure at point 1, p_2 as the pressure at point 2, and ρ as the flowing fluid's density.

By placing the tube at the exact center of the pipe, the maximum velocity can be measured and the average velocity can be calculated via the Reynolds number. The governing equation and chart are below.

$$Re = \frac{D v_{\max} \rho}{\mu}$$

with Re as the Reynolds number, D as the pipe diameter, v_{max} as the maximum velocity, ρ as the flowing fluid's density, and μ as the flowing fluid's viscosity.



Adapted from Geankoplis.

Finally, the flow rate can be found by accounting for the area of the pipe.

$$Q = v_{\text{avg}} \frac{\pi}{2} r^2$$

with Q as the volumetric flow rate, v_{avg} as the average velocity, and r as the pipe's radius.

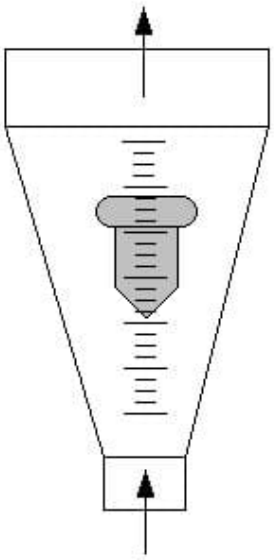
It should be noted that all the equations apply to incompressible fluids only, but these can be used to approximate gas flows at moderate velocities. This flow meter must also be placed at least 100 pipe diameters in distance, downstream of the nearest flow obstruction. This ensures no unwanted pressure fluctuations and accurate pitot tube readings. Furthermore, fluids containing large particles can clog the pitot tube and should be avoided.

DIRECT FORCE

These flow meters are governed by balancing forces within the system.

Rotameter

A rotameter is a vertically installed tube that increases in diameter with increasing height. The meter must be installed vertically so that gravity effects are easily incorporated into the governing equations. Fluid flows in through the bottom of the tube and out through the top. Inside the glass tube there is a float that changes position with the flow rate. When there is no liquid flow, the float rests in the bottom of the meter.



The applied concept for the rotameter is differential area. As the flow rate of the fluid changes, the position of the float changes and annular area change directly, keeping a constant pressure drop across the meter. Changes in float position and annular area are approximately linear with changes in flow rate. Upon achieving a stable flow rate, the vertical forces are balanced and hence the position of the float remains constant. The volumetric flow is given by the following equation:

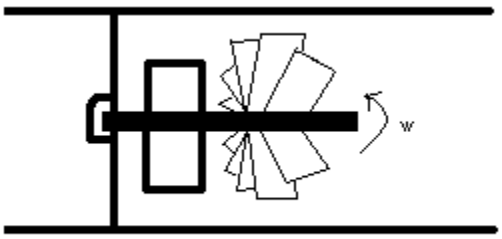
$$Q = CA_b \sqrt{\frac{2g\left(\frac{V_f(\rho_f-\rho)}{A_f} - \rho h_f\right)}{\rho\left[1-\left(\frac{A_b}{A_a}\right)^2\right]}}$$

with C being the discharge coefficient, A_b being the cross sectional area of the top of the float, V_f volume of the float, ρ_f the density of the float, ρ the density of the fluid h_f the height of the float, A_a the cross sectional area of the bottom of the float.

Generally, rotameters are inexpensive and simple to use. This allows them to be used in many plant applications.

Turbine Meter

A turbine wheel is placed in a pipe that holds the flowing fluid. As the fluid flows through the turbine, the turbine is forced to rotate at a speed proportional to the fluid flow rate. A magnetic pick-up is mounted to the turbine wheel, and a sensor records the produced voltage pulses. Voltage information can then be translated into the actual flow meter reading.



The following is the equation used to model the turbine meter:

$$Q = VA = \frac{\omega(\bar{r})^2 A^2}{\bar{r}A\tan\beta - 0.037Re^{-0.2}n\left(R_o + R_i\right)D\sin\beta}$$

with A the pipe area, $\bar{r}$ the root mean squared radius, ω rotational speed,β the angle between the flow direction and the turbine blades,R_O the outer blade radius, R_i the inner radius, and D the distance between blades.

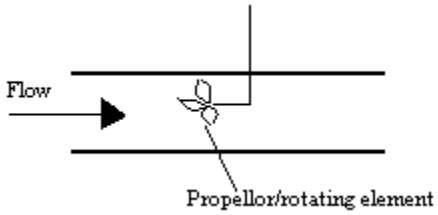
There are two main advantages of the turbine meter over conventional differential head devices

1. The extended are more accurate registration of flow in the low flow range of process operation. This results from the registration being proportional to the velocity rather than the velocity square
2. The comparatively low head loss across the meter

Another advantage to using this type of flow meter is reliability. Extensive testing has proven these results. Additionally, the turbine flow meter does not have a high installation cost. However, due to the turbine wheel motion, a low to medium pressure drop can result. Turbine wheel replacement may also be required due to abrasion caused by particles within the fluid.

Propeller Flow Meter

Propeller flow meters have a rotating element similar to the wheel in turbine meters. Again, rotation is caused by fluid flow through the propeller, and voltage pulses are created as the propeller passes a magnetic or optical sensor. Similarly, the frequency of the pulses is proportional to flow rate of the fluid and the voltages can be directly correlated with the fluid flow rate.



Propeller flow meters are often used specifically with water, though other fluids may also be used. Low cost coupled with high accuracy make propeller flow meters a common choice in many applications.

Paddle Wheel Sensors

A kind of propeller sensor is the paddle wheel sensor. Paddle wheel sensors are similar to turbine sensors, except for one thing. The shaft of the paddle wheel sensor is perpendicular to the flow of the fluid while the turbine sensor's shaft is parallel to the flow of the fluid. This adds several advantages for the paddle wheel flow sensor. Due to the shaft being perpendicular to the flow, it sustains less axial from the fluid, and thus less friction. Paddle wheel sensors also have a smaller number of blades, so there is less force needed to turn the paddle wheel. This means that a paddle wheel can be accurate at lower flows, have a high efficiency, as well as a longer lifetime.

There are two kinds of paddle wheel sensors, insertion and inline sensors. There is more than one design for an insertion sensor, but one popular design has the bearing built into the rotor and the shaft sliding through it as the center axis of spin. The blade sticks out and is inserted into the pipe with the flowing fluid. An inline paddle wheel sensor is used for smaller pipes. It contains a rotor assembly with the magnet sealed inside and a main body.

Coriolis Mass Flow Meter

A Coriolis flow meter harnesses the natural phenomenon wherein an object will begin to "drift" as it travels from or toward the center of a rotation occurring in the surrounding environment. A merry-go-round serves as a simple analogy; a person traveling from the outer edge of the circle to its center will find himself deviating from his straight-line path in the direction of the ride's rotation.

Coriolis flow meters generate this effect by diverting the fluid flow through a pair of parallel U-tubes undergoing vibration perpendicular to the flow. This vibration simulates a rotation of the pipe, and the resulting Coriolis "drift" in the fluid will cause the U-tubes to twist and deviate from their parallel alignment. This Coriolis force producing this deviation is ultimately proportional to the mass flow rate through the U-tubes.

$$\text{Mass Flow} = \frac{F_c}{2wx}$$

where F_c is the Coriolis force observed, w is the angular velocity resulting from rotation, and x is the length of tubing in the flow

meter.

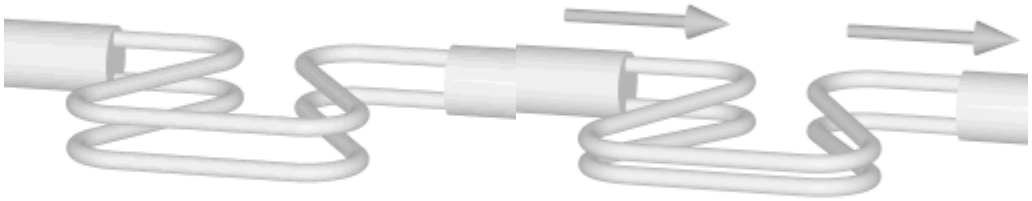


Figure 3.5.1: (left) The vibration pattern during no-flow and (right) The vibration pattern with mass flow. (CC BY-SA 2.5; Cleonis via Wikipedia)

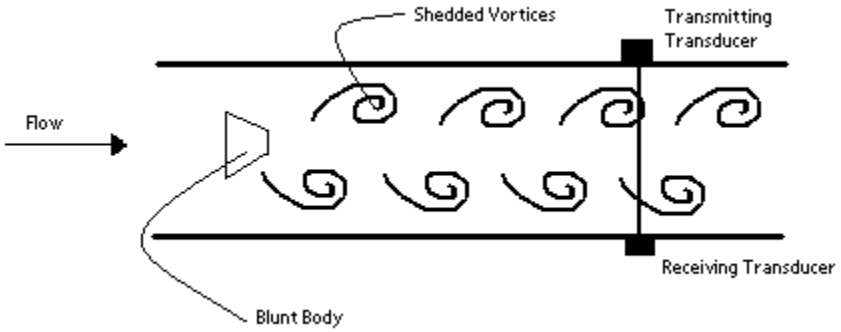
Because the Coriolis flow meter measures the mass flow rate of the fluid, the reading will not be affected by fluctuations in the fluid density. Furthermore, the absence of direct obstructions to flow makes the Coriolis flow meter a suitable choice for measuring the flow of corrosive fluids. Its limitations include a significant pressure drop and diminished accuracy in the presence of low-flow gases.

FREQUENCY

These flow meters use frequency and electronic signals to calculate the flow rate.

Vortex Shedding Flow Meter

A blunt, non-streamline body is placed in the stream of the flow through a pipe. When the flow stream hits the body, a series of alternating vortices are produced, which causes the fluid to swirl as it flows downstream. The number of vortices formed is directly proportional to the flow velocity and hence the flow rate. The vortices are detected downstream from the blunt body using an ultrasonic beam that is transmitted perpendicular to the direction of flow. As the vortices cross the beam, they alter the carrier wave as the signal is processed electronically, using a frequency-to-voltage circuit. The following diagram shows the basic principle of the vortex-shedding flow meter:



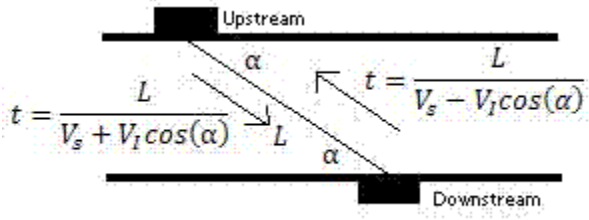
Vortex-shedding flow meters are best used in turbulent flow with a Reynolds number greater than 10,000. One advantage of using this type of flow meter is its insensitivity from temperature, pressure, and viscosity. The major disadvantage to using this method is the pressure drop caused by the flow obstruction.

Ultrasonic Flow Meters

There are two types of Ultrasonic meters, the transit time/time of flight and Doppler models, both of which have unique equations representing the principles behind them. The basis for these meters is monitoring ultrasonic waves in fluid passing through a pre-configured acoustic field. These meters are based on the technique of sound waves that change.

Transit Time/Time of Flight Flow Meters

Transit time meters have two opposing transducers outside of the pipe to measure the time of a signal sent from a transducer upstream to a transducer downstream and vice versa.

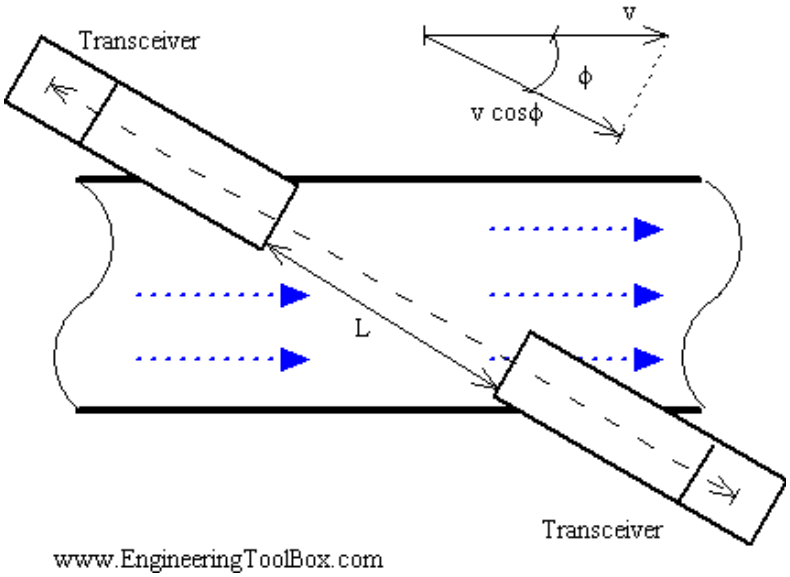


This allows the average velocity and hence the flow rate, Q, to be determined.

$$Q = K\left(\frac{\pi d^3 \tan\alpha}{8}\right)\left(\frac{1}{t_{UD}} - \frac{1}{t_{DU}}\right)$$

where d is the diameter of the pipe, α is the angle between direction of the flow and the pipe, t_{UD} is the time for the signal to reach downstream transducer from the upstream transducer, and t_{DU} is the time for signal to reach upstream transducer from the downstream transducer.

With the Time of Flight Ultrasonic Flowmeter the time for the sound to travel between a transmitter and a receiver is measured. This method is not dependable on the particles in the fluid.



Two transmitters / receivers (transceivers) are located on each side of the pipe. The transmitters sends pulsating ultrasonic waves in a predefined frequency from one side to the other. The difference in frequency is proportional to the average fluid velocity.

Downstream pulse transmit time can be expressed as

$$t_d = \frac{L}{c + v \cos \Phi}$$

where t_d is the downstream pulse transmission time and L is the distance between transceivers.

Downstream pulse transmit time can be expressed as

$$t_u = \frac{L}{c - v \cos \Phi}$$

where t_u is the upstream pulse transmission time.

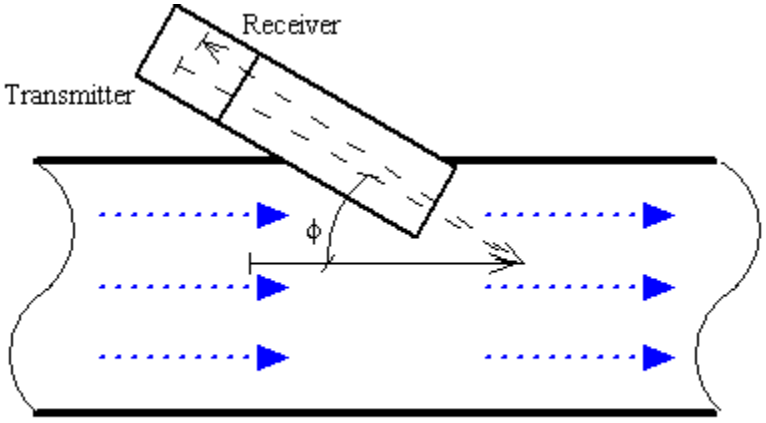
Since the sound travels faster downstream than upstream, the difference can be expressed as

\[t = \dfrac{t_d - t_u) t = \dfrac{ 2 v L \cos Φ}{c^2 - v^2 \cos 2Φ} t = \dfrac{2 v L \cos Φ}{c^2} \label{4} \]

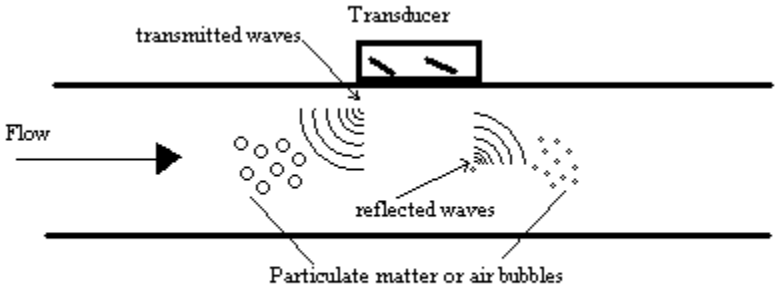
(since v is very small compared to c)

Doppler Meters

Doppler meters use the frequency shift of an ultrasonic signal when it is reflected by suspended particles or gas bubbles (discontinuities) in motion. The Doppler Effect Ultrasonic Flowmeter uses reflected ultrasonic sound to measure the fluid velocity. By measuring the frequency shift between the ultrasonic frequency source, the receiver, and the fluid carrier, the relative motion are measured. The resulting frequency shift is named the *Doppler Effect*.



www.EngineeringToolBox.com



These reflected signals travel at the velocity of light.

$$V_; = \frac{V_s \Delta f}{2 f_{\text{actual}} \cos \alpha}$$

where f is the actual frequency and Δf is the change in frequency or frequency shift.

The fluid velocity can be expressed as

$$v = \frac{c(f_r - f_t)}{2ft\cos\Phi}$$

where

- f_r is the received frequency
- f_t is the transmission frequency
- v is the fluid flow velocity
- Φ is the relative angle between the transmitted ultrasonic beam and the fluid flow
- c is the velocity of sound in the fluid

This method requires that there are some reflecting particles in the fluid. The method is not suitable for clear liquids.

Advantages with the Doppler Effect Ultrasonic Flowmeter

Doppler meters may be used where other meters don't work. This could be in liquid slurries, aerated liquids or liquids with some small or large amount on suspended solids. The advantages can be summarized to:

Limitations with Doppler Effect Ultrasonic Flowmeters

Benefits with Ultrasonic Flowmeters As A Whole

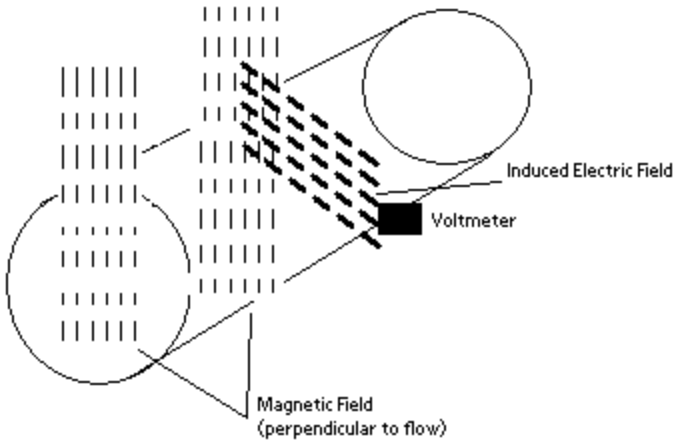
- Doppler flowmeters performance are highly dependent on physical properties of the fluid, such as the sonic conductivity, particle density, and flow profile.
- Non uniformity of particle distribution in the pipe cross section may result in a incorrectly computed mean velocity. The flowmeter accuracy is sensitive to velocity profile variations and to the distribution of acoustic reflectors in the measurement section.
- Unlike other acoustic flowmeters, Doppler meters are affected by changes in the liquid's sonic velocity. As a result, the meter is also sensitive to changes in density and temperature. These problems make Doppler flowmeters unsuitable for highly accurate measurement applications.
- Obstruction less flow
- Pressure drop equal to an equivalent length of straight pipe
- Unaffected by changes in temperature, density or viscosity
- Bi-directional flow capability
- Low flow cutoff
- Corrosion-resistant

Both meters are effective in measuring open channels and partially filled pipes but are very sensitive to flow conditions and hence should be calibrated with care. Also, there is no pressure drop since there are no obstructions in the flow path.

OTHER TYPES

Magnetic Flow Meter

One magnetic model flow meter positions electric coils around the pipe of the flow to be measured. A pair of electrodes is set up across the pipe wall. The fluid flowing has a minimum value of electrical conductivity, the movement of the fluid through the pipe acts as a conductor moving across the magnetic field. There is an induced change in voltage between the electrodes, which is proportional to the flow velocity.



The flow velocity is found by measuring the changes of induced voltage of the conductive fluid passing through a controlled magnetic field at right angles. According to Faraday's Law, the magnitude of the voltage induced is directly proportional to the product of the magnetic flux, distance between probes and the velocity of the medium (fluid).

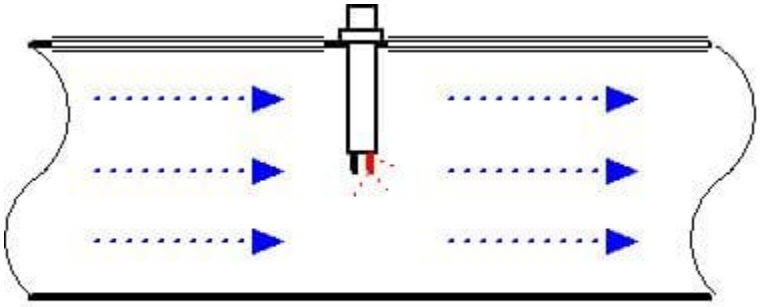
$$E = -N\frac{d\phi}{dt} = -NB\frac{dA}{dt} = -NBD\frac{dz}{dt} = -NBDv$$

where E is the voltage of induced current, N is the number of turns, B is the external magnetic field, φ is the magnetic flux, D is the distance between electrodes and v is the velocity of the fluid.

Some of the advantages are minimum pressure drop because of minimum obstructions in flow path; low maintenance cost because of no moving parts. One of the disadvantages is that it usually requires a fluid with an electrical conductivity of higher than 3 µS/cm.

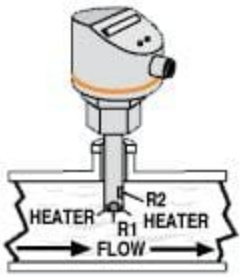
Calorimetric Flow Meter

This type of flow meter is suitable for monitoring the flow rates of liquid and gaseous mediums. The calorimetric principle is based on two temperature sensors in close quarters of one another but thermally insulated from one another.



www.EngineeringToolBox.com

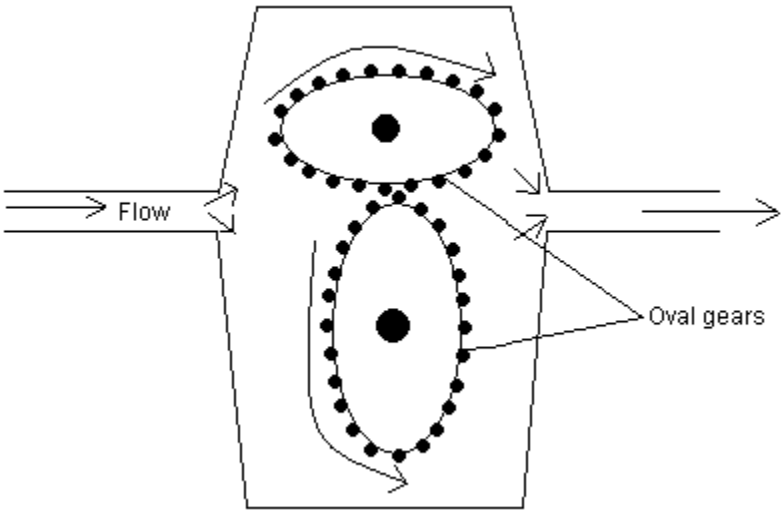
One of the two temperature sensors is constantly heated and the cooling effect of the flowing fluid is used to monitor the flow rate. In a stationary phase fluid condition there is a constant temperature difference between the two temperature sensors. When the fluid flow increases, heat energy is extracted from the heated sensor and the temperature difference between the sensors are reduced. The reduction is proportional to the flow rate of the fluid. The calorimetric flow meter can achieve relatively high accuracy at low flow rates.



Common applications: air compression, argon compression, nitrogen compression, carbon dioxide compression and flow detection of all fluids (liquids and gases)

Gear Flow Meter

This type of flow meter has oval shaped gears with fitting teeth which control the amount of fluid passing through. The flow rate is calculated by number of times the gears are filled and emptied. These meters have high accuracy and are used for measuring low flow and for high viscosity fluids. It is very easy to install these types of meters because it requires no pipe.



Thermal Flow meters

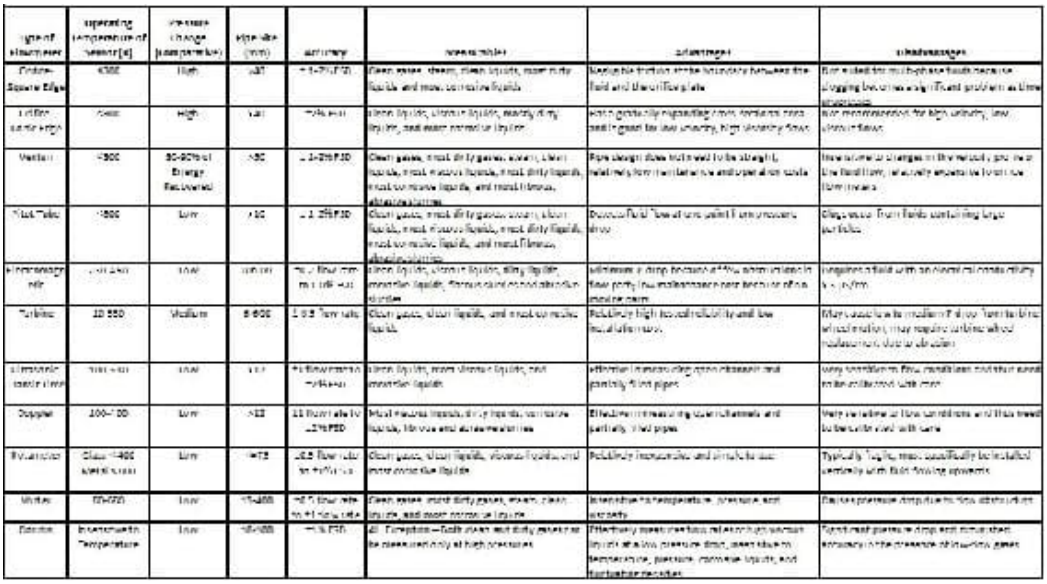
These flow meters take advantage of the thermal properties of the fluid to measure the flow of the fluid in a pipe. In most thermal flow meters, a measured amount of heat is applied to the heater of the sensor. Portions of this heat is lost to the fluid as it flows. Accordingly, as the fluid flow increases more heat is lost. The amount of heat lost is tracked using a temperature measurement instrument in the sensor. Then an electronic transmitter uses the heat input and temperature measurements to calculate the fluid flow, taking into account the thermal properties of the fluid.

Common applications of thermal flow meters are measuring the flow rate of clean gases like air, nitrogen, hydrogen, helium, ammonia, and argon. Most thermal flow meters are used to measure pure gases that would be used in laboratory experiments or semiconductor production. Mixtures like flue stack flow can also be measured but the mixture compositions must be known to use the appropriate thermal properties of each component in the mixture. The key advantage of this technology is its dependence on the thermal properties which are mostly independent of the gas density.

However caution should be taken if this type of flow meter is used to measure a fluid of unknown or varying composition. Additionally, thermal flow meters should not be used to measure abrasive fluids because they may damage the sensor. Some fluids can coat the sensor causing it to be inoperable and must be regularly cleaned to be useful.

Table of Flow Meters

The following table outlines specifics for most of the flow meters. This may be helpful in answering example 1. Accuracy is given in terms of FSD (full scale deflection) which is the deflection of meter's pointer to the farthest point on the scale. This implies the highest measurement.



Flow Profile Distortion

In the real world, the flow profile is not always symmetrical. Pipe fittings such as elbows, tee-pieces, and reducers can change the flow profile. One example of a fitting that alternates the flow profile is a sharp elbow which causes pure swirls throughout the fluid. Some flow meters are more sensitive to particular types of flow distortion. More complex flow conditions produce better velocity profile but there is a trade off since they are more expensive and give higher pressure drops.

Turndown Ratio

The turndown ratio is a term used to describe the range of accurate operability of a specific flow meter. This rangeability is critical when selecting flow meters for specific applications. Typically, in a plant setting gas flow may not be constant, but accurate measurement of gas flow is needed ranging from no flow to full flow.

For example if nitrogen gas is being used in a plant conducting multiple batch reactions, sometimes the little nitrogen (100 m3/min) will be needed and other times full nitrogen flow (1000 m3/min) will be needed. For this system, the turndown ratio is 10:1. Accordingly, a flow meter must have a turndown ratio of at least 10:1.

For each type of flow meter, the turndown ratio is limited by the theoretical and physical constraints. For example, for an orifice meter the accuracy may be compromised near the limits of the rangeability. Orifice meters create a pressure drop in the measured fluid which is proportional to the velocity squared. If the range of differential pressures becomes too large the overall accuracy of the flow meter at its range limits may be inconsistent.

Turndown ratio can be expressed as:

$$TR = qmax/qmin \tag{3.5.1}$$

where

The table below shows a list of typical turndown ratios for different flow meters.

Type	Turndown Ratio
Orifice Meter	3:1
Turbine Meter	10:1
Rotary Displacement Meter	10:1-80:1
Diaphragm Meter	80:1
Ultrasonic Meter	50:1
Thermal Meter	10:1-100:1

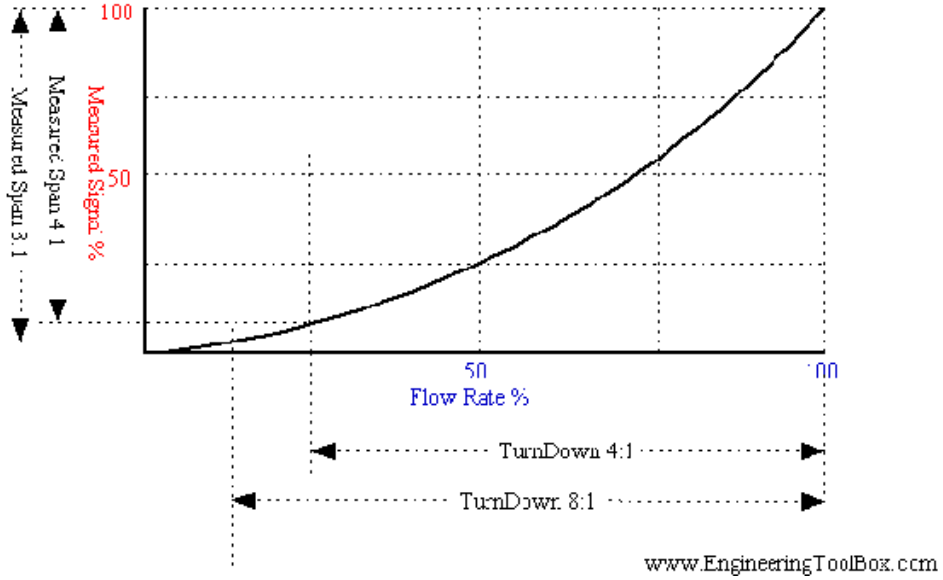
Example - Turndown Ratio for an Orifice Meter

The turndown ratio - TR - for an orifice meter with maximum flow of 12 kg/s and a minimum flow of 3 kg/s can be calculated as:

TR = (12 kg/s) / (3 kg/s) = 4 - normally expressed as turndown ratio of 4:1

This is a typical turndown ratio for a orifice plate. In general a orifice plates has turndown ratio between 3:1 and 5:1.

Turndown Ratio and Measured Signal



The graph above shows how the turndown ratio effects the measured signal % and flowmeter %. With an increased turndown ratio there is a larger range that the flowmeters can operate within. In a flow meter based on the orifice or venturi principle, the differential pressure upstream and downstream of an obstruction in the flow is used to indicate the flow. According the Bernoulli Equation the differential pressure increases with the square of flow velocity. A large turndown ratio will cramp the measurement signal at low flow rate and this is why venturi and orifice meters are not accurate at low flowrates.

Example 3.5.1

a) Oil is flowing in a medium sized pipe which is insensitive to temperature and pressure. Pressure drop measured is relatively low. Which type of flow meter is recommended?

Coriolis flow meter because it effectively measures flow rates of high viscous liquids at a low pressure drop. Coriolis flow meters are also insensitive to temperature and pressure.

(b) Why is Doppler only recommended for liquids containing suspended solids or gases?

The Doppler sensor transmits sound into the flow stream, which must be reflected back to the sensor to indicate flow velocity. Gas bubbles or suspended solids act as reflectors for the Doppler signal. Doppler flow meters are recommended for liquids containing solids or bubbles with a minimum size of 100 microns and a minimum concentration of 100 ppm.

(c) When would you use Vortex flow meters?

Vortex is mainly used to measure steam flow rates. They are also used to measure gases and liquids with low viscosity that contains hardly any solids.

(d) A conductive liquid is flowing in a pipe (1000 mm size) at the temperature of 430K. Which type of flow meter is recommended to measure the flow rate?

Electromagnetic flow meter is used in this case as it measures flow rates of conductive liquids.

Example 3.5.1

Kerosene at 20 deg C flows at 18 m^3 / h in a 5 cm diameter pipe. If a 2 cm diameter thin plate orifice with corner taps is installed, what will the measured drop be in Pa? Why is orifice meter used in this case?

Data given: C_d=0.6030; μ = 192E-3 kg/ (m s) for kerosene; ρ = 804 kg/m^3

Solution

Orifice meter is used in this case because Kerosene is a clean and non viscous liquid.

Usually, pipe velocity is calculated using the velocity formula to calculate Re and C_d values. The pipe velocity is:

$$v = \frac{Q}{A} = \frac{\left(\frac{18}{3600}\right)}{\left(\frac{\pi}{4}\left(0.05^2\right)\right)} = 2.55\mathrm{m/s}$$

In this case, we already have the C_d. So,using orifice pressure drop formula, calculate

$$Q = \frac{18}{3600} = 0.6030\frac{(2\pi)(0.02)^2}{\sqrt{1-\frac{(0.04)^2}{(0.02)^2}}}\sqrt{\frac{2\Delta p}{804}} = 0.005m^3/h$$

Final step is to solve for Δp to give 273 kPa.

$$\Delta p = \frac{(0.005)^2 * \left(1-\frac{(0.04)^2}{(0.02)^2}\right)(804)}{(0.6030)^2(2\pi)^2(0.02)^4(2)} = 273kPa$$

Sage's Corner

Selection of an Industrial Flow Sensor

References

Contributors and Attributions

- TR = Turndown Ratio
- qmax = maximum flow
- qmin = minimum flow
- Cole-Palmer Catalog, 2005-2006
- [5]Engineering fundamentals' website
- Perry, Robert H., Perry's Chemical Engineers' Handbook, 7th Edition. McGraw-Hill, 1997.
- Richardson, J.F. and Peacock, D.G., Chemical Engineering, 3rd Edition, Volume 3. Butterworth- Heinemann, 1994.
- [6]Process Measurement and control Products
- [7]Universal Flowmeters Presents Flowmeters.com

- [8]Transactions Volume 4: Flow and Level Measurement
- [9] Wikipedia entry on Coriolis Mass Flowmeters
- [10]Omega Literature: Flow and Level Measurement
- Geankoplis, Christie John, Transport Processes and Separation Process Principles, 4th Edition, Pearson Education, 2003.
- McKetta, John J., Cunningham, William A., Encyclopedia of Chemical Processing and Design: Volume 43 - Process Control: Feedback Simulation to Process Optimization, CRC Press, 1993.
- Authors: (September 14, 2006) Ashwini Miryala, Kyle Scarlett, Zachary Zell, Brandon Kountz
- Stewards: (September 5, 2007) Jeremy Goldfarb, Daniel McClure, Lilian Ngobi, Daniel Wheeler

This page titled 3.5: Flow Sensors is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

3.6: Composition Sensors

Introduction

There are many useful analytical tools, such as photometric analysis, electrometric analysis, chromatography, mass spectrometry, thermal conductivity, and various physical property measurements (density and specific gravity), which can be used to determine the composition of mixtures. A wide array of methods to measure composition are available, so it is important to choose the best method given a set of conditions. The first step is to decide between measuring physical or chemical properties to determine the composition. If you choose to measure a physical property, make sure that it is unique to the desired component of the mixture and will accurately allow you to determine the composition. The goal of this article is to explain the various analytical methods and tools used to determine the composition of a given sample. After reading this, you should be able to determine which method of composition measurement is most appropriate for a given circumstance.

Types of Testing: On-line v. Off-line

On-line analysis is the continuous monitoring of the composition of a sample, which is under the influence of a control system and directed by an actuator which can respond and regulate the operating conditions in real time such that the desired set points are maintained. On-line testing can either be performed in-line or by slip stream testing. In on-line testing, the sensor is attached directly to the line and provides feedback via a transmitter. On the other hand, in slip stream testing, a side stream of the process runs alongside the main line. Such an apparatus closely resembles the set up of a bypass. The slip stream process conditions can be continuously manipulated to make the measurements easier to obtain. Similar to in-line testing, the sensor is directly attached to the slip stream and provides feedback through a transmitter. Advantages of on-line analysis include an immediate and continuous feedback responding to changes in process conditions. The main disadvantage of on-line testing is that it is usually much more complicated and more expensive than off-line testing. Also, off-line testing is more robust and has more varied applications, while on-line testing may not work in every situation. On-line testing prevents the continued production of undesired product, with an immediate response and correction of the flawed material.

Off-line analysis involves the extraction of a sample from the process or reaction, and its subsequent testing in a machine that may be situated at a location far away from the process line in a lab. In this case, a sample is manually removed which is later sent to the composition analyzer. The results of the analysis are examined and then they are sent to the control system or actuator to make the appropriate adjustments. On-line analysis is comparable to off-line testing, the primary difference being that in on-line testing, the samples are analyzed on a machine that is next to the process line. This greatly reduces time lost in transporting the sample, though it still permits the prospect of introducing contaminants into the sample. In off-line analysis, some of the disadvantages include sample dead-time, which is the time lost during transportation, variability of sample testing locations, and lag time for adjustments to be made to the process. In off-line analysis, the lag time between removing the sample from the system and receiving the results of the test could cause significant losses to the company, since the defective product is produced and the process is not corrected until the results are received. This prolonged defective processing could result in the accumulation of losses worth thousands of dollars, which could have been avoided if an on-line analysis system was set up in place.

While both methods are widely used in industry, there is a push towards more on-line testing. A paper published in 2003 urging drug companies to adopt more in-process analytical testing such as on-line testing, endorsed by the Food and Drug Administration reflects this trend. This has been a part of a broader effort by the FDA to encourage companies to move towards better manufacturing practices. While on-line testing maybe comparatively expensive to install at first, the savings from the process would not only recoup the costs associated with installation, but also prevent significant losses. Many companies have adopted lean manufacturing techniques, and one of its hallmarks is to correct a defect as early in the production process as possible, not allowing defective products to move on to the next processing step. Allowing defective products to perpetuate through the system may even contaminate other non-defective materials in the production process.

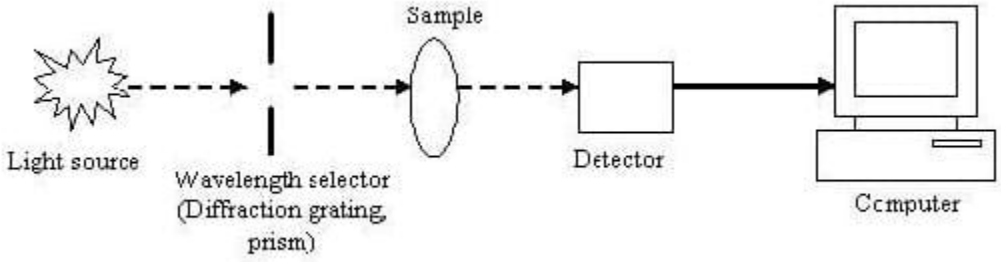
Standards and Calibration Curves

Before discussing the different ways to measure the composition of a sample, one needs to understand that composition sensors use standards or calibration curves to measure an unknown composition against a known one(s). The standard can either be an internal standard against which all other measurements are compared, such as a reference cell or a series of standards used

to create a calibration curve. With a calibration curve, a series of known standards are measured with the sensor and the signal produced by the sensor is graphed on the y-axis of a plot; the known information about each sample is graphed on the x-axis. From this information, a relationship can be developed between the signal output from the sensor and the known quantity that you are trying to measure. The data measured for the unknown sample can then be compared to the calibration curve graphically or using the equation written to describe the curve. Using this information, the measurement made by the sensor can be interpreted as a useful composition measurement of weight percent, volume percent, mole percent, etc.

Photometric Analysis

Photometric analysis is the measurement of the intensity of visible light and other electromagnetic (EM) waves. Through the measurement of these values, the composition of samples or flows can be determined in various different ways. A wide range of photometric devices based on many differing principles are used in the chemical engineering industry for this purpose. A few of the most common instruments are covered in this article.



Types of Photometers

A photometer is any type of instrument used to measure characteristics of light. There are two broad categories of photometers: dispersive and non-dispersive.

Dispersive Photometers
In dispersive photometers, the light from the source is dispersed and a narrow spectral band is selectively directed to the sample and detector. A monochromator is usually the tool that performs this function. The dispersing element of the monochromator can be a diffraction grating, prism, and interference filters. Diffraction gratings are the most common; for each position of the grating a narrow band of dispersed light passes through an exit slit. Dispersive devices can scan across a spectrum and make measurements at several wavelengths. This capability allows them to be used for the analysis of multiple components.

Non-dispersive Photometers
Non-dispersive photometers use a narrow-band-pass filter to block out a large amount of undesired radiation. They make measurements at selected discrete wavelengths. The filter passes radiation at the selected reference and measurement wavelengths. The reference wavelength filter selects a band where none of the components present in a process stream absorb radiation; the measurement wavelength filter is selected to match the absorption band of the component being analyzed. The ratio of the transmitted light at the reference and measurement wavelengths is measured by the photometer. Non-dispersive photometers are usually used to measure a single component in the process stream.

An example of such a device is the Non-Dispersive Infrared (NDIR) analyzer that is frequently used to measure the concentrations of certain gases in a mixed gas flow. The NDIR analyzer uses a reference non-absorbent gas such as O_2, Cl_2 and N_2 and compares that to a sample gas like CO, CO_2, SO_2, CH_4, etc. Two beams of infrared radiation are used: a reference beam for the non-absorbent gas and an analyzing beam for the sample.

Spectrophotometers

Spectrophotometers are widely used as composition sensors. Spectrophotometers are photometers that measure the intensity of light at specific wavelengths. Many different types of spectrophotometers exist, and they are typically categorized based on the ranges of wavelengths they monitor: ultraviolet (280-380 nm), visible light (400-700 nm), infrared (0.7-1000 µm), or microwave (1-300 mm). Other important distinctions are the measurement technique, the method of obtaining the spectrum, and the source

of intensity variation.

The emission spectrum recorded by the spectrophotometer typically consists of a series of narrow peaks of varying heights, also known as spectral lines. The height or amplitude of those lines can be related to the concentration or abundance of the specific material through a calibration curve and their position on the spectrum distinguishes the components of the material. While there are numerous methods for obtaining a spectrum, the most common application of a spectrophotometer is for the measurement of light absorption.

Absorption

The degree of absorption of electromagnetic waves exhibited by the components in the sample is a distinctive feature that can be used to reveal the composition of the sample. Absorption is related to the inverse of transmittance by:

$$A = \log\left(\frac{1}{T}\right)$$

where A is the Absorbance (Absorbance unit, AU) and T is the Transmittance (as percentage).

Spectrophotometers based on absorption have a radiation source that will emit a specified range of EM waves to be absorbed by the sample by measuring the intensity of light that passes through the sample. A useful relationship used to quantitatively determine the concentration of an absorbing species in solution is the Beer-Lambert Law:

$$A = -\log\left(\frac{I}{I_0}\right) = \epsilon\, c\, L$$

where

- I = Intensity of incident light at a particular wavelength
- I_0 = Intensity of transmitted light
- *epsilon* = molar absorptivity or extinction coefficient (1 / *M* * *cm*)
- *c* = concentration of the absorbing species
- *L* = path length through the sample

According to the Beer-Lambert Law, the absorbance of a solution is directly proportional to its concentration. The constant *epsilon* is a fundamental molecular property for a given solvent. It can be found from references tables or determined through the construction of a calibration curve. The *Beer-Lambert Law* is useful for characterizing many compounds, but does not hold for all species. For example, very large, complex molecules such as organic dyes often exhibit a 2nd order polynomial relationship between absorbance and concentration.

Two widely-used devices that measure the absorbance of samples are *Ultraviolet* and *Infrared* spectrophotometers.

Ultraviolet (UV) Spectrophotometers

UV spectrophotometry is a useful method for detecting and quantifying substances in solution. Certain compounds in the gaseous or liquid state absorb ultraviolet radiation, resulting in absorption peaks that can be correlated with specific functional groups. These species include aromatic and carbonyl compounds in organic substances and transition metal ions in solution. This property is useful for categorizing and measuring the composition of these types of samples, however this property cannot be used as a specific test for a given compound. The nature of the solvent, the pH of the solution, temperature, high electrolyte concentrations, and the presence of interfering substances can influence the absorption spectra of compounds, as can variations in slit width in the spectrophotometer. Many organic solvents have significant UV absorption and therefore will add peaks to the species dissolved in them. Solvent polarity, high electrolyte concentration, and pH can effect the absorption spectrum of an organic compound by causing protonation or deprotonation of the compound. Temperature can also cause complex molecules to change in conformation and can also alter the properties of the solvent used. Variations in the slit width will alter the effective bandwidth, distorting wavelength measurements.

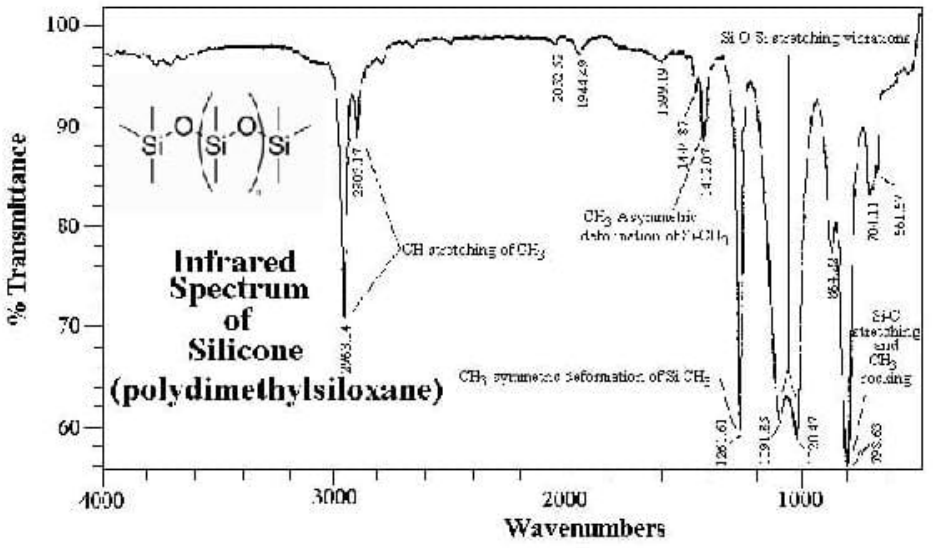
The basic components of the overall analyzer are the UV light source, sample holder, optical filters, and detector (typically a phototube). Optical filters (typically a monochromator or diffraction grating) separate the wavelengths of light going through the sample and ensure that only the desired wavelengths reach the detector.

Most UV analyzers are dispersive photometers that also function in the visible light spectrum and are called UV-Vis spectrophotometers. Since these instruments are generally found in laboratories, they are generally used for offline analysis. However, there are inline UV machines and sensors available for real-time inline measurements of process flow streams. UV spectrophotometers used throughout many industries for various applications.

Infrared (IR) Spectrophotometers

Infrared spectroscopy is based on the fact that there are specific frequencies at which molecular bonds rotate or vibrate when exposed to discrete energy levels of radiation. A beam of infrared light is passed through the sample, either one wavelength at a time with a monochromatic beam or using a Fourier transform instrument that measures all wavelengths at once. Inspection of the transmitted light reveals how much energy was absorbed at each wavelength, this data is then plotted (see example below). This plot allows the frequency of the vibrations to be associated with a particular bond type. The characteristic IR absorption spectra obtained for a sample can be used to identify the compound, while the amplitude of absorbance is proportional to the analyte concentration. IR analyzers are used to analyze gaseous or liquid compounds that can absorb IR waves. The system components are generally similar to those of an UV spectrophotometer, with the exception of the IR light source.

IR instruments have been used regularly in a fixed laboratory setting for decades now. Both dispersive and non-dispersive IR techniques exist. More recently, manufacturers of instruments have significantly reduced their size and power requirements. In addition to inline IR sensors for process streams, portable instrumentation is available for field analysis such as the monitoring of stack gases. Another relatively modern technology, the Fourier Transform Infrared (FTIR) Spectrometer, is widely used in industry today because of its high sensitivity and broad range of application. For example, FTIR analysis is used in the quantification of harmful organic peroxides in lubricants and fuels.



Radiation

Instead of measuring the absorption of light, many spectrophotometers measure the radiation of EM waves. Samples may radiate EM waves in many forms such as heat, luminescence, fluorescence, and phosphorescence. The measurement of such radiation is often employed in biological and pharmaceutical areas because of its high sensitivity and specificity. For example, fluorescence spectrophotometers are used to assay the presence and concentration of target biological compounds. Fluorescent probe molecules are introduced to the sample that will attach selectively to the biological molecule of interest and the device will then detect and measure the amount of fluorescent radiation. In these experiments, the appropriate standards must be measured to properly calibrate the radiation measured with sample concentration. Additionally, contamination of the sample can affect the radiation patterns.

Examples of other techniques utilizing the radiation EM waves include flame emission spectroscopy, particle-induced X-ray emission, and X-ray fluorescence. These methods are strictly used for laboratory analysis.

Photometry Using Visible Light

The manner in which certain materials interact with visible light can be used to analyze the composition of the sample material. Refractometers, turbidimeters, and opacity meters are all analyzers which utilize this principle.

Refractometry

The theory behind refractometry is that visible light bends (or refracts) when it passes two opaque media of differing densities. Snell's Law relates the angle of incidence α to the refractive index η by:

$$\frac{n_1}{n_2} = \frac{\alpha_2}{\alpha_1}$$

α_1, α_2 = angles of incidence

η_1, η_2 = refractive indices of the two media (unitless)

Alternatively, the critical angle may be used to analyze the composition. The critical angle is defined as the angle of incidence that produces a 90° refraction with respect to normal of the interface i.e. $\alpha_2 = 90^o$. Thus, Snell's Law becomes:

$$\frac{n_1}{n_2} = \frac{1}{\sin \alpha_c}$$

α_c = critical angle

Please refer to the diagram below.

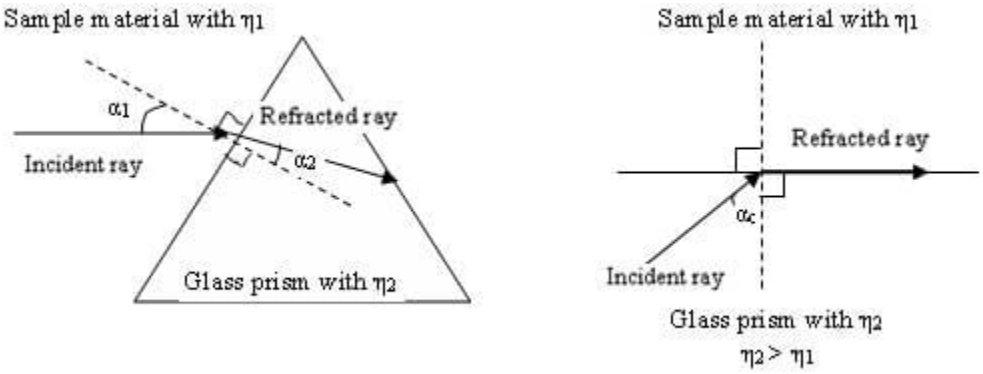


Figure 1: Light refraction at a glass-medium interface

Different materials will have different refractive indices; the experimental refractive index measured can be compared to a list of refractive index values at a data bank. Continuous process or online refractometers generally measure the changes in the critical angle at a glass-sample interface caused by the variations in composition concentrations. Refractometers are used in numerous industries such as food, chemical, metalworking industry because of its high accuracy and repeatability, low maintenance and easy installation. An example is the use of a refractometer to monitor the amount of juice concentrate in juice.

Turbidimeters

Turbidimeters measure the turbidity, or sample clarity, of fluids. A light beam passes through the analyzer and if the water, or other chemical that is to be tested, is "pure," then the light would travel undisturbed. However if the sample contains suspended solids or particles, the light would interact with the particles causing the particles to absorb the light and reflect it into different directions. An example is smoke in air that gives the air a hazy look.

Most modern turbidimeters are nephelometers. Nephelometers or nephelometric turbidimeters measure the *scattered* light reflected off the suspended particles in the sample. This instrument contains a light source, some sort of container to hold the sample, and photodetectors to sense the scattered light. A tungsten filament lamp is most commonly used as a light source. If the suspended particles, typically measured in nanometers, are smaller than the wavelength than the incident light, the reflected light

is scattered equally in all directions. If the particles are larger than the wavelength of the incident light, the pattern of scattering will be focused away from the incident light.

Nephelometers are usually used when solids are present in the sample in small concentrations. When there is high turbidity in a sample, multiple scattering occurs and the turbidimeter may no longer be able to analyze the fluid. Disadvantages of nephelometers include being prone to fouling and incapability of analyzing fluids that have high concentrations of solid particles.

Turbidimeters are important to industry, specifically for chemical plants, oil refineries, and waste water treatment plants. For example, nephelometers are used in detecting the pollutants in water in water quality monitoring.

Opacity Monitors

Opacity monitors measure the *attenuation* of light due to *scattering and absorption* by the sample. Attenuation is defined as the decrease in intensity and amplitude of the measured light beam when compared to the emitted light beam. Opacity monitors differ from turbidimeters in that they analyze samples by measuring the percentage of transmission of the light source that passes through the sample and NOT the scattering of the light source. A common application is the measurement of smoke density in chimney stacks. The density of the pollutant or particle is expressed as percent opacity, percent transmittance, or optical density.

$$\% \text{ opacity} = 100 \times \text{opacity} = 100 - \% \text{ transmittance}$$

Electrometric Analysis

Electrometric analysis uses the principles of electrochemistry to analyze the composition of different samples.

Conductivity Cells

Conductivity cells measure a liquid's ability to conduct an electric current. Since liquids conduct current due to the presence of ions, the composition can be determined by measuring the concentration of the ions present. Conductivity cells determine concentration from the definition of molar conductivity (lambda). Molar conductivity is a property that can be influenced by temperature. It also varies with concentration if the chemical being measured is a weak electrolyte. It is defined by the following equation:

$$\Lambda = \frac{\kappa}{C}$$

- Λ = molar conductivity [Sm^2/kmol]
- κ = conductivity [S]
- C = concentration [kmol/m^3]

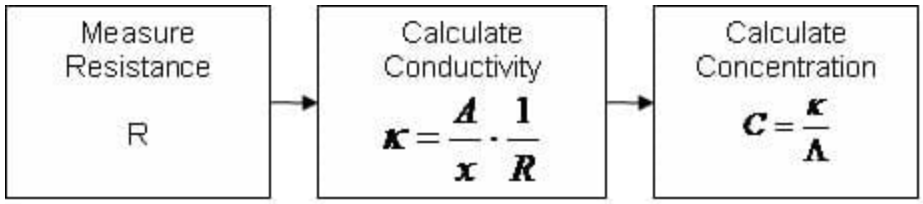
Since it is difficult to directly measure conductivity, conductivity cells measure the resistance (R) of a sample liquid between two electrodes when a small current is applied. Resistance is related to conductivity by the following equation:

$$\kappa = \left(\frac{A}{x}\right)\left(\frac{1}{R}\right)$$

- κ = conductivity [S]
- A = area of electrode [m^2]
- x = distance between electrodes [m]
- R = resistance [ohm, Ω]

Note that the electrode area, A, and distance between the two electrodes, x, are fixed for a particular sensor and are often expressed as a cell constant (k_c).

The following flow chart shows how composition is determined in terms of concentration:



Conductivity cells can be used in flow system and installed directly into the system of pipelines. Conductivity cells can come in three types: two electrode probe, multiple electrode probe, or a non-contact, electrode-less probe (toroidal). The two electrode probe is one of the most common types and typically uses carbon or stainless steel as the electrodes. The multiple electrode probed has either two, four, or six electrodes. This probe has the advantage of reducing errors due to polarization (electrochemical reaction on the electrode surface) and fouling. Finally, the toroidal probe has no electrodes and instead measures the AC current in a closed loop of solution using two embedded toroids (a magnetic core with wire wrapped around it to make an inductor). The electrode-less probe is good for solutions that tend to have problems with polarization or oily fouling. In-line conductivity cells are suitable for temperatures up to 480K and pressures up to 1700 kN/m_2.

Since concentration is found from molar conductivity, which is a property of the fluid that is strongly dependent on temperature, a temperature sensor is often paired with a conductivity cell so temperature effects can be accounted for in concentration calculations.

Conductance cells are best used in applications where a single type of ion is present in a dilute solution. Higher concentrations lead to complications that cannot be described by the simplistic linear relation between concentration and conductivity, since molar conductivity starts to vary with higher concentrations of weak electrolytes.

Ion Selective Electrodes

pH Electrode

The pH electrode is commonly used in both the lab and industrial setting. This electrode is sensitive to the concentration of the H^+ ions. The pH electrode basically measures the potential difference between a reference solution and the sample. Please refer to the pH sensor section for examples and more information on relevant theory.

Other Ion Selective Electrodes

Electrodes can also be specific to other ions such as: Na^+, K^+, Ag^+, ClO_4, BF_4, SO_2, and CO_2. An electrode's ion sensitivity depends mainly on the method of separating ions in the solution. The biggest problem with ion selective electrodes is the problem of not enough selectivity for the desired ion, and subsequent interference by other ions with similar chemical properties to the desired ion. Most electrodes use a membrane to separate a specific ion, though other techniques exist. Selective ion electrodes can be divided into the following three categories based on the mode of separation: solid membrane (glass or crystalline), liquid membrane (classical ion exchanger with a neutral or charged carrier), and membrane in a special electrode (gas sensing or enzyme electrode). Examples of ion selective electrodes used in industry include nitrate sensors for sewage treatment or river quality monitoring, hydrazine and sodium sensors for in boiler water, and choride and fluorine sensors for in water treatment plants.

Oxidation-Reduction Potential (ORP) Sensors

ORP sensors, or also called Redox sensors, can determine concentration of ions that exist in two different oxidation states in solution. For example, this type of electrode can be used to analyze the composition of a solution with Fe^{2+} and Fe^{3+} ions. This type of electrode measures the potential caused by a solution's reduction or oxidation ability.

This type of sensor directly measures electric potential which indicates the reducing or oxidizing strength of a solution. The reducing or oxidizing power of a solution depends on whether the higher or lower oxidation state is preferred by the ion.

When an ion favors the higher oxidation state (or has reducing properties), the ions tend to give up electrons and the electrode becomes more negatively charged.

When an ion favors the lower oxidation state (or has oxidizing properties), the ions tend to gain electrons and the electrode becomes more positively charged.

When purchasing an ORP sensor from a vendor, many models come already packaged with a pH sensor, since these two measurements are often both monitored to keep a process running at the right conditions. For instance, in the pool and spa industry, its important to monitor both the concentration of sanitizers in the water with an ORP sensor, as well as measure the pH of the water.

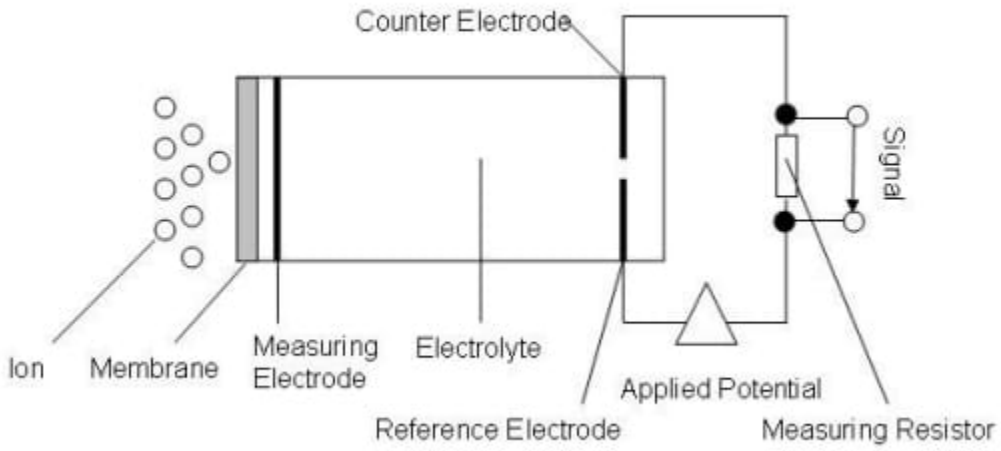
Polarographic Sensors

Polagraphic sensors work by performing electrolysis with two types of electrodes: one which is polarizable and one which is not. They operate very similarly to voltameters, but with voltameters the potential is changed in a controlled manner and the voltage measured. The oldest system of polarographic measurement used dropping mercury electrodes. Here the principle is that that the surface tension of mercury changes with how much charge is applied between the top and bottom of the electrode, and this surface tension change can be seen in how the weight of the mercury drops varies with time.

More modern types of polarography use two different types of metals which are connected in a solution with electrolytes, and electrons will move towards the more positively charged metal. This movement of electrons induces a current, which eventually stops once the charges are balanced (polarized). If all the charges are balanced within the system, no current will flow. Therefore a small potential is applied to a polarographic sensor so that a current exists for the sensor to measure.

The sample being analyzed needs to have the ability to depolarize the system, or tip the balance of charges. The ion of interest crosses a selective membrane and reacts by an oxidation/reduction reaction with a metal surface (often called the measuring electrode). The reaction between the ion and electrode causes an uneven distribution of charges. Once a system is depolarized, the potential of the system increases since a more positive and a more negative region exists separately. The increase in potential causes current to flow which is the signal measured by the sensor. This creates a graph showing the relationship between applied current versus measured potential. A disadvantage to this type of measurement tool is that this graph requires a calibration curve in order to interpret the data, and so analysis of composition using polarography can't be done in-line. The advantage of this type of method is that it is highly accurate and very reproducible.

Below is a basic schematic of a polarographic sensor.



Polagraphic sensors are most commonly used assessing pollutants in air such as Cl_2, *HCl*, *HBr* and *HF*. It an also be used to measure O_2 and other inert gases in flue.

High Temperature Ceramic Sensors

These types of sensors have a heated section of zirconium oxide (ZrO_2), also known as zirconia, that is slightly doped with yttria (Y_2O_3). This stable lattice structure is able to measure oxygen ion conduction at high temperatures. This sensor is exclusively used for measuring O_2 and is often used to measure O_2 composition in flue gases.

Chromatography

Chromatography is a set of physical methods used to separate and thus analyze complex mixtures. Chromatography consists of a mobile phase, usually a gas or a liquid, and a stationary phase, typically a solid or a liquid. These two phases are combined in a column of the stationary phase based on the type of chromatography one wishes to complete (i.e. Gas-liquid, Gas-solid, Liquid) in order to separate the mixture. The retention time of a component is the time before the component elutes from the column. The retention time and elution order are both based on the relative interaction between each solute and each phase. The stronger the solute's interaction with the mobile phase, relative to the stationary phase, the sooner the solute will elute from the column. Solute-phase interaction is a function of the charges of the molecules, the absorption of the stationary phase, and the relative solubility.

Chromatography by itself does not detect substances, but it is routinely coupled with a detection unit. Each substance is detected as leaves the chromatograph. The retention time for each substance can then be compared to a standard to determine what substances were in the mixture. For unknown substances, the elution order can determine some properties of the substance. For example, a substance that initially elutes after ethanol in a nonpolar column is less polar than ethanol. However, other composition detectors are far better suited for analysis of unknown substances.

Chromatography is the most widely used composition detecting process for on-line analysis of samples within a complex mixture. Due to its precision, chromatography is commonly used to separate or purify samples within the chemical or biochemical industries. Depending on the type of chromatography used, even delicate mixtures like proteins can be analyzed. However, chromatography is primarily useful when there are some expectations for the components in a mixture; chromatography is not best for a completely unknown sample, but it is useful for detecting the composition of a process stream where most components are known and a suitable standard exists.

Types of Chromatography

Gas-Liquid Chromatography (GLC)
GLC is one of the more common types of chromatography. It is commonly referred to as GC (gas chromatography). In a typical run, the liquid mixture to be analyzed is added to the system and vaporized before entering the column. A carrier gas, often Helium, carries the vaporized sample through a column. The column can either be packed with stationary phase, or the column can have a very small diameter in which case it is lined with stationary phase (capillary tube). The gas mobile phase runs through the stationary phase within the column at various rates determined by relative volatility and affinity to the stationary phase. The longer a sample molecule spends in the gas phase, the faster it will elute from the column. For this reason, temperature and the chemical identity of the coating on the stationary phase are two important variables which may need to be manipulated to receive clear separation of the components in the mixture. In general, lower temperature will result in better separation but longer elution times and sample spreading due to flow effects. Efficient use of this apparatus requires a balance between these competing effects.

GC is relatively simple for a technician to use; a calibrated gas chromatograph can run several samples without a large amount of sample preparation. Gas chromatography works well on substances that vaporize below 300°C and are free of ions; however, it does not work well on substances that decompose below 300°C.

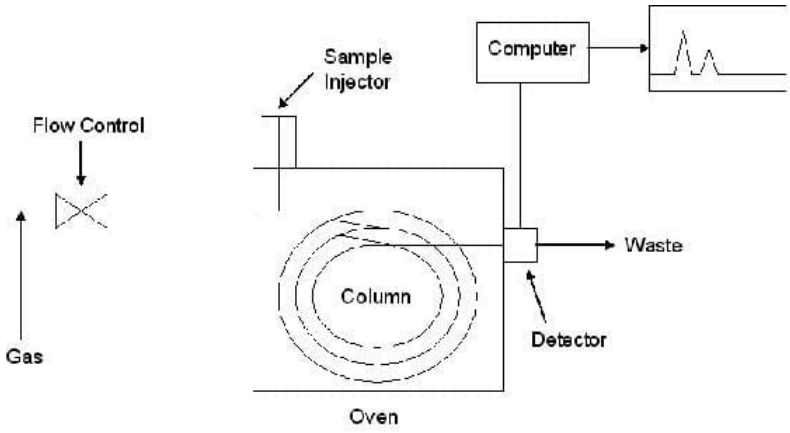


Figure: Diagram of Gas Chromatography

Gas-solid Chromatography (GSC)
GSC uses a gas mobile phase which runs through an absorbent solid (e.g. silica gel or alumina) stationary phase within the glass column as determined by absorption. This method operates along the same principles as gas-liquid chromatography. It is most commonly used with gases (as opposed to a vaporized substance) and with low molecular weight hydrocarbons.

Liquid-Chromatography (LC)
LC analyzes low volatility liquids within a solid stationary phase. Separation occurs from adsorption/desorption due to the solutes having different degrees of attraction to the stationary phase, typically due to the stationary phase having different polarity than the mobile phase. It is common to use a mixture of solvents to accurately separate mixtures. Liquid chromatography is used in several industries, including the pharmaceutical and petrochemical industries.

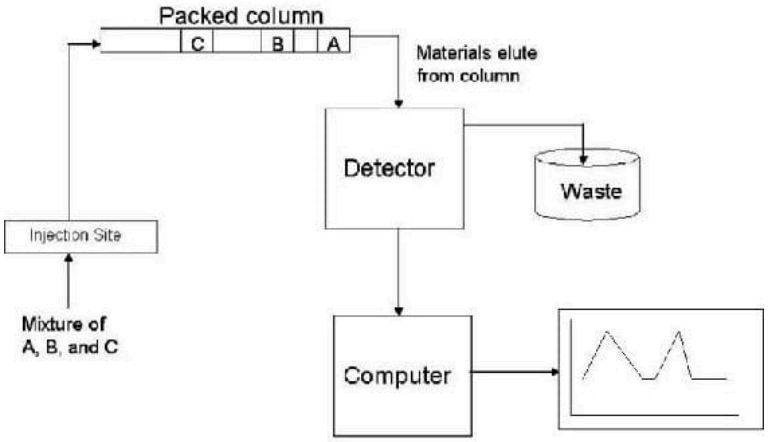
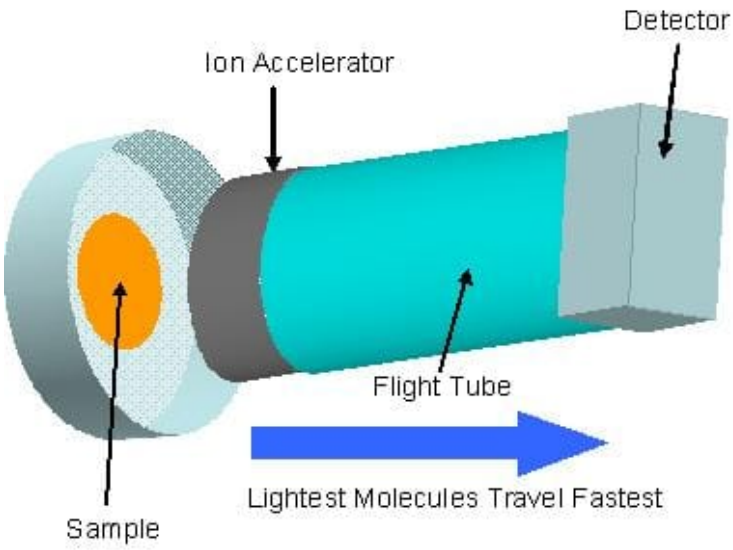


Figure: Example of Liquid Chromatography

Mass Spectrometry

Mass spectrometry is among the most precise compositional sensing tools, and, accordingly, one of the most expensive. The basic idea is illustrated below:



A sample is ionized—this may be performed by a number of techniques—the ions are subjected to an electrical force and accelerated through a tube. Because the electrical force applied to each molecule is the same, a molecules acceleration through the tube depends on its mass (F=ma). The acceleration is also dependent on any charge on the molecule due to magnetic attractive or repulsive forces. At the end of the tube, a detector calculates the time of flight for each of the molecules.

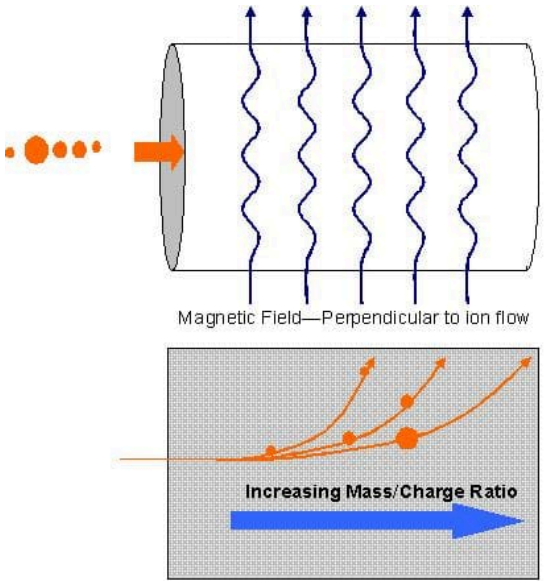
Mass spectrometry is essentially a tool to create a spectrum of distinct mass/charge ratios. It is very often used after chromatography separation techniques to serve as a molecule identification technique.

MS Components

There are 3 fundamental components for mass spectrometry—an ion source, a mass analyzer, and a detector.

The ion source is the component responsible for ionizing the sample. There are many different methods including chemical, electrical, or laser ionization. Selection of the proper ion source depends on the characteristics of the sample (phase, biologically active, etc.).

The mass analyzer is the technique and tool used to differentiate the ions in a spectrum based on the mass/charge ratio. The technique described earlier is the Time-of-Flight technique. Another common method is a sector field mass analyzer where the molecules are accelerated through a tube and a magnetic force is applied perpendicular to the direction of flight of the molecules. The molecules will deflect at different magnitudes depending on their size.

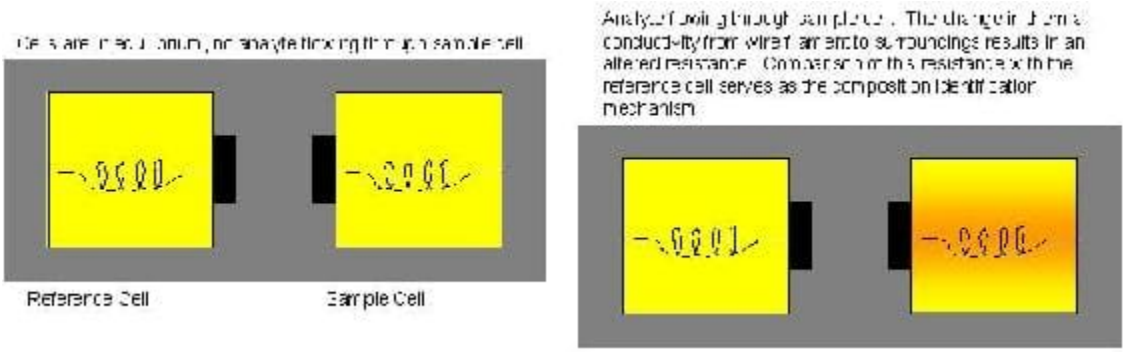


The detector generally amplifies changes in an electrical current that occur when molecules hit the detector.

Thermal Conductivity for Gases

Changes in thermal conductivity may be used to detect variations in a mixture of gases. When a heat source is present, gases and vapor will conduct heat. The change in heat between the source and gases will stabilize to a common temperature. The temperature is mainly dependent on the thermal conductivity and therefore the composition of the gas.

Typical equipment for thermal conductivity analyzers includes: a reference cell, a sample cell, a combined heat source (wire filaments or thermistors), and detector. The sample cell and reference cell are usually placed in a holder where the detector may be mounted. The reference is an identical cell of the sample cell, through which a known gas will flow. The reference-detector resistance will be constant and the sample-detector resistance will vary depending on the composition. Resistance is a function of temperature and the output from the detector bridge will be a function of sample composition.



Thermal conductivity, like mass spectrometry, is often used in combination with gas chromatography

Physical Property Measurements

Physical properties can be used to indirectly measure the composition, but this technique is not as versatile as those described above, since it only applies to a limited number of circumstances. The advantage to using physical properties to indirectly measure composition is that these types of measures are often cheaper than the types of sensors described above.

Density and Specific Gravity

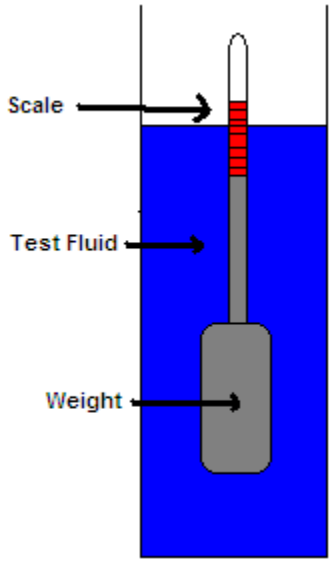
Density and specific gravity measurements require a binary mixture with at least one fluid phase to make accurate measurements. Given the temperature and pressure, the density of a gas can be found using the ideal gas law or some other relationship, and since density is a function of composition, the composition can be measured indirectly. See Thermodynamicsfor more information. Specific gravity is simply the ratio of density of water over the density of the non-compressible component at the same physical conditions, which means that it is simply a dimensionless measurement of density.

Four different measuring devices for finding density are described below.

Liquid Column
In a liquid column, the pressure measurements in a column are used to determine the density. The column may be open to the atmosphere, where the gauge pressure at the bottom of the fixed-height column is measured. Or, the column may be closed, in which case the differential pressure measurement is made between the bottom of the fixed-height column and at the top of the column immediately below the liquid surface.

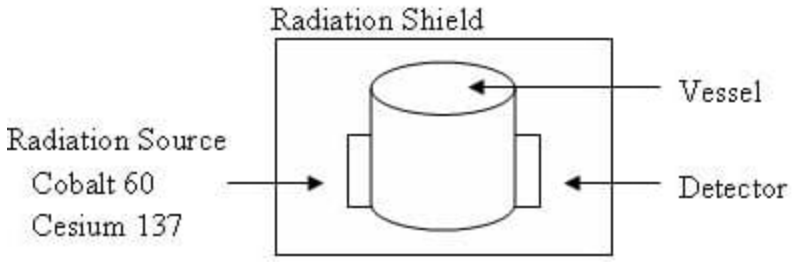
Displacement – Hydrometer
A hydrometer is usually made of glass and consists of a cylindrical stem and a bulb weighted with mercury or lead shot to make it float upright. The liquid to be tested is poured into a tall jar, and the hydrometer is gently lowered into the liquid until it floats freely. The point at which the surface of the liquid touches the stem of the hydrometer is noted. Hydrometers usually contain a paper scale inside the stem, so that the specific gravity can be read directly.

The operation of the hydrometer is based on the Archimedes principle that a solid suspended in a liquid will be buoyed up by a force equal to the weight of the liquid displaced. Thus, the lower the density of the substance, the lower the hydrometer will sink.



Direct Mass Measurement
With this technique, fluid is continuously passing through a U-shaped tube, where the tube is oscillating at its natural frequency. An electrochemical device strikes the tube periodically at the curved portion of the tube. The change in frequency between the electrochemical strike and the natural frequency is proportional to the changes in density which correlates to the composition of the substance.

Radiation-Density Gauges
Gamma radiation inside a pipe/vessel is used to determine the density of the liquid. Below is a diagram of the setup.



If the path-length for the radiation source is under 610 mm then cesium 137 is used. Above 610mm, cobalt 60 is used. The detector is usually an ionization gauge. The density is a function of the absorption of the radiation being detected. Decay of the vessel must be taken into consideration when calibrating

Viscosity

Please refer to the Viscosity Sensor section of the process control website for in-depth information.

Comprehensive Charts

- **For simple, easy to read charts summarizing the advantages and disadvantages of each type of sensor, please follow these links for a PDF version of each chart.**

Media:Photometric sensors.pdf

Media:Electroanalytical Measuring Techniques.pdf

Media:Other Types of Measuring Techniques.pdf

- **For complete summary charts including more sensors not discussed in the article, please follow the links here for a PDF version of these summary charts.**

Media:CompositionSensorsGeneral.pdf

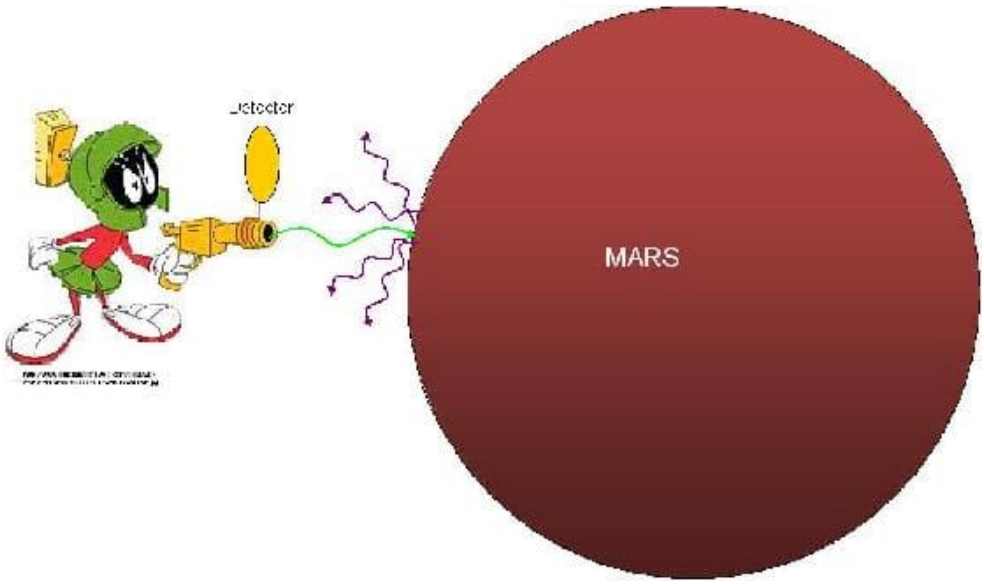
Media:CompositionSensorsPhysical.pdf

Media:CompositionSensorsChemical.pdf

Media:CompositionSensorsElectrical.pdf

Applied Example: Mars Pathfinder Spectrometer

The Mars Pathfinder uses an Alpha Proton X-Ray spectrometer to identify the chemicals that comprise the Martian surface. The incredible tool shoots X-Rays at the surface of the planet and uses a backscatter detector to collect the alpha particles, protons, and X-rays characteristic of specific elements. These atom specific scattering events allow NASA scientists to determine the surface composition.



Example 3.6.2

Your supervisor put you in charge of monitoring the emissions from an old smoke stack that the plant just re-opened to keep up with product demand. It is important to measure the amount of smoke and oxygen leaving the stack. Choose the proper types of sensors to install and keep in mind that the contents of the stack are exiting at very high temperatures.

Solution

To monitor the smoke density one could install an opacity monitor that measures the absoprtion and scattering of light. Please refer to *Sensor Assmebly* on this website Opacity Monitor Assembly for a basic schematic diagram.

To monitor the oxygen concentration one could install either a polarographic sensor or a high temperature ceramic sensor because both types of sensors are specific to O_2 particle. The high temperature ceramic sensor is more suitable for this application since the emissions are exiting at high temperatures.

Example 3.6.2

You are in charge of calibrating a new conductivity cell acquired by the company. The conductivity cell will be used to measure the concentration of NaCl in the beginning reagents. NaCl has a molar conductivity of 12.6 Sm^2 / kmol. The cell dimensions are as follows: area of the cells (A) is 100 m^2 and the distance between the cells (x) is 0.25 m. Calculate the concentration needed to produce a resistance of 10 ohms in the sensor.

Solution

Referring to the Conductivity Cell section, the equation can be arranged to find concentraton:

$$\text{concentration} = \frac{\text{conductivity}}{\text{.olar conductivity}}$$

Unknown = conductivity

$$\begin{aligned}\text{conductivity} &= \left(\frac{A}{x}\right)\left(\frac{1}{R}\right)\\ &= \left(\frac{100}{0.25}\right)\left(\frac{1}{10}\right)\\ &= 40\end{aligned}$$

Plug back into initial equation

concentration = conductivity / molar conductivity

concentration = 40/12.6

concentration = 3.17 kmol

References

- Richardson, JF. Coulson & Richardson's Chemical Engineering Vol. 3, Oxford: Butterworth-Heinemann. ISBN 0080410030
- Perry, GH; Green, DW. Perry's Chemical Engineering Handbook 7th Edition, New York: Mcgraw-Hill ISBN 0070498415
- So, Peter; Dong, CY. *Fluorescence Spectrophotometry*. Encyclopedia of Life Sciences, London: McMillan
- OMEGA Engineering. Turbidity Measurement.
- Analytical Specialities. Electrochemical Sensors
- Shimadzu. Quantification of Organic Peroxides Using FTIR and ATR.
- Merriam-Webster's Dictionary.
- Wikipedia, Mass Spectrometry
- Berg, Jeremy M.; Tymoczko, John L.; Stryer, Lubert. *Biochemistry*, Sixth Edition. W.H. Freeman and Company.
- Wilks, Paul. Process Monitoring: In-line, At-line, or Slip-stream?.
- Thayer, Ann M. KNOW THY PROCESS:Regulatory push for process analytics sets new goals for pharmaceutical manufacturing.
- Randall, Willis C. PAT Pending.
- ABB Instruments. Types of Industrial Composition Sensors available for purchase.
- Topac Inc. Conductivity Cells.
- Wroblewski, Wojciech. Ion Selective Electrodes.
- Chemical Engineering Department at Case Western Reserve. Electrochemistry Encyclopedia.
- Basic Turbidimeter Design and Concepts. www.epa.gov/ogwdw/mdbp/pdf/turbidity/chap_11.pdf

Contribuotors

Authors: (September 14,2006) Nicole Blan, Jessica Nunn, Pamela Anne Roxas, Cynthia Sequerah

Stewards: (September 27, 2007) Eric Gauthier, Kevin Lin, Matt Naughton, Rebecca Rutishauser

Anchors: (September 26, 2007) Megan Boekeloo, Stacy Young

This page titled 3.6: Composition Sensors is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

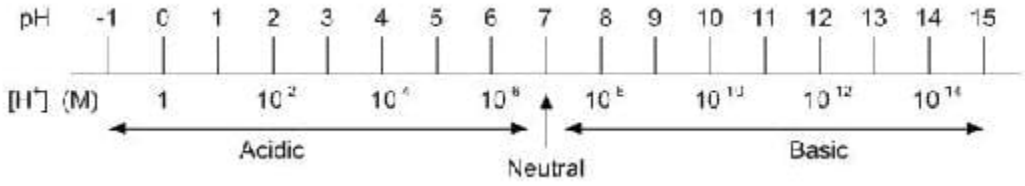
3.7: pH and Viscosity Sensors

Introduction to pH

pH is a measure of concentration of H^+ ions ($[H^+]$) in solution, and can be defined as follows:

$$pH = -\log\left([H^+]\right)$$

Typically, pH values range between 0 (strongly acidic) and 14 (strongly basic, or alkaline). However, that is not to say that 0 and 14 are the limits of pH values. A solution that has a H^+ concentration of 10 M, such as very strong HCl, would have a pH of -1. Pure water, which has a pH of 7, is said to be neutral because its H^+ concentration is equal to its OH^- concentration. An acid solution is said to have a greater H^+ concentration than its OH^- concentration. Conversely, a base, or alkaline solution, is said to have a greater OH^- concentration than its H^+ concentration. See for reference:



For weak acids and bases, the above equation is appropriate to use when calculating pH. However, many solutions are not weak acids and bases, which means that the other ions in solution play a large role in acidity and alkalinity, and thus a more accurate method is needed to calculate the pH. pH sensors are used to determine the acidity or alkalinity of a particular process solution. When dealing with a strong solutions, the sensors measure the effective concentration, or activity, of H^+ ions in solution, and do not measure the actual concentration. Hence, for the purposes of pH sensors, pH is properly defined as:

$$pH = -\log\left(a_{H^+}\right)$$

with

$$a_{H^+} = \left(\gamma_{H^+}\right)\left[H^+\right]$$

where γ_{H^+} is the activity coefficient.

The activity coefficient accounts for the interactions between all the ions present. These interactions may prevent some of the H^+ ions from contributing to the acidity of the solution, and thus decrease the effective concentration of H^+ ions. This effective concentration is the concentration used when calculating pH.

Why is pH Relevant for Chemical Engineering?

pH sensors, while being relatively simple pieces of equipment, have numerous uses. Knowing the pH of a solution is valuable to an analytical chemist attempting to determine the contents of an unknown solution, as well as a farmer trying to determine the appropriate applications to his fields and many people in between. In industry pH sensors can be used as a simple way to monitor reactions involving acids or bases as either a reactant or product of a chemical process. They can also be used to monitor the conditions of a reactor to ensure optimal conditions. This is especially important in biological systems where very minor changes in pH can result in significantly lowered production levels or even a complete stoppage of production due to the death of the organisms on which the reaction is dependent.

An example of pH measurement's importance in industry is seen in food processing during the manufacture of fruit jelly. The gel formation brought about by pectin occurs over a small pH range, which is further complicated by the concentration of sugar. Too high of a pH value results in an unacceptably runny liquid. Too low of a pH causes the mixture to stiffen prematurely, resulting in

a nearly-solid product. Continuous pH adjustment using edible acids such as citric acid yields the optimal gel consistency.

It is also important to keep in mind that application requirements should be carefully considered a pH sensor is chosen. In the case above, it is important that the pH sensor be made of a material that will not interfere with the quality of the product. For instance, the sensor should not be made of glass in order to avoid broken glass entering the product. Accurate pH measurement and the precise control that it allows for helps to optimize a process and can result in increased product quality and consistency. Continuous pH monitoring also controls chemical usage (such as the edible acids above), resulting in minimized system maintenance (i.e. less cleaning of hardened gel from containers, less corrosion of containers, etc.).

Typical pH Sensor Construction

Because it is difficult to directly measure the H+ concentration of a solution, pH is typically measured by comparing the potential of a solution with a known [H^+] (this is known as a reference solution) with the potential of a solution with an unknown [H^+]. The potentials of these solutions are measured using a reference half-cell and a sensing half-cell, respectively.

Reference Half-Cell

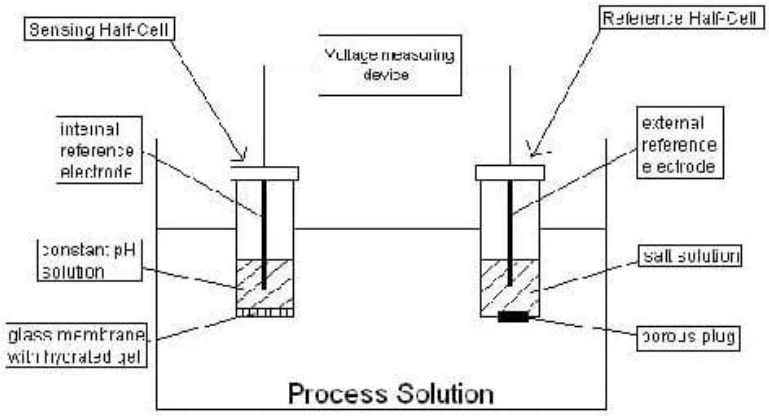
The reference half-cell generally consists of a chamber with a conductor submerged in a reference electrolyte. The conductor, or external reference electrode wire, is typically silver coated with silver chloride (Ag/AgCl(s)) or mercury coated with mercurous chloride (Hg/Hg2Cl2(s)). The reference electrolyte is a standard solution such as KCl. The last component of the reference half-cell is a porous plug, which serves as a liquid-liquid interface between the standard solution and the process solution being analyzed. This plug allows standard solution to travel from the chamber out into the process solution, but does not allow process solution into the chamber. The purpose of this interface is to establish the electrical connection, which provides the reference potential.

Sensing Half-Cell

The sensing half-cell is of similar construction to the reference half-cell. A glass chamber contains an electrode (Ag/AgCl(s)), known as the internal reference electrode, which is submerged in a standard solution of constant pH. However, instead of a porous plug acting as the liquid-liquid interface between the standard and process solutions, there exists a glass membrane. This membrane is coated on the inside and outside with a hydrated gel. When the sensing half-cell is placed in solution, metal ions (Na+) in the gel on the outside of the membrane diffuse out into the sample being analyzed and are replaced by H+ ions from the sample. The change in energy associated with this substitution creates the change in electric potential that is detected by the electrode. The combination of the hydrated gel and the glass membrane are what make the sensor specific only to H+ ions. The solution of constant pH ensures that the surface of the glass membrane and the electrode are at the same pH before the cell is exposed to the process solution being analyzed.

The electrode wires from the reference and sensing half-cells are each connected to the voltage measuring device.

The diagram below shows the setup of a typical pH measurement system:



While this figure shows the essential components of a pH sensor, it is not an accurate depiction of what pH sensors actually look like. The reference and sensing cells in most pH sensors are generally encased together in a sturdy housing (usually made of some sort of plastic). This does not affect the operation of the sensor, but allows for ease of use, as there is only one major piece of equipment that the user needs to carry around.

pH sensors can be used both online and offline. If the sensor is being used to carefully control the pH of a system or to provide an accurate record of the pH of a solution it is often desirable to have an online pH sensor. In order to have an online pH sensor determine accurate readings it must be in a position to take a representative sample. This is most often done by diverting a side stream from the main process in which the pH sensor can be inserted. Often times continuous pH measurements are not necessary, but rather infrequent measurements to make sure that the system is running correctly are of interest. In these cases it is much more efficient to use an offline pH sensor.

How do pH sensors calculate pH?

pH is calculated using the potential drop between the reference and sensing electrodes. The potential across the overall pH measurement system can be expressed as:

$$E = E_{\text{internalreference}} + E_{\text{membrane}} + E_{\text{porousplug}} - E_{\text{externalreference}}$$

The potential of the porous junction is usually negligible, and the potentials of the internal and external references can be lumped into a constant value, E_0. Thus, we get:

$$E = E_0 + E_{\text{membrane}}$$

Using the Nernst equation, the potential of the membrane in a half-cell reaction can be expressed as:

$$E = E_0 + \frac{RT}{nF}\ln(a)$$

where

- R is the Universal Gas Constant
- $F = 9.6490 \times 10^7$ coulombs/kmol\) is Faraday's Constant
- n is the net number of negative charges transferred in the reaction
- a is the activity of the ions involved in the reaction

This equation is the standard equation used in electrochemistry to find the potential with respect to a reference electrode. It can be viewed as a calculation of the change in free energy at an electrode, due to changing ion concentrations at that electrode. Assuming that the electrode is completely selective to H^+ ions ($a = 1$) and that there is only one positive ion carrying one

elementary charge in each half-cell reaction ($n = 1$) the pH of the unknown solution can be expressed as:

$$pH = \frac{F\left(E_0 - E\right)}{2.303RT}$$

where T is the absolute temperature of solution being analyzed.

Temperature Compensation

As one can see from the Nernst equation, the pH as determined using the potential difference between the reference and sensing half-cells varies linearly with changes in temperature. That is, the voltage drop/pH unit across the cells increases with increasing temperature. Thus, pH measuring system needs to have some way to compensate for temperature changes in the process solution. Many pH control kits come equipped with automatic temperature compensation (ATC), which continuously measures and incorporates temperature changes into their pH calculations. Other older pH control kits require manual temperature compensation, meaning that the operator or user must manually input the process solution temperature into the control system.

Alternative Methods for Determining pH

While glass membrane pH sensors are an effective method of measuring pH they are not always the best choice. Certain properties of the system of interest make using a glass membrane pH sensor impractical. Glass membrane pH sensors can become clogged by viscous solutions or solutions containing suspended solids. Harsh conditions can also scratch or break the glass membrane. There are also cases where a glass membrane pH sensor would work but using one would be an unnecessary cost. In these situations there are several alternatives that can be used to determine pH. One method is using pH paper. This paper changes color when submerged in a solution. The color of the paper can be compared to a standard scale included with the paper to determine an approximate pH. pH paper is an excellent choice when all that is needed is a "ballpark" pH value. Another method of determining pH is the use of indicators. These are substances that change color depending on the pH of the solution. They can be very useful if the primary interest is determining if the pH of the system is above or below a particular point. Numerous indicators exist that change colors at different pH values.

Other Ion-Specific Sensors

The same concepts that allow pH to be measured using a reference and measuring electrode can also be applied to measure the concentration of other ions. The differences between the pH sensors and ion specific sensors are the composition of the measuring electrode the membrane used in the measuring electrode. For some ions of interest glass membranes can still be used, however for other ions solid state membranes or liquid ion exchange may be necessary. The choice of membrane depends on both the size and the charge of the ion of interest.

While ion specific electrodes can be fairly accurate a number of problems can compromise their accuracy. Ion specific electrodes can suffer from interference. This can be especially problematic in complex solutions, especially those containing many other ions similar in charge and size to the ion of interest. The measurements reported by ion specific sensors are also functions of the ionic strength of the solution. This can also result in significantly decreased accuracy of the sensor if the effect of the ionic strength of the solution is not accounted for. As with pH the property that is actually being measured with an ion specific sensor is the activity of the ion, not its concentration. The activity can no longer be considered negligable in solutions with high ionic strength. In order to obtain proper readings the user must consider what effect the ionic strength of the solution has on the activity coefficient and adjust the measurement accordingly. Despite these limitations ion specific sensors can still provide accurate and valuable information, provided an appropriate measuring electrode membrane is used and the ionic strength of the solution is accounted for, and are widely used.

Problems with pH Sensors

Several problems can occur with pH sensors that can negatively affect the accuracy of these devices. The majority of these problems involve the reference electrode. One such problem is a partial clogging of the porous plug. Another error associated with the plug is due to a junction potential which can vary depending on the electrode setup. This junction potential adds error to the pH measurement when the ionic composition of the solution being tested is significantly different than the ionic composition of the standard buffer used to calibrate the electrode. In general, a junction potential develops at the interface of two solutions with different ionic compositions because the various ions in each solution have different mobilities, and as oppositely charged ions

diffuse at different rates a charge separation develops at the interface between the solutions. A different type of error can occur in measurements of solutions with very low hydrogen ion concentration and very high sodium ion concentration; pH electrodes can "mistake" the sodium for hydrogen, leading to a lower pH reading than the true value.

Other problems include the contamination of the electrolyte in the reference electrode of a complete depletion of the electrolyte. Any of these problems will result in an incorrect potential in the reference electrode and ultimately inaccurate pH measurements.

Problems can also occur with the measuring electrode. Any damage to the glass membrane including cracks, etching, or abrasion, as well as coating of the glass membrane will result in poor pH measurements. In addition, if the electrode is dry it will not measure pH correctly; it must be soaked in some aqueous solution for a few hours. Problems with pH sensors can be fairly easily detected. Any problems such as difficulty calibrating the sensor, unsteady pH readings, or significant drift in the readings usually indicate damage to the pH sensor. It is now possible to test some forms of pH sensors online to quickly and easily determine which electrode is the cause of the problems and potentially the source of these problems. In some cases these problems can be fixed by a careful cleaning of the sensor, however if problems persist it may be necessary to replace the pH sensor.

Example 3.7.1: pH Sensor Application

Would it be wise to use a pH electrode in the following situations? If not what method would you use to determine the pH of the system of interest?

a. You are an engineer responsible for a wastewater treatment facility. It is important that you know the pH of the water at several points in the treatment process to detect any potential problems.
b. You work for a pharmaceutical company who uses bacteria to produce a desired protein. The bacteria are very sensitive to changes in pH and production falls off significantly if the pH varies from optimal conditions.
c. You work in a plant that uses hydrochloric acid as a catalyst. After the reaction has gone to completion and the product is extracted it is desirable to separate the acid from the remaining solution to be reused in the plant and allow for easy disposal of the remaining solution. The EPA requires accurate data on the pH (among other things) of all discharged solution.
d. You work in a plant that produces acetic acid. You want to ensure that the solution exiting the reactor contains at least 75 mol% acetic acid.

Solution

a. NO. A pH electrode would not hold up well under the harsh conditions of a wastewater treatment plant. It is also not necessary in this case to have constant measurements or exact pH values as the pH is being used simply as an early problem detection system. Simple pH paper would be sufficient in this situation.
b. YES. The sensitivity of the bacteria, as well as the likely high value of the proteins being produced calls for the accuracy and frequent measurements afforded the engineer by a glass membrane pH sensor.
c. YES. Here not only constant measurements are required, but records of these measurements must be kept to prove that environmental standards are being met. The most cost effective way to measure and record this data would be with a pH sensor connected to a computer.
d. NO. While pH could be used as an effective way to monitor the progress of the reaction a sensor is not necessary. Here you are only interested in ensuring that a certain benchmark, in this case a reaction completion percentage measured using pH, is reached. Choosing an appropriate dye indicator would be a much simpler and more cost effective way to make sure this standard is met.

Introduction to Viscosity

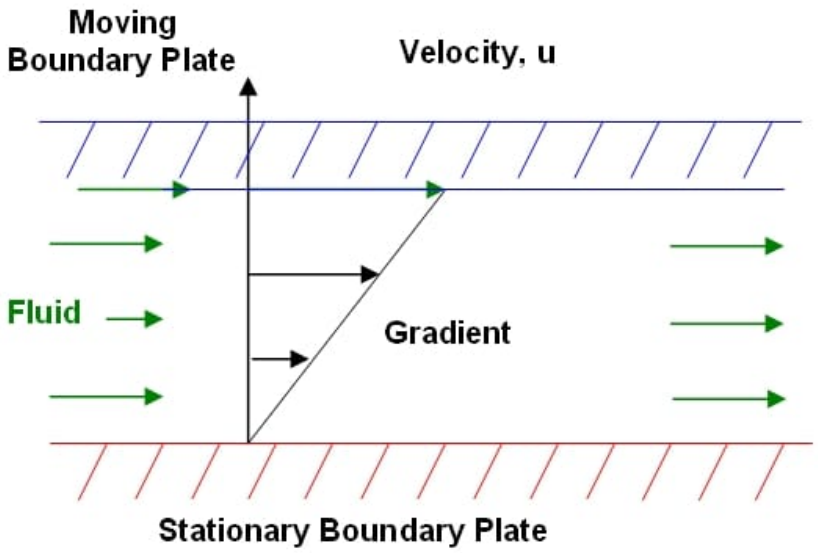
Viscosity quantifies the systematic deformation of the surface of a fluid upon an applied shear stress. Fluid, while in motion, typically travels at varying velocities due to the geometry of contact surfaces. It may be characterized either as a Newtonian fluid or a Non-Newtonian fluid. A Newtonian fluid flows continuously in a uniform manner irrespective of the forces acting upon it (a common example, water). Conversely, the viscosity of a non-Newtonian fluid changes upon differing flow rates (common examples, table ketchup or mud).

Also, the effect of temperature on fluid viscosity is imperative. From a thermodynamic perspectice, temperature measures the random movement of molecules of a given substance, and as movement increases internal energy increases, and temperature

also increases. For most traditional substances, having higher internal energy/temperature also implies a lower viscosity, as the substance will be less resistant to movement (i.e. will deform sooner to an applied shear stress). For example, consider honey: at room temperature it oozes out of the squeeze bottle, but after a minute in the microwave it may be easily poured out, similar to a syrup.

Nearly all modern viscometers have built in controls to address temperature issues, and a competent operator or engineer may input temperature settings to that the appropriate calculations are carried out correctly.

Viscosity may be further subdivided into two distinct forms: dynamic viscosity and kinematic viscosity. **Dynamic viscosity** (μ), a figure representing shear stress as proportional to the strain rate, has the SI unit Pa*s. **Kinematic viscosity** ($v = \mu/\rho$) describes shifts in momentum and has SI units m^2s^{-1}, but is also commonly represented by the Stoke, cm^2s^{-1} (content adapted from http://scienceworld.wolfram.com/physics/DynamicViscosity.html).



As noticed above, the typical Newtonian fluid experiences friction upon contacting a surface and resultantly develops a characteristic velocity profile. This profile may be described by examining the known properties of the fluid and the surrounding structure through which it travels, commonly piping for plant chemical engineering.

In the above diagram we notice a non- cylindrical surface featuring two boundary plates, one stationary to the observer, and one mobile. Fluid contained within the two boundaries may behave according to the constraining conditions of the boundary plates, and subsequently characterized by fluid mechanics:

$$\tau = \mu \frac{\partial u}{\partial y}$$

Assuming μ as the coefficient of viscosity, the above equation describes the relationship between the shear (parallel) stress and velocity of fluid travel per unit height above the stationary boundary plate. Fluids having a relatively high viscosity resultantly also require greater force to induce motion—resulting in a larger shear stress.

Why is Viscosity Relevant for Chemical Engineering?

For engineers a thorough knowledge of the physical and chemical properties of products is essential to the successful implementation of any design. Viscosity, similar to volatility, density, or any other physical property, offers significant insight into the potential behavior of a substance, whether classified formally as a solid, liquid or gas. By taking viscosity into account, engineers may correctly select and place instrumentation in an environmentally friendly, sustainable method.

Many professional assignments necessitate a genuine understanding of viscosity and its effects on process engineering. For a more detailed and insightful glance into the daily upkeep and activities associated with plant maintenance, please consult Michael Pearce at www.reynoldsmixers.com/whitepapers/030910_hotmelt.doc. He speaks at length regarding the cost- effective

selection of specific types of blades used in mixing processes.

As an example, a key parameter in food process monitoring is viscosity. The viscosity is directly related to the flow characteristics of the product, which impact pumpability, pourability, and spreadability. In the food-processing environment there are a number of challenges to viscosity measurement, such as harsh process conditions and the complex rheological properties (relationships between deformations and stresses of materials). Offline measurements are often cumbersome, labor-intensive, and prone to operator error. Online viscometers must be able to deliver continuous measurement day after day with minimal maintenance. Therefore, considerations when choosing viscosity sensors for food processing should include ease of cleaning, minimal risk of fouling, and whether or not they meet sanitary requirements.

Off-line Instruments

For the measurement of viscosity in controlled settings (typically laboratory) where the majority of variables are maintained as constant, engineers and scientists use off-line viscometers. Typical examples include: Capillary, Couette, Falling Ball, Cone and Plate and Oscillating Cylinder. It is very important to keep these instruments in a regulated environment with a stable temperature (such as in a water bath) because of viscosity's sensitivity to changes in temperature.

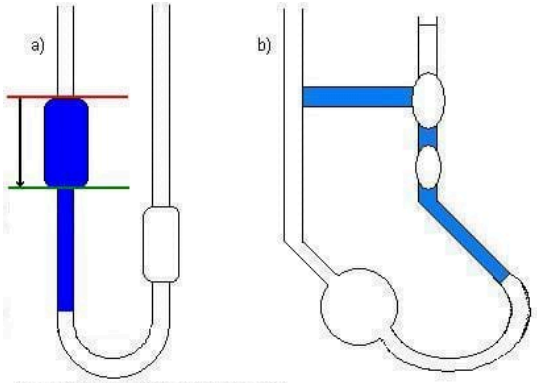
Capillary

The Ostwald U- Tube viscometer functions by measuing the amount of time a specified quantity of fluid takes to move from its initial position to a final position.

Typically, a suction device holds the bottom of the fluid meniscus at the start position (indicated by the solid horizontal red line), and upon its release the fluid drops an approximate distance of 5 mm (to reach the solid horizontal green line). The time for this drop is recorded so that the acquired data may then be used in Poiseuille's Law to determine the outcome of kinematic viscosity.

Both the dynamic and kinematic viscosities may be found by appling these formulae (where K_0 is the non-dimensionless viscometer constant):

- Dynamic Viscosity: $\mu = K_0 \rho t$
- Kinematic Viscosity: $v = K_0 t$



Examples of Capillary Viscometers:

a) Simple Ostwald Type
b) Cannon- Fenske Type

Content for Diagram Adapted from:

Coulson and Richardson's Chemical Enginering, Volume 3, Third Edition, Chemical, Biochemical Reactors, and Process Control; contributing editors: J.F. Richardson & D.G. Peacock,

Couette

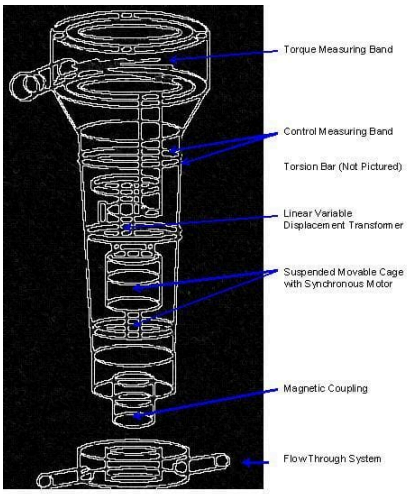
This device may be used in both off-line and on-line applications with appropriate modifications in the output settings. A couette type viscometer measures viscosity by spinning a cylinder encapsulated in fluid. This is accomplished through the synchronization

(identical frequency, no phase difference) of a motor with magnetic coupling to rotate a magnet which in turns forces the inner cylinder to revolve within the fluid.

The torque reaction of the motor is resultant of the viscous drag on the rotating cylinder. This torque on the motor is effectively counteracted by the torsion bar (a thin rod connecting the control/ measuring bandto the linear variable displacement transformer). The deflection of the torsion bar, a function of the fluid viscosity, is then subsequently converted into local signal available for laboratory analysis.

Engineers use viscometers of the couette type in in-line or in-tank applications. Also, the meters are appropriate for both Newtonian and non-Newtonian fluids. The acceptable range of viscosity spans from 10^{-3} to $5 * 10^{3} Ns / m^{2}$.

Information for this device has been adapted from Richardson's Chemical Engineering.



Couette- Type Viscometer for a Continuous Process

Content for Diagram Adapted from:

Coulson and Richardson's Chemical Engineering, Volume 3, Third Edition, Chemical, Biochemical Reactors, and Process Control; contributing editors: J.F. Richardson & D.G. Peacock,

Falling Ball

A substantially more direct approach to viscosity measurement, testers allow a ball to fall through a column containing liquid of unknown viscosity and then measure the amount of time necessary for the ball to reach the bottom of the column.

A commonly used method involves a stainless steel ball which sinks through the liquid under scrutiny; researchers measure the time necessary for the ball to drop from one preselected level to another. To attain the terminal velocity of the ball in the liquid, u_0, timers write the initial timing mark a minimum of six ball diameters below the ball release point.

Also important to note, the ratio of ball diameter to column tube diameter should not exceed 1:10 to account for the drag effects of the column tube walls on the terminal velocity of the falling steel ball. A simple viscosity derivation from buoyancy principles is presented in Richardson's text, and is shown again here:

$$v = \frac{K_v \left(\rho_s - \rho\right)}{\rho} t$$

where

- K_v is the dimensional viscometer constant
- ρ_s is the density of the ball
- ρ is the density of the liquid
- t is the time required for the ball to sink the measured distance

Cone and Plate

Another popular method used employs a cone placed in a manner so that the tip touches the center of a stationary plate. Fluid encapsulates the cone and researchers take a measurement of the amount of torque required to keep the plate stationary.

This process allows the dynamic viscosity to be quantitatively described as:

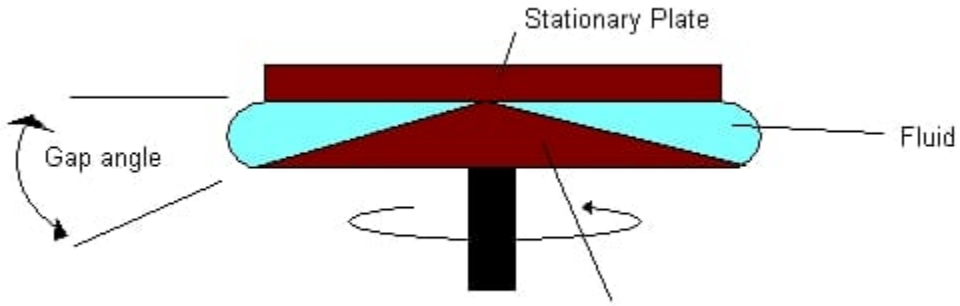
$$v = \frac{3}{\pi \gamma r^3} T$$

- T = Torque required to keep the top plate in place
- γ = Rate of shear

The rate of shear is the ratio of the angular velocity of the cone to angle between the cone and the plate (note: this angle must be small enough so that the sin of the angle approximately equals the angle itself).

Important Additional Notes:

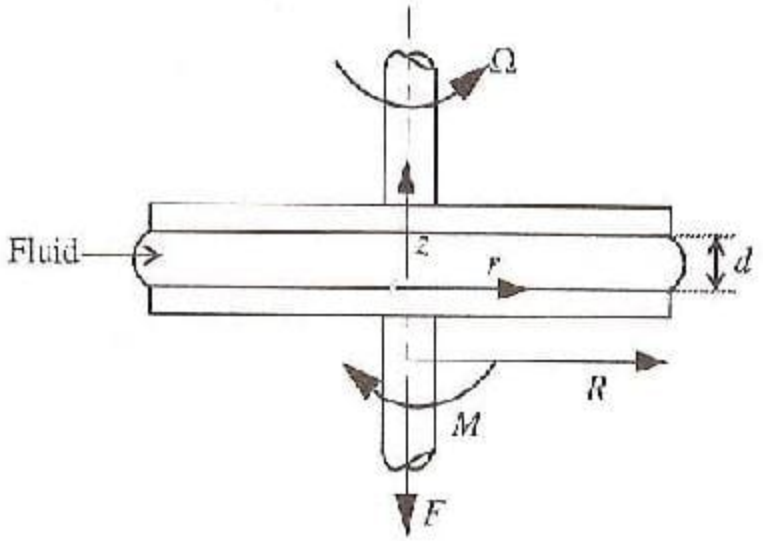
Edge effects of the cone or plate are neglected. The equation may also be used for Non-Newtonian fluids provided that γ is relatively constant.



Example of a Cone and Plate Viscometer

Parallel Plates

Parallel plates or discs viscometer is similar to the cone and plate method. The cone in the case is replaced with another plate. In this viscometer, the fluid flows in a gap between two parallel discs. One of the discs rotates with an angular velocity ω which creates the shear. Torque is applied to the other plate so it stays stationary. A normal force, F, is created by fluid elasticity and acts as to separate the two plates. No slip at boundaries is assumed.



Parallel Plates viscometer

ω, T and F are all experimental parameters. Shear rate is given by:

$$\dot{\gamma} = \frac{\omega r}{d}$$

- $\dot{\gamma} =$ Shear rate
- r = The radius of the disc
- d = The distance between the two plates

Shear stress, τ, is given by :

$$\tau = \frac{T}{2\pi r^3}\left(3 + \frac{dlnT}{dln\dot{\gamma}}\right)$$

Viscosity is, therefore, the ratio of shear stress to shear rate. Some limitations associated with parallel plates include:

1. Sufficient data of T vs w and F vs w must be available
2. Uniform temperature at different points in the gap between the two plates is required
3. Error might be caused from edge fracture, wall slip, misalignment of the plates and viscous heating.

Oscillating Cylinder

This instrument involves an arrangement requiring a rotating cylinder placed in a quantity of a viscous fluid. The level of disturbance noticed in the fluid resulting from the rotation of the cylinder is then noted, and may be used if the following equation to determine viscosity:

$$v = \frac{(r_2^2 - r_1^2)}{4\pi\omega_v r_2^2 r_1^2} T'$$

- r_2 = Outer cylinder radius
- r_1 = Inner cylinder radius
- ω_v = Angular velocity

The outer cylinder of radius r_2 rotates with an angular velocity of ω_v while the inner cylinder of r_1 remains stationary thus allowing for the torque (alternatively reffered to as 'viscous drag'), T', to be measured at predetermined values of ω_v.

On-line Instruments

On-line instruments are those that are capable of giving a precise measurement of viscosity under plant conditions. This means that they will be able to withstand the wide range of viscosities and the variable temperature, pressure, and flow rate that occur within a process plant while still obtaining the accuracy that is required. In all on-line instruments automatic temperature compensation is always required because of viscosity's temperature dependence.

Capillary

This type of viscometer is derived from Poiseuille's law:

$$\frac{dv}{dt} = v_s * \pi * r^2$$

- v_s = Fluid velocity
- r = Internal radius

The instrument has many temperature and pressure sensors in the intricate system because there must be a precisely controlled pressure drop and temperature in the stream. Response time is minimized by adding a fast loop sampling system. A capillary viscometer would be useful when using Newtonian fluids such as lubricating oils or fuel oils. It can endure viscosities ranging from $2 * 10^{-3}$ to 4 Ns / m^2.

Couette

Identical to the description provided earlier, except now with a change in the output data processing. The deflection of the torsion bar, a function of the fluid viscosity, is now converted into a 4-20 mA output signal and is readily available for additional plant data mangement/processing.

Information for this device has been adapted from Richardson's Chemical Engineering.

Vibrating Rod or Cylinder

This type of viscometer can also be used in-line or in-tank. It is suitable for Newtonian, non-Newtonian fluids, and mineral slurries. The vibrating rod or cylinder works by using an electromechanical system as the driving force, or by a pulse of current through a specific alloy. The lowest viscosity that it can handle is $10^{-4} Ns / m^2$ and it can extend to $2 * 10^3 Ns / m^2$. Normally on lower viscosity ranges the cylinder is used, while with higher ranges the rod is used. Given a plants' natural vibration care must be taken when deciding on the location of an oscillation viscometer because of their sensitivity.

Table of Viscometer Uses

This table found in Richardson's text offers some guidance on the device selection process.

Common Viscometer Ranges (Content from **Richardson's Chemical Engineering**)

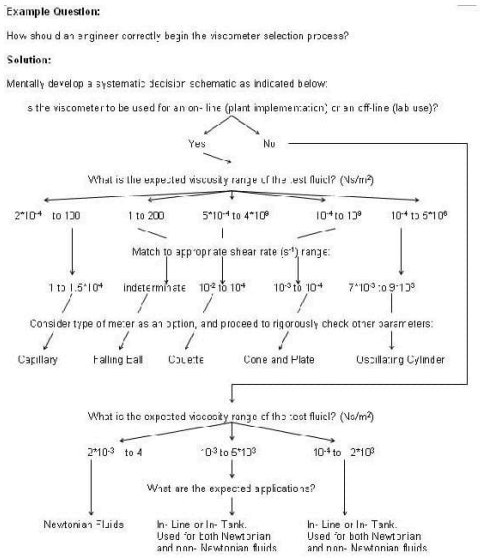
Viscometer Type	Lowest Viscosity (Ns/m^2)	Highest Viscosity (Ns/m^2)	Shear Rate Range (s^-1)
Offline (Laboratory Instruments)			
Capillary	2*10^-4	100	1- 1.5*10^4
Falling Ball	1	200	Indetermin nate
Couette	5*10^-4	4*10^6	10^-2 - 10^4
Cone and Plate	10^-4	10^7	10^-2- 10^4
Oscillating Cylinder	10^-4	5*10^6	7*10^-2- 8*10^3
Online Instruments			**Applications**
Capillary	2*10^-3	4	Newtonian fluids
Couette	10^-3	5*10^3	In-line or in-tank*
Vibrating Rod or Cylinder	10^-1	2*10^3	In-line or in-tank*#

*Also Newtonian and Non Newtonian Fluids
#May be used also for mineral slurries

One may notice that the viscosity ranges of some of the viscometers overlap. In a situation where this occurs and it is necessary to choose between multiple viscometers, one needs to take into account process conditions such as changes in temperature, pressure, and solution flow rate. The appropriate viscometer should be able to handle these conditions without disrupting the process.

Problem Solving Schematic

The following schematic may be employed to select an appropriate viscometer. Additional considerations must also take place, as the usable viscometer ranges occasionally overlap. This requires the competent operator or engineer to base his or her selection from a host of other valuable criteria, including but not limited to: cost- effectiveness, difficulty of implementation, heat generation, availability of qualified local maintenance personnel, environmental impact, and long term sustainability.



Example 3.7.1: Viscosity Sensor

You are an engineer running a process that requires hot water at some times and cold water at others. To heat and cool your water, you run it through a heat exchanger. When you want to heat the water, you use 100°F glycol stream to heat the water (viscosity = 0.002 Ns/m^2). When you want to cool the water, you use 20°F glycol stream to cool the water (viscosity = 5 Ns/m^2). To ensure that the glycol does not clog the pumps in your process, you need to monitor its viscosity on-line. What type(s) of viscometers could you use on the hot glycol stream? On the cool glycol stream?

Solution

After looking at the summary table given above, we can immediately narrow down our search to just the on-line instruments from the information given in the problem statement. For the hot glycol stream we find which instruments have a viscosity range which allows for a viscosity of 0.002 Ns/m^2. We find that the following viscometers would be acceptable: Capillary, vibrating rod or cylinder

The cool glycol stream involves the same process, but it must have a different range of viscosity which includes 5 Ns/m^2. Viscometers which are acceptable for this stream include: Couette, vibrating rod or cylinder

pH/Viscosity Sensor Key Learnings

As is apparent after reading this article, pH and viscosity measurements do not have much in common with each other. However, there is a common thread between these two types of sensors that needs to be acknowledged, and that is the method that needs

to be taken when deciding what type of sensor to use. While there is only one single type of pH sensor available, before it is to be used the user needs to evaluate whether there is an easier and cheaper way (eg. using pH paper or a pH activated dye) to accomplish the task at hand. With viscometers, there are multiple types of viscometers to use over various ranges of viscosity. The user needs to evaluate process conditions such as temperature, pressure, and flow rate, in order to decide which type of viscometer is optimal for the application.

The math associated with both types of sensors is good to know for background purposes and to gain a fundamental understanding of how a sensor produces a measurement, however it is not information that greatly helps the user decided what type of pH indicator or viscometer to use. Hence, this article has tried to give the background theory behind the measurements, but a thorough analysis of this theory is not included and not necessary for the purposes of Process Control.

References

- Cole-Palmer Catalog, 2005/2006
- Liptak, Bela. *Instrument Engineers Handbook - Process Control and Optimization, Fourth Edition, Volume 2*. Taylor & Francis, New York, 2005.
- Perry, Richard. *Perry's Chemical Engineers' Handbook, Seventh Edition*. McGraw-Hill, 1997.
- Richardson, J.F. and Peacock, D.G. *Chemical Engineering, 3rd Edition, Volume 3*. Butterworth-Heinemann, 1994.
- Rodriguez, Ferdinand; et al. *Principles of Polymer Systems, 5th Edition*. Taylor & Francis, 2003. ISBN 1560329394
- Stephanopoulos, George. *Chemical Process Control - An Introduction to Theory and Practice*. Prentice Hall, New Jersey.
- Hydramotion. Online Viscosity Control Increases Process Productivity and Efficiency. ScientistLive 17 November 2008. 8 December 2008 http://www.scientistlive.com/European-Food-Scientist/Processing_Equipment/Online_viscosity_control_increases_process_productivity_and_efficiency/21292/

Contributors and Attributions

- Authors: Navin Raj Bora, Dallas Burkholder, Nina Mohan, Sarah Tschirhart

This page titled 3.7: pH and Viscosity Sensors is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

3.8: Miscellaneous Sensors

This page is a summary of different sensors not as commonly used as others in industry. This page details the methods of operation, industrial uses, and other relevant information about the sensors.

Humidity Sensors

“Humidity sensors are used to control the amount of water vapor present in many industrial processes. Textile, wood, and chemical processing is very sensitive to humidity” [1]

Psychrometers

Psychrometers use latent heat of vaporization to determine relative humidity in the system. This can be done by using a dry bulb thermometer with a wet bulb thermometer. The two temperatures recorded can be used with a psychrometric chart to obtain the relative humidity, water vapor pressure, heat content, and weight of water vapor in the air.

Hygrometers

Hygrometers are devices that sense the change in either their physical or electrical properties. Some materials such as hair or thin strips of wood change length depending on water content. The change in length is directly related to the humidity.

Dew Point Measuring Devices

These devices measure humidity by cooling the air until water starts to condense on the object. The amount that the air needed to be cooled before water started to form on the object can be used to determine the relative humidity.

Other Humidity Sensors

Microwave absorption by water vapor can be used to measure the humidity in a material

Infrared absorption can be used by hitting the object with infrared radiation and measuring the energy of the reflectance.

Sound Sensors

Sound sensors are important because they can be used in industrial applications such as detection of flaws in solids and location and linear distance measurements. Sound pressure waves can also cause vibration and failure.

Microphones

Microphones are pressure transducers, and are used to convert sound pressures into electrical signals. There are six types of microphones: electromagnetic, capacitance, ribbon, crystal, carbon, and piezoelectric.

Smoke Sensors

Smoke sensors are useful for not only safety for workers, but also environmental concerns and purity issues of processes.

Infrared Sensors

These sensors detect changes a signal received from an LED due to smoke or other objects in the light path from the LED.

Ionization Chambers

These devices can detect the difference in current between two plates that have a voltage between them. The difference is due to carbon particles from smoke that provide a conductive path between the two plates.

Taguchi-Type

Taguchi-type sensors are used to detect hydrocarbon gases, such as carbon monoxide and carbon dioxide. The sensor is coated with an oxidized element that when combined with a hydrocarbon creates a change in the electrical resistance of the sensor.

References

- [1] Dunn, William C. Introduction to Instrumentation, Sensors and Process Control. Boston: Artech House, 2006. Print.

Contributors

Authors: Michael Kennett, Conor O'Bryan, Ryan Weigold

This page titled 3.8: Miscellaneous Sensors is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.9: Valves - Types, Kinds, and Selection

Authors: Daniel Katzman, Jessica Moreno, Jason Noelanders, and Mark Winston-Galant

Stewards: Jeff Byrd, Khek Ping Chia, John Cruz, Natalie Duchene, Samantha Lyu

Introduction

Control valves are imperative elements in any system where fluid flow must be monitored and manipulated. Selection of the proper valve involves a thorough knowledge of the process for which it will be used. Involved in selecting the proper valve is not only which type of valve to use, but the material of which it is made and the size it must be to perform its designated task.

The basic valve is used to permit or restrain the flow of fluid and/or adjust the pressure in a system. A complete control valve is made of the valve itself, an actuator, and, if necessary, a valve control device. The actuator is what provides the required force to cause the closing part of the valve to move. Valve control devices keep the valves in the proper operating conditions; they can ensure appropriate position, interpret signals, and manipulate responses.

When implementing a valve into a process, one must consider the possible adverse occurrences in the system. This can include noise due to the movement of the valve, which can ultimately produce shock waves and damage the construction of the system. Cavitation and flashing, which involve the rapid expansion and collapse of vapor bubbles inside the pipe, can also damage the system and may corrode the valve material and reduce the fluid flow.

There are four general types of valves.

1. Electronic, or electrical valves. The movement of the ball or flap that controls flow is controlled electronically through circuits or digitally. These types of valves have very precise control but can also be very expensive.
2. Non-Return valves. These valves allow flow in only one direction, and are common in various industries. When pressure in the opposite direction is applied, the valve closes.
3. Electromechanical valves. These valves have electro magnets controlling whether the valve is open or closed. These valves can only be fully open or fully closed.
4. Mechanical Valves. These valves use mechanical energy in the process of opening and closing the actual valve. Larger valves can be opened and closed using mechanical processes such as levers and pulleys, whereas smaller mechanical valves can be opened or closed via a turning wheel or pulling a level by hand.

There are four major valve types: ball valves, butterfly valves, globe valves, and plug valves. There is also an array of many other types of valves specific to certain processes. Selecting which type of valve to use depends on what task the valve must carry out and in what space said valve can fit to carry out the task.

Some general features that one can take into consideration when selecting a valve are the following:

1. Pressure rating
2. Size and flow capacity
3. Desired flow condition
4. Temperature limits
5. Shutoff response to leakage
6. Equipments and pipes connected
7. Material compatibility and durability
8. Cost

Valve Types

There is a vast abundance of valve types available for implementation into systems. The valves most commonly used in processes are ball valves, butterfly valves, globe valves, and plug valves. A summary of these four valve types and their relevant applications is in the table below.

Valve Type	Application	Other information
Ball	Flow is on or off	Easy to clean
Butterfly	Good flow control at high capacities	Economical
Globe	Good flow control	Difficult to clean
Plug	Extreme on/off situations	More rugged, costly than ball valve

Following is a detailed description of the four main valve types.

Ball Valves

A ball valve is a valve with a spherical disc, the part of the valve which controls the flow through it. The sphere has a hole, or port, through the middle so that when the port is in line with both ends of the valve, flow will occur. When the valve is closed, the hole is perpendicular to the ends of the valve, and flow is blocked. There are four types of ball valves.

A full port ball valve has an over sized ball so that the hole in the ball is the same size as the pipeline resulting in lower friction loss. Flow is unrestricted, but the valve is larger. This is not required for general industrial applications as all types of valves used in industry like gate valves, plug valves, butterfly valves, etc have restrictions across the flow and does not permit full flow. This leads to excessive costs for full bore ball valves and is generally an unnecessary cost.

In reduced port ball valves, flow through the valve is one pipe size smaller than the valve's pipe size resulting in flow area becoming lesser than pipe. But the flow discharge remains constant as it is a multiplier factor of flow discharge (Q) is equal to area of flow (A) into velocity (V). A1V1 = A2V2; the velocity increases with reduced area of flow and decreases with increased area of flow.

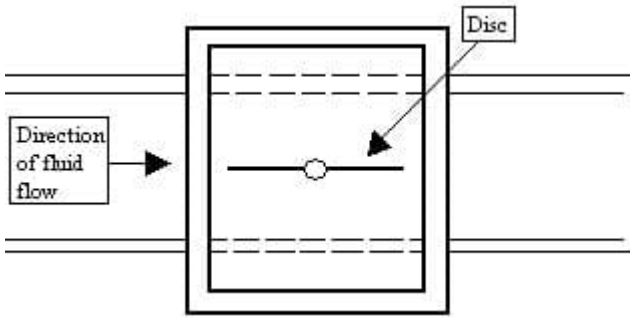
A V port ball valve has either a 'v' shaped ball or a 'v' shaped seat. This allows the orifice to be opened and closed in a more controlled manner with a closer to linear flow characteristic. When the valve is in the closed position and opening is commenced the small end of the 'v' is opened first allowing stable flow control during this stage. This type of design requires a generally more robust construction due to higher velocities of the fluids, which would quickly damage a standard valve.

A trunnion ball valve has a mechanical means of anchoring the ball at the top and the bottom, this design is usually applied on larger and higher pressure valves (say, above 10 cm and 40 bars).

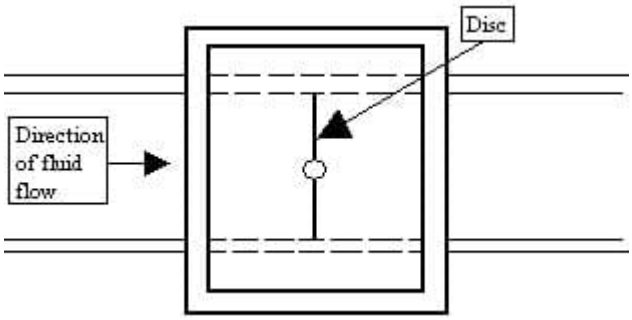
Ball valves are good for on/off situations. A common use for a ball valve is the emergency shut off for a sink.

Butterfly Valves

Butterfly valves consist of a disc attached to a shaft with bearings used to facilitate rotation. The characteristics of the flow can be controlled by changing the design of the disk being used. For example, there are designs that can be used in order to reduce the noise caused by a fluid as it flows through. Butterfly valves are good for situations with straight flow and where a small pressure drop is desired. There are also high performance butterfly valves. They have the added benefit of reduced torque issues, tight shutoff, and excellent throttling. It is necessary to consider the torque that will act on the valve. It will have water moving on both sides and when being used to throttle the flow through the valve it becomes a big factor. These valves are good in situations with high desired pressure drops.They are desirable due to their small size, which makes them a low cost control instrument. Some kind of seal is necessary in order for the valve to provide a leak free seal. A common example would be the air intake on older model automobiles.



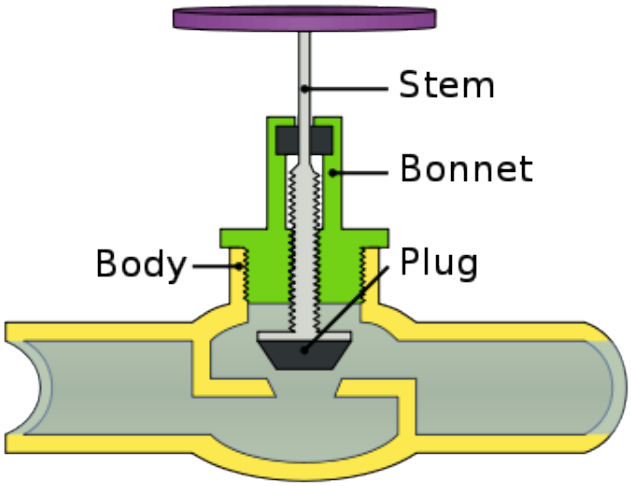
Top view, open configuration



Top view, closed configuration

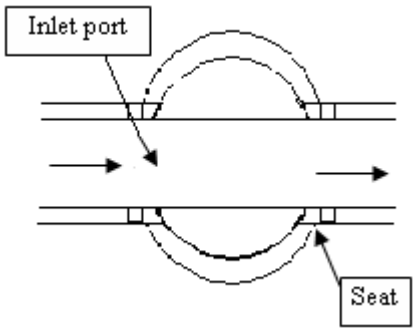
Globe Valves

A globe valve is a type of valve used for regulating flow in a pipeline, consisting of a movable disk-type element and a stationary ring seat in a generally spherical body. The valve can have a stem or a cage, similar to ball valves, that moves the plug into and out of the globe. The fluid's flow characteristics can be controlled by the design of the plug being used in the valve. A seal is used to stop leakage through the valve. Globe valves are designed to be easily maintained. They usually have a top that can be easily removed, exposing the plug and seal. Globe valves are good for on, off, and accurate throttling purposes but especially for situations when noise and caviatation are factors. A common example would be the valves that control the hot and cold water for a kitchen or bathroom sink.

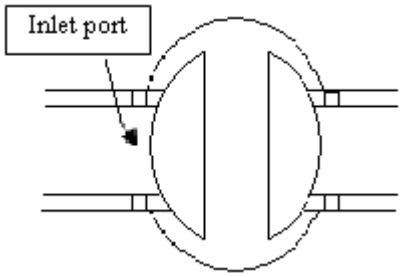


Plug Valves

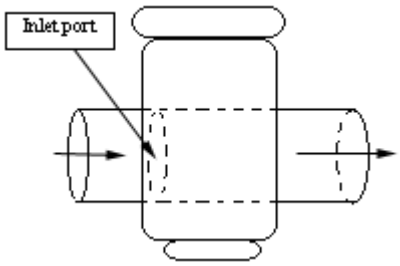
Plug valves are valves with cylindrical or conically-tapered "plugs" which can be rotated inside the valve body to control flow through the valve. The plugs in plug valves have one or more hollow passageways going sideways through the plug, so that fluid can flow through the plug when the valve is open. Plug valves are simple and often economical. There are two types of plug valves. One has a port through a cylindrical plug that is perpendicular to the pipe and rotates to allow the fluid to proceed through the valve if in an open configuration. In the closed configuration, the cylinder rotates about its axis so that its port is no longer open to the flow of fluid. An advantage of these types of valves is that they are excellent for quick shutoff. The high friction resulting from the design, however, limits their use for accurate modulating/throttling. Schematics of this type of plug valve are below.



Top view, open configuration

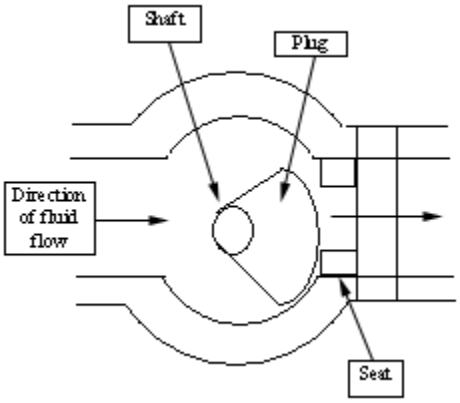


Top view, closed configuration



Side view

The other type of plug valve is the eccentric plug valve. In this design, the plug rotates about a shaft in a fashion similar to a ball valve. To permit fluid flow, the plug can rotate so that it is out of the way of the seat. To block fluid flow, it rotates to the closed position where it impedes fluid flow and rests in the seat. A schematic of this valve is below.



Top view, near-closed configuration

A common example would be a spray nozzle at the end of a garden hose.

Safety Valves

Pressure Relief Valves

Pressure relief valves are used as a safety device to protect equipment from over-pressure occurrences in any fluid process. Loss of heating and cooling, mechanical failure of valves, and poor draining and venting are some of the common causes of overpressure. The relieving system depends on the process at hand; pressure relief valves either bypass a fluid to an auxiliary passage or open a port to relieve the pressure to atmosphere. Some areas of common usage include reaction vessels and storage tanks. In the Petroleum Refining Industry, for example, the Fluidized Catalytic Cracker (FCC) reactor has several pressure relief valves to follow saftey codes and procedures on such a high pressure/high temperature process. Each of the pressure relief valves have different levels of pressure ratings to release different amounts of material to atmosphere in order to minimize environmental impact.

Here are three examples of pressure relief valves:

Conventional Spring Loaded Safety Valve

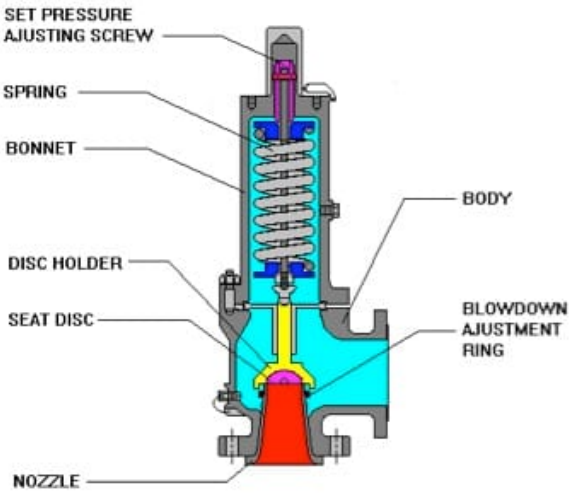
As the pressure rises, this causes a force to be put on the valve disc. This force opposes the spring force until at the set pressure the forces are balanced and the disc will start to lift. As the pressure continues to rise, the spring compresses more, further lifting the disc and alleviating the higher pressure. As the pressure inside the vessel decreases, the disc returns to its normal closed state.

Advantages:

- Most reliable type
- versatile

Disadvantages:

- Pressure relief is affected by back pressure
- Susceptible to chatter



Bellows Spring Loaded Safety Relief Valve

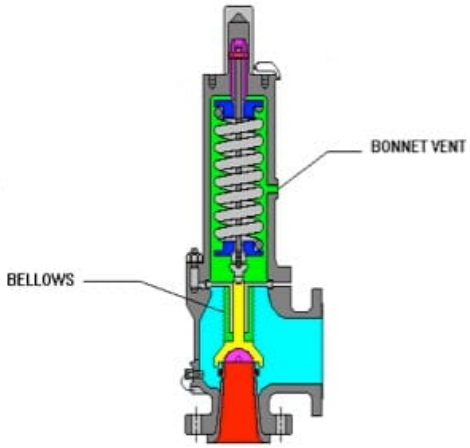
The bellows spring loaded safety relief valve has the same principle as the conventional spring valve, with the exception of a vent located on the side of the valve. This vent lets releases the contents of the valve out to the surrounding environment.

Advantages:

- Pressure relief is not affected by back pressure
- Can handle higher built-up back pressure
- Spring is protected from corrosion

Disadvantages:

- Bellows can be susceptible to fatigue
- Not environmentally friendly (can release of toxics into atmosphere)
- Requires a venting system



Pilot Assisted Safety Relief Valve

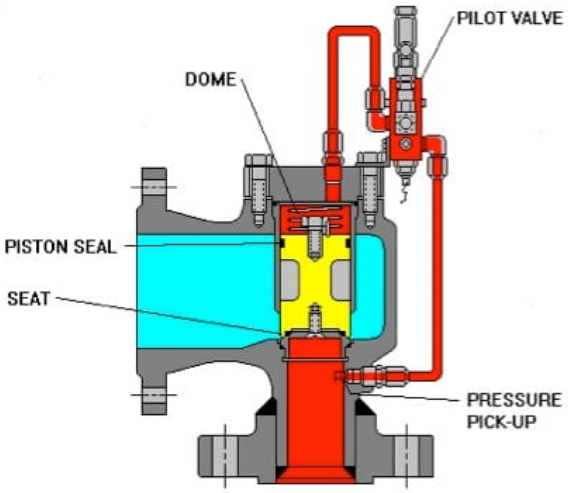
The pilot operated safety relief valve is also similar to the conventional safety relief valve except a pneumatic diaphragm or piston is attached to the top. This piston can apply forces on the valve when the valve is closed to balance the spring force and applies additional sealing pressure to prevent leakage.

Advantages:

- Pressure relief is not affected by back pressure
- Can operate at 98% of set pressure
- Less susceptible to chatter

Disadvantages:

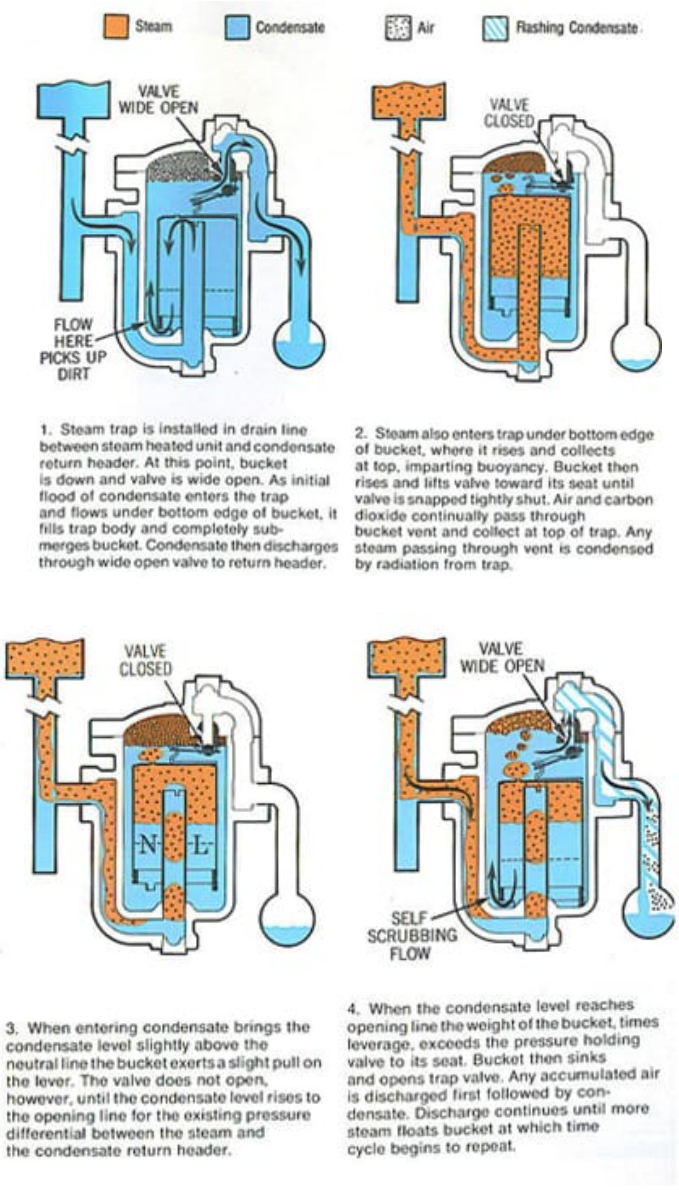
- Pilot is susceptible to plugging
- Has limited chemical use
- Condensation can cause problems with the valve
- Potential for back flow



Steam Traps

Steam traps are devices that exist in low lying places within a pressurized steam line to release condensate and non-condensable gases from the system. Steam lines in industry are used to open/close control valves, heat trace pipelines to prevent freezing, etc. These steam traps are used in industry to save money on the prevention of corrosion and loss of steam. When these traps fail, it can mean a lot of money for the industry. There can be several hundred to several thousand in one process unit, therefore it is important to maintain and check the condition of each trap annually. The checks can be done by visual, thermal, or acoustic techniques. Many suppliers have equipment to read the flow within the pipeline to see if it is: blocked, cold, leaking, working.

There are many types of steam traps that differ in the properties they operate on including Mechanical (density), Temperature (temperature), and Thermodynamic (pressure). For more information on the purpose and function of steam traps please view [Steam Traps].



Other Safety Valves

- *Rupture Disc* - A rupture disc (also called a "bursting disc" or "safety disc") is a thin membrane of material (usually metal) that acts as a one-time use pressure relief device. At a critical pressure, the disc will fail and rupture allowing flow and the release

of pressure. Often, rupture discs are used as a back-up to a conventional spring-controlled safety valve. Given primary safety valve failure (either no or inadequate pressure relief) the rupture disc will burst open and relieve pressure. Rupture discs are commonly used in chemical process plants and large air-cooled steam condensers.
- *Pressure Vacuum Valves* - Pressure vacuum (P-V) valves protect against both over-pressure and under-pressure conditions. They are commonly used in atmospheric storage tanks, to prevent the build up of excessive pressure or vacuum which can dangerously unbalance the system or damage the storage vessel.

Other Valves

In addition to the four main types of control valves, there are numerous other valves that may be necessary to manipulate fluid flow in a certain system. Some valves important to chemical engineering processes are stated below with a brief description of their design and application.

- *Angle valves* include inlet and outlet ports that are oriented at a 90 degree angle. The fluid leaves at right angles to the direction in which it enters the valve.
- *Bleed valves* vent signal line pressure to atmosphere before removal of an instrument or to assist in calibration of control devices. Common bleed valves include ball and plug bleed valves.
- *Check valves* are one way valves. Check valves only allow fluid in one way and out the other. They are often placed on individual fluid streams when mixing fluids so as to prevent the mixture from flowing back into the original streams. Also, the speed with which the valve closes is significant to prevent reverse-flow velocity. There are different types of check valves which include lift check, swing check, tilting disk and diaphragm.
- *Lift check valves* are quick closing. This valve should only be used for low viscosity fluids because they can be slowed by viscous fluids.
- *Swing check valve* has a disc like closing method from a hinge which may or may not be spring loaded.
- *Tilting disk check valve* is spring loaded for quick response. These are often more expensive and harder to fix when broken.
- *Diaphragm valves* have excellent shut-off characteristics and are used when there is particulate matter in the fluids. Not a good choice for controlling flow. A diaphragm valve has both a flexible and a rigid section. One advantage is that there are no crevices that affect the flow of the fluid when open. Used mostly in the pharmaceutical and food industry for their ability to be sanitary.
- *Piston valves* have a closure member shaped like a piston. When the valve opens, no flow is observed until the piston is fully moved from the seat bore, and when the valve closes, the piston removes any solids that might be deposited on the seat. Therefore, piston valves are used with fluids that have solid particles in suspension.
- *Gate valves* work by raising a gate or barrier out of the way of the flowing fluid. The water spigot outside a house works as a gate valve. Have the positive quality that, when fully open, the fluid flow is totally unobstructed. Two major types of gate valves are used: parallel and wedge gate valves. The wedge gate valve, in which the closure member has a wedge shape, provides the advantage of sealing against low pressure, as well as high pressure, flow. Gate valves have the ability to open very quickly. Advantages of gate valves are that they have a high capacity, have good seals, relatively inexpensive, and don't have very much resistance to flow. Some disadvantages of gate valves are that they sometimes can have poor control, they can cavitate at lower pressures, and they cannot be used for throttling.
- *Needle valves* are similar to gate valves; however, they are usually applied to a much smaller orifice. Needle valves are excellent for precise control of fluid flow, typically at low flow rates.
- *Flush Bottom valves* are usually at the lowest point of a tank or reactor and used to drain out contents. Unique because it leaves no dead space in the valve when it is closed, this eliminates the problem of product buildup within the valve.
- *Pinch valves* are mainly used in order to regulate the flow of slurries in certain processes and systems. Pinch valves have flexible bodies that can be shut by pinching them. They are often used when it is necessary for the slurry to pass straight through when the valve is not pinched. They are often used for sanitary processes such as medical and pharmaceutical applications. Pinch valves can be controlled mechanically or with fluid pressure.
- *Knife valves* are used in systems that deal with slurries or powders. They are primarily used for on and off purposes; whether or not the slurry or powder flows or not. A knife gate valve can be used for fibrous material because it can cut through to close the valve.
- *Ballcock valves* are used in controlling levels in tanks. The valve is connected to a float in the tank using a lever. When the level in the tank rises, the float rises and forces the valve to be shut at the maximum level of the tank allowed. Ballcock valves are used mostly in water tanks and other low-risk areas within a certain process.
- *Solenoid valves* are used very frequently in industry. The valves have a loop of wire that is wrapped around a metal core. A current is shot through the valve creating a magnetic field, which in turn opens or closes plungers in pipelines controlling flow in that pipe. There are three types of solenoid valves:

1. Electromechanical solenoid valves use electrical energy

2. Pneumatic solenoid valves use compressed air
3. Hydraulic solenoid valves use energy from pressurized fluids

Flow Characteristics

For all valves, flow rates of fluid through the valve depend on the percentage of a full valve opening. In ball and butterfly valves, the valve opening is based on rotation. In the case for butterfly valves, an open valve is a result of a 90 degree rotation. When the valve is at the open position, the flow is parallel to the valve. The flow is uninterrupted, and therefore, no pressure is on the valve. When the valve is throttling, as is common for globe valves, the flow rate is reduced or increased depending on the opening of the valve, and there is unequal pressure on the ends of the valve.

The ball valve flow stream involves at least two orifices; one for inlet and one for outlet flow. While traveling through an open ball valve, the fluid will continue in its flow straight through, with little pressure loss. When the ball valve is throttling, however, the fluid is subject to shearing and a change in flow rate in accordance with the percentage of the valve that is open. With high velocity liquids, the valve is susceptible to cavitation and erosion, and could produce noise. Also, at sonic velocities, the vena contracta will stop expanding, and create choking.

Globe valves are dependent on shape of the plug of the valve, in addition to the opening size, for flow variations. Lifting the globe valve will cause it to open. A flat plug is used for a quick opening; a conically shaped plug will create linear flow as the valve is raised; a rectangular shaped plug, where the bottom converges to one point directly in the middle of the plug, creates an equal percentage of flow through the valve.

Sizing

Once a specific valve type is chosen for a process, the next step is to figure out what size the valve will need to be. Topics to consider in valve sizing are the pressure-flowrate relationship and the flow characteristics to ensure an appropriate size for the desired valve.

Pressure-Flowrate Relationship

Control valves predominately regulate flow by varying an orifice size. As the fluid moves from the piping into the smaller diameter orifice of the valve, the velocity of the fluid increases in order to move a given mass flow through the valve. The energy needed to increase the velocity of the fluid comes at the expense of the pressure, so the point of highest velocity is also the point of lowest pressure (smallest cross section). This occurs some distance after leaving the smallest cross section of the valve itself, in a localized area known as the vena contracta.

Beyond the vena contracta, the fluid's velocity will decrease as the diameter of piping increases. This allows for some pressure recovery as the energy that was imparted as velocity is now partially converted back into pressure. There will be a net loss of pressure due to viscous losses and turbulence in the fluid.

Overview of Sizing Formulas

In order to determine the correct size of a valve for a specific system many factors must be considered. The most important factor is the capacity parameter, C_v, or the **flow coefficient**. The flow coefficient is a way of measuring how efficient a valve is at allowing fluid to flow through it and generally determined experimentally. Valve manufacturers can provide you with Cv charts for the valves they sell. To determine the valve size needed for your system, you can estimate C_v with the following equations:

Liquid Flow

$$C_v = 11.7Q\sqrt{\frac{G_f}{\Delta P}}$$

where:

The next equation is also for calculating the flow coefficient. It should be used when you have a design flow rate, that will be the maximum flow through the valve, and a maximum pressure drop that will occur across the valve, which will be ΔP. This equation

is also for turbulent flow and the units used in the calculation are important.

$$C_v = F_{\max}\sqrt{\frac{G_t}{\Delta P}}$$

where:

Another important piece of information about sizing and specifications is what diameter pipe may be used with a certain flow, seen in the following equation:

Note: The second and third equations are from Peter Woolf's Fall'08 lecture titled Modeling Case Study: Surge Tanks, Valves, Level sensors, and modeling.

Air and Gaseous Flow

(when $P_o < 0.53P_i$)

$$C_v = Q\frac{\sqrt{G_f(T+460)}}{660P_i}$$

(when $P_o > 0.53P_i$)

$$C_v = Q\frac{\sqrt{G_f(T+460)}}{1360\sqrt{\Delta P\,(P_o)}}$$

where:

The relationship between the inlet and outlet pressure is important, as seen above, in determining which equation to use for gaseous flow. When the outlet pressure is less than 53% the inlet pressure, this is said to be a **critical** pressure drop. The gas will behave differently when the pressure drop is critical, therefore it is necessary to use the correct equation depending on the extent of pressure drop. Once you have computed the desired value of C_v, you can choose the valve. The chosen valve must have a valve coefficient greater than or equal to the computed value.

Other sizing information

When sizing a valve it is important not to choose a valve that is too small. Take the following tips into consideration when choosing a valve.

- Valves should not be less than half the pipe size.
- Valves are easiest to control when they are open from 10 to 80% of maximum flow.

Using the Flow Coefficient to Determine Valve Sizes

The flow coefficient, C_v, is used as a standard sizing parameter. It is used to determine valve sizes and is also used in modeling programs for large piping systems. The C_v is an essential and practical variable that is essentially required when using pipes and valves. Try searching for a valve information and you will most likely find information requiring the C_v as well as the valve type and use.

Try looking at information for specific valves at the website of American Valve. [1]

You'll notice that every valve type and size has a listed C_v for correct sizing and function. Tables using this value are available for practically every valve imaginable and should be used to prevent over-use or under-use of valves in engineering processes.

Cv Charts

When sizing a valve, it is important to choose the correct Cv chart. Different Cv charts are used for different types of valves. Some different valve types are explained below. Note that *valve stroke* means the amount of flow through a pipe, and *valve travel* means the amount in which the valve is turned to achieve the valve stroke.

- **Equal Percentage Valves:** Valve which produces equal valve stroke for equal increments in valve travel. This is the most common type of valve.
- **Linear Valves:** Valves stroke is directly proportional to valve travel.
- **Quick Opening Valves:** A small amount of valve travel produces a large valve stroke. Example: Gate Valves

Equal percentage and linear valve types offer the best throttling control. Examples are globe and butterfly valves.

Valves vary from manufacturer to manufacturer. Valve manufacturers can provide you with Cv charts to size each type of valve they manufacture. An example CV chart is shown below. VT stands for valve travel.

Size (NPT)	Cv @ 10% VT	Cv @ 25% VT	Cv @ 50% VT	Cv @ 75% VT	Cv @ 100% VT
1/2"	0.03	0.075	0.3	0.5	1
1/2"	0.05	0.12	0.48	0.8	1.6
1/2"	0.08	0.19	0.75	1.25	2.5
1/2"	0.13	0.3	1.2	2	4
3/4"	0.21	0.47	1.89	3.15	6.3
1"	0.33	0.75	3	5	10
1 1/4"	0.53	1.2	4.8	8	16
1 1/2"	0.83	1.88	7.5	12.5	25
2"	1.33	3	12	20	40
2 1/2"	2.10	4.73	18.9	31.5	63
3"	3.33	7.5	30	50	100
4"	5.33	12	48	80	160
5"	8.33	18.75	75	125	250
6"	13.33	30	120	200	400

Materials

It is usually not enough to simply select the type of valve suited to given process parameters. Selecting compatible materials of construction helps ensure the lifespan of the valve, as well as the safety of the workforce, the environment and the public.

Selecting the most appropriate materials of construction of control valves is guided primarily by the service of the valve, then secondarily by cost; the least expensive material that is compatible with the service will be chosen to be used. Proper material selection promotes safety by avoiding materials of construction that may react with or be corroded by the process fluid.

The principle materials that need to be selected carefully are the wetted materials, that is, the materials that come into contact with the process fluid. These generally include the ball (for ball valves), the disk (for butterfly valves), and the plug (for plug and globe valves). Also included are the seats, which is the area where the plug or disk "sits" when closed to provide the actual shut off. The seals and the valve body are usually wetted as well.

There are many resources that contain what resources are compatible with a wide variety of process fluids, such as the Cole Parmer Chemical Resistance Database and the Cat Pumps' Chemical Compatibility Guide

There are design parameters inherent in the valve designs themselves that increase safety. For high service pressures (or in case of fire) some valves are designed with initial flexible seal rings that function as the primary seals. Behind these primary seals would be a backup seal of a more durable material such as 316 stainless, inconel or hastelloy. These backup seals assist in handling the additional pressure and heat.

In the highly specialized case where the process fluid is so dangerous or unsafe that any release of process fluid is unacceptable, the valve's packing can be slightly pressurized with a barrier fluid. As long as the pressure of the barrier fluid is higher than the process fluid, any leakage between the valve and the process will leak barrier fluid into the process, and not process fluid into the environment. Though as a side note, these applications usually require double containment piping and a whole host of other precautions beyond simply the safety of the valve. The most common barrier fluid is water or a water/antifreeze mix for freeze protection.

Some other considerations when selecting a material for valve are longevity/reliability of the valve and the temperature range of usage. If the valve is a control valve that gets constant use, it is important to select durable materials or to plan for replacement of the valve frequently. Service temperature is also important; materials need to be selected so the mechanical integrity of the valve is maintained throughout the entire service temperature.

Concerns

Noise

Noise in a system is created by the turbulence from a throttling valve. Throttling is when a valve is not used as a simple OPEN/CLOSE valve, but instead to control a flow rate.

It can be quite loud and damage people's hearing as well as the equipment. Two methods can be used to control noise. The first is to use a quiet trim that hinders system noise before it begins and the second is to treat the path (piping) where the noise already occurs. To do this, one can use thicker piping or insulation.

Cavitation

As previously mentioned, at the point where the fluid's velocity is at its highest, the pressure is at its lowest. Assuming the fluid is incompressible (liquid), if the pressure falls down to the vapor pressure of the fluid, localized areas of the fluid will vaporize forming bubbles that collapse into themselves as the pressure increases downstream. This leads to massive shock waves that are noisy and will certainly ruin process equipment. This is a process called cavitation.

For a given flow rate, a valve that has the smallest minimum flow area through the valve will have the most problems with cavitation (and choking, as mentioned in the next section).

Choking

If the fluid is compressible and the pressure drop is sufficient, the fluid could reach sonic velocity. At this point the static pressure downstream of the valve grows above design limits. This is a process known as choking, since it slows down, essentially "choking," the flow through the valve.

For a given flow rate, a valve that has the smallest minimum flow area through the valve will have the most problems with choking.

Choking is important for a gas-using process since the mass flow rate only depends on the upstream pressure and the upstream temperature. It becomes easier to control the flow rate since valves and orifice plates can be calibrated for this purpose. Choke flow occurs with gases when the ratio of the upstream pressure to the downstream pressure is greater than or equal to [(k + 1) / 2]^(k / (k - 1)). If the upstream pressure is lower than the downstream, it is not possible to achieve sonic flow.

When choking occurs in liquids, it is usually due to the Venturi effect. If the Venturi effect decreases the liquid pressure lower than that of the liquid vapor pressure at the same temperature, it will cause the bubbles in the liquid. These bubble burst can cause enough turbulence (known as cavitations) that can lead to physical damage to pipes, valves, controllers, gauges, and/or all other equipments linked to that section of the flow.

As mentioned above the limiting case of the **Venturi effect** is choked flow as the smaller space in a pipe will but a cap on the total flow rate. The Venturi effect is fluid pressure resulting from an incompressible fluid travelling through a tight or constricted section of a pipe. The Venturi effect can be explained by principles such as Bernoulli's principle and the continuity equation.

Bernoulli's equation : $$\frac{v^2}{2} + gz + \frac{p}{\rho} = \text{constant}$$

Continuity equation: $$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0$$

for incompressible fluids: $\nabla \cdot \mathbf{u} = 0$

To satisfy the continuity equation, the velocity of the fluid must increase in the space but the pressure must decrease to agree with Bernoulli's equation, which is related to the law of conservation of energy through a pressure drop.

Venturi tubes are used to control the fluid speed. If an increase in speed is wanted, then the tube will decrease in diameter, or projected area, as the fluid moves away from the pump or energy source. If a decrease in fluid velocity is wanted, the tube will increase in diameter as it moves away from the energy source (which could be a pump). Using a venturi tube sometimes introduces air into the system which is mixed with the fluid in the pipe. This causes a pressure head at the end of the system.

Venturi tubes can be replaced by **orifice plates** in some cases. Orifice plates use the same principles as the venturi tubes but cause the system to suffer more permanent energy loss. Orifice plates, however, are less expensive to build then venturi tubes.

Upstream Pressure Increase

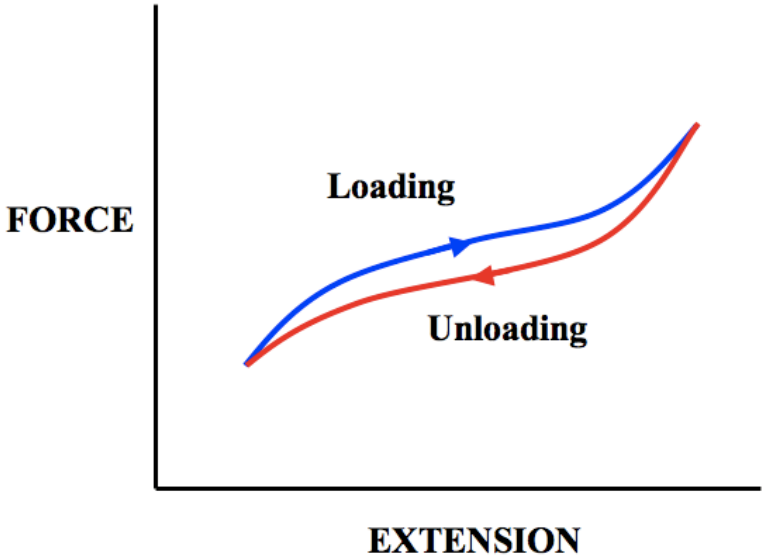
If a control valve is sized incorrectly and is too small, fluid velocity upstream of the valve will slow, causing an increase in pressure (much like when the end of a garden hose is partially obstructed). This increase in upstream pressure can be detrimental to certain processes, i.e. membrane filtration processes, where a large pressure difference across the membrane is desired. If the valve increases the pressure on the permeate side of the membrane, the driving force for separation will be reduced.

Decreasing the upstream static temperature will cause an increase in upstream static pressure therefore increasing the mass flow rate, so it is important to keep track of your temperature.

Hysteresis

Hysteresis, in general, is defined as the phenomenon where previous events can influence subsequent events in a certain process. This can create a relationship in a process going one direction, however, when carrying out the same process in the exact opposite direction, the same relationship does not occur. When processes exhibiting hysteresis are graphed there is an area between the two equally opposite curves that is used to describe the extent of hysteresis seen in the process.

The most commonly described process exhibiting hysteresis involves elastics and can be the loading and unloading of force (weight) to a stretchable rubber band. As you load more and more weight onto a rubber band, it will stretch in response to increased weight. You can subsequently remove the weight in similar increments and observe the ability of the rubber band to compress itself. Rubber bands, however, do not obey Hooke's Law perfectly therefore as you unload, the rubber band will be longer at the corresponding weights as you loaded. Below is a graphical representation of this phenomenon.



As seen in the above graph, for any given force applied to the rubber band, the extension of unloading is larger than the extension

when loading. This is showing how the process of loading effected the outcome of the process of unloading. The area of the graph between the two curves is the area of hysteresis and, in this case, represents the energy dissipated as heat. The larger this area the less efficient this rubber band was at maintaining shape, and consequently the worse it obeys Hooke's Law.

Hysteresis, regarding control valves, is the principle that a control valve is dependent on the prior valve position. A common valve opening may correlate to different flow rates depending on if the valve was opened or closed to get to that position. This shows that the direction a valve is being changed may need to be accounted for in a control system to obtain a desired flow rate. If hysteresis becomes too large, it can cause the control architecture to force a system to oscillate around a desired point.

Chatter

Chatter is mainly a concern when selecting and sizing safety valves. Chatter is the rapid opening and closing of a safety valve or pressure relief device. Commonly, chatter is caused due to a large pressure drop between the vessel being relieved and the valve relieving the pressure.

If the pressure behind a closed safety valve builds up to some critical pressure, it will open. Once open, fluid will flow from the vessel being relieved out through the safety valve in order to reduce pressure within the vessel. It is possible to have enough pipe length, fittings, or other energy-reducing obstacles so that once the high pressure flow reaches the safety valve, the pressure in the fluid is again below the critical pressure of the valve. Since there is not enough pressure in the fluid to keep the safety valve open, it closes and the vessel re-pressurizes, causing the safety valve to pop open again, locally depressurize and close again. This opening and closing is called chatter. The vibrations caused by chatter can be extremely damaging, causing unnecessary safety valve wear, possible seat-misalignment within the fittings, and even full failure in the valve or surrounding piping.

Chatter can be avoided by making sure that at critical release rates the pressure drop across the fittings to the safety valve is not large enough such that the valve will shut once fluid release is initiated.

Actuators

Actuators are the mechanical equipment that supply the force necessary to open or close a valve. Actuators are, essentially, the alternative to manual operation of a valve. The method of applying the opening/closing force to a valve is what differentiates the various types of actuators. When selecting the actuator, the most important feature to specify is whether you want fail-safe open or closed. This is determined entirely by a careful analysis of the process to decide which is safer. If all the power goes out or some other emergency occurs, the decision on the fail-safe mode of the valve is a huge factor in saving lives. For example, if a reaction is exothermic, the fail-safe mode of the cooling water in the jacket of the reactor should be fail-safe open. Pneumatic actuators have a delayed response which make them ideal for being resilient against small upsets in pressure changes of the source. Hydraulic actuators on the other hand use an incompressible fluid, so the response time is essentially instantaneous.

Pneumatic

Pneumatic actuators are the most popular type of actuators. The standard design of a pneumatic actuator consists of a pre-compressed spring that applies force against a disk on a sealed flexible chamber. The disk is usually attached to the stem of the valve it is intended to control. As the chamber is compressed with air, the chamber expands and compresses the spring further, allowing axial motion of the valve stem. Knowing the relationship between the air pressure in the chamber and the distance the stem moves allows one to accurately control flow through the valve.

The biggest advantage of the pneumatic actuators is their failsafe action. By design of the compressed spring, the engineer can determine if the valve will fail closed or open, depending on the safety of the process. Other advantages include reliability, ease of maintenance, and widespread use of such devices.

Motion conversion

Motion conversion actuators are generally used to adapt a common translational motion from the actuator’s output to a rotary valve. The rod that moves axially from the translational motion actuator is connected to a disk and the connection is pivoted. The disk itself is also pivoted about its center. This system of pivots allows the translational motion to be converted into the rotation of the disk, which would open or close the rotary valve.

The main advantage of this setup is that an inexpensive translational motion actuator can be used with rotary valves. The key drawback is that the applications in which this can be used is very limited. Specifically, this setup is useless in the common case where the rotary motion required is greater than 90°.

Hydraulic

Hydraulic actuators work using essentially the same principal as pneumatic actuators, but the design is usually altered. Instead of a flexible chamber, there is a sealed sliding piston. Also, instead of using a spring as the opposing force, hydraulic fluid is contained on both sides of the piston. The differential pressure across the area of the piston head determines the net force.

Hydraulic actuators offer the advantages of being small and yet still providing immense force. Drawbacks of hydraulic actuators are primarily the large capital cost and difficulty maintaining them.

Electric

Electric actuators typically use standard motors, powered by either AC induction, DC, or capacitor-start split-phase induction. The motor is connected to a gear or thread that creates thrust to move the valve. As a failsafe, some motors are equipped with a lock in last position on its gear. This means that the gear cannot move from forces outside of the electric motor. This helps prevent overshoot on the motor as well as helps create better positioning for the gear.

Another type of motor that can be used is called a stepper motor. It uses increments on gear reduction to alleviate problems with positioning and overshoot. The increments are in a range of 5,000 to 10,000 increments in a 90 degree rotation.

A problem with using electric actuators is that a battery operated back-up system is needed or else the system is useless during power failure. Also, the actuator needs to be in an environment that is rendered safe, meaning a non-explosive environment.

Manual

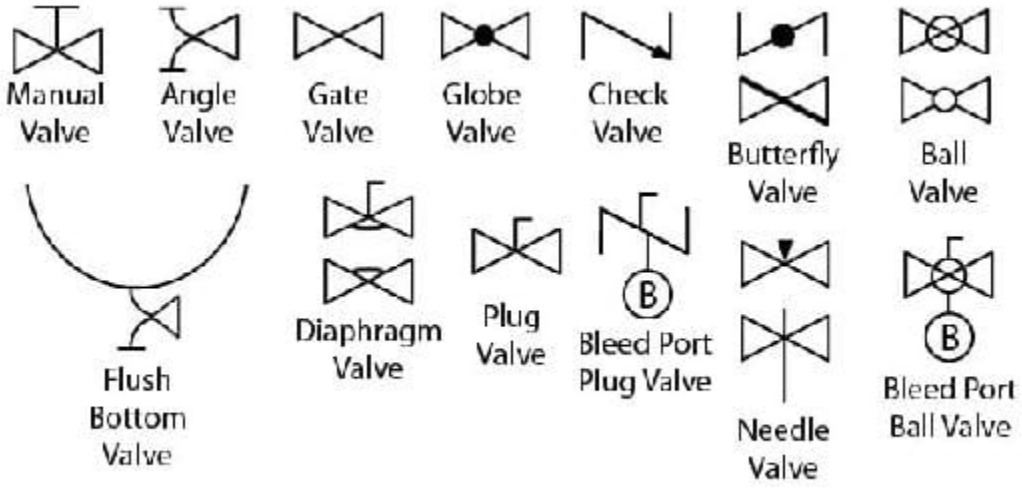
Manual actuators are usually used for overrides of power actuators described above. This is an important safety measure in case the power actuator fails. Manual actuators typically consist of either a lever or a wheel (used for larger valves) connected to a screw or thread that turns the valve.

Summary Tables

Valve Type	Description	Advantages	Disadvantages
Ball	A ball that can be spun to allow or close flow	Operated in a 1/4 turn for quick on/off application	Size limitations
Butterfly	A disk that is rotated about it's diameter to allow, throttle, or close flow	Low pressure drop	Torque issues
Globe	A plug that moves into and out of a globe to allow, throttle, or close flow	Precise throttling	High pressure drop
Plug	A plug that rotates to allow, throttle, or close flow	Quick shutoff	Limited throttling ability

Actuator	Description	Advantages	Disadvantages
Pneumatic	Compressed air fills a chamber and pushes against a spring, operating the valve	Failsafe action, reliability, low cost, and ease of maintenance	Limited force applications
Motion Conversion	Consists of a disk connected to a rod. The rod moves axially and pivots the connection with the disk, causing the disk to also pivot.	Inexpensive and can be used for rotary valves	Limited to applications of 90° or less rotations.
Hydraulic	Similar to pneumatic, but a uses an incompressible fluid	Small size with large force	High cost
Electric	Electric motor with a rotating gear that moves the valve.	Low overshoot and accurate positioning	Needs a backup battery in case
Manual	Typically either a level or wheel connected to a screw	Good for an override in case of failure of	Not efficient as primary

The table below shows a list of typical icons for different valves found in industry.



Valve and Actuator Selection Example

Note: This problem is completely fictionalized. Any relation to a real occurrence is completely coincidental.

A chemical engineer wants to use a valve at a start/stop flow of liquid water to a CSTR. The determination of the start/stop flow is a level sensor in the CSTR. Once the water reaches a certain level inside the CSTR, the flow of the water will stop. If the water level drops to a certain point, water will need to quickly flow into the CSTR. Your job is to list possible valves, and determine if each valve type can be used to start/stop the flow of water. Also, list all the potential actuators for the valves and determine which actuators *could* be used and how. Some issues to consider are choking, throttling, longevity, and reliability.

Solution

Valves:

Ball valve: A ball valve can work for this example. A 2-way ball valve may not be the most ideal type of valve depending on the flow rate. If the water is moving at a high velocity, the 2-way ball valve will be susceptible to choking. If a ball valve is used, a cage valve would be the best choice, due to its durability and its on/off mechanics.

Globe valve: Just like the caged ball valve, a globe valve is a conceivable solution. The plug that would best help prevent water

from flowing in the CSTR quickly would be the quick opening plug. This way, there will be no throttling when closing the valve, creating a quick on/off.

Butterfly valve: If a butterfly valve is selected, it would need to be a high performance valve. The benefits of this valve include a tight shutoff and improved torque over the non-high performance butterfly valves. This is important for the on/off capabilities for water flow the example calls for. Using a regular butterfly valve could lead to water leaking into the CSTR or torque issues after constant use.

Plug valve: A plug valve could work, but would encounter similar problems as a 2-way ball valve with choking. Also, the friction created when constantly rotating could cause a problem in the future with the valve not closing properly.

Actuators:

Pneumatic: A pneumatic actuator has a good potential to be used in this example. Its failsafe action is great in case there is a failure in the system. Its low maintenance and high reliability means that the pneumatic actuator is an ideal actuator for the valve.

Motion Conversion: Motion conversion actuators would be ideal for any rotating valve, like the high performance butterfly valve. Therefore, this actuator should be used with rotational valves for best reliability.

Hydraulic: Hydraulic actuators have similar advantages that the pneumatic actuators have. Hydraulic actuator, however, cost more than pneumatic actuators. Therefore, if another option exists, hydraulic actuators should not be used for this example.

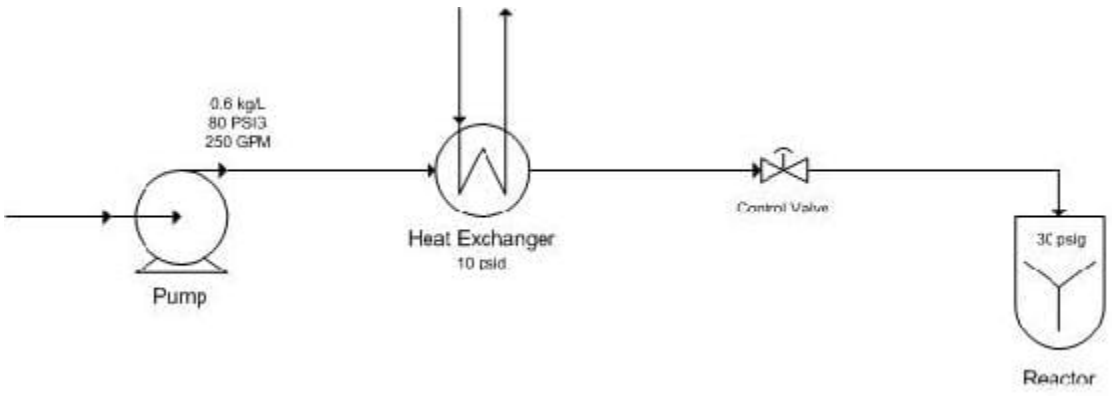
Electric: An electric actuator will work best with rotational valves. So, if a high performance butterfly valve was chosen, then an electric actuator could possibly be used. As an effort to limit throttling issues, rotational models for ball valves would be less ideal. This is, of course, on the assumption that a back up battery would be included with the electric actuator

Manual: A manual actuator could be used in conjunction with any other actuator, but would it not be a good idea to use a manual actuator alone. The manual actuator would be a good backup method if any of the other actuators fail, but tedious if used as the primary actuator.

Selecting Control Valve Example

Note this example problem and the data contained within fictitious, with the exception of some of the Cv data, which is taken from Clark's Product Specification Site

An engineer in a chemical factory is asked to decide what type of control valve to use in throttling the line from a heat exchanger to a reactor. The chemical is part of a very sensitive reaction in which the precise addition of the chemical in question is crucial. If an imprecise valve is utilized, the entire reactor's $5 million worth of product would need to go down the sewer. The process considerations are as follows:



The process is designed to pump the liquid chemical at 80 psig and 250 GPM, though the flow occasionally needs to be varied from 125 to 400 GPM. The chemical has a density of 0.60 kg/L (S.G. = 0.6). The heat exchanger should take 10 psi of the pressure drop across it. Also, the reactor is operated at 30 psig Assume turbulent non-choked flow.

Solution:

First one must choose the type of valve. The four main valves of interest are ball, butterfly, plug and globe. Since the application is a throttling application a ball valve probably should not be our first choice. Since the chemical needs to be throttled at an extremely accurate rate, a plug valve would not be best suited here. Also, since there is significant variability of the desired flowrate, a butterfly valve probably is not the best decision either. A globe valve seems to work well here.

For reference, a list of Cv's by valve size (when fully opened) is listed below. The Cv's of each valve are listed also by % valve travel (as can be commonly acquired from a vendor), where 100% means the valve is fully open.

Size (NPT)	Cv @ 10% VT	Cv @ 25% VT	Cv @ 50% VT	Cv @ 75% VT	Cv @ 100% VT
1/2"	0.03	0.075	0.3	0.5	1
1/2"	0.05	0.12	0.48	0.8	1.6
1/2"	0.08	0.19	0.75	1.25	2.5
1/2"	0.13	0.3	1.2	2	4
3/4"	0.21	0.47	1.89	3.15	6.3
1"	0.33	0.75	3	5	10
1 1/4"	0.53	1.2	4.8	8	16
1 1/2"	0.83	1.88	7.5	12.5	25
2"	1.33	3	12	20	40
2 1/2"	2.10	4.73	18.9	31.5	63
3"	3.33	7.5	30	50	100
4"	5.33	12	48	80	160
5"	8.33	18.75	75	125	250
6"	13.33	30	120	200	400

Now one must size the globe valve. Recall the equation

$$C_v = Q\sqrt{\frac{G_f}{\Delta P}}$$

We can see that the pressure drop across the valve is the pressure at the pump minus the pressure at the reactor minus the pressure drop across the heat exchanger. Thus

$$Delta\ \ P = 80 - 30 - 10 = 40\,\psi$$

So getting Cv,

$$C_v = 250\sqrt{\frac{0.6}{40}} = 30.62$$

Looking at the chart it would appear that the 2" Globe valve would work, but remember, the process fluid may sometimes need to be controlled up to 400 GPM. And the Cv at 400 GPM is:

$$C_v = 400\sqrt{\frac{0.6}{40}} = 49.0$$

Looking at this it is clear that the 2" globe valve would simply be fully open at 400 GPM. It appears that the 2 ½" globe valve would be a better choice since it has the capability to throttle up to the maximum design flowrate.

Checking the Cv at the minimum flow that will need to be throttled, 125 GPM:

$$C_v = 125\sqrt{\frac{0.6}{40}} = 15.3$$

We see that the 2 ½" globe valve will be between 25 and 50% open, so it will throttle well at these low flow rates at also. This reaffirms that the 2 ½" globe valve will work well for this application.

Example 3.9.1

When designing a jacketed CSTR that will contain a highly exothermic reaction, what type of actuator should be used to operate the valve controlling the cooling water?

a. pneumatic
b. motion conversion
c. hydraulic
d. electric

Solution

a

Exercise 3.9.2

Which of the following is a good application for a diaphragm valve?

a. inlet to a CSTR where precise volumes of reactants are required
b. food industry where ease of cleaning is important
c. steam heating line for jacketed CSTR
d. vent on a signal line to release pressure to atmosphere

Answer available at the end of this section.

References

- Q = flow rate (ft^3 / *h*)
- G_f = specific gravity of the gaseous fluid (dimensionless)
- ΔP = absolute pressure drop through the valve (psia)
- P_o= outlet pressure (psia)
- P_i = inlet pressure (psia)
- T = gas temperature (degrees F)
- Edgar, Thomas F., Sale E. Seborg, and Duncan A. Mellichamp. *Process Dynamics and Control*. New York: John Wiley & Sons.
- Liptak, Bela G., *Process Control and Optimization*. Vol. II. New York: Taylor & Francis.
- Perry, Robert H.(1997), *Perry's Chemical Engineers' Handbook*, 7th Edition. McGraw-Hill. ISBN 0070498415
- "Valves", *Cole-Parmer*, 2005
- Considine, Douglas M. (1993), *Process/Industrial Instruments & Controls Handbook*, 4th Edition, McGRAW-HILL, INC. ISBN 0070124450
- Smith, Peter; Zappe, R.W. (2004),*Valve Selection Handbook*. Elsevier
- Matthews,Clifford . (2004),*A Quick Guide to Pressure Relief Valves (PRVs)*. John Wiley and Sons. ISBN 1860584578, 9781860584572
- Websites for steam traps: www.energysolutionscenter.org/BoilerBurner/Eff_Improve/Steam_Distribution/Steam_Traps.asp http://en.Wikipedia.org/wiki/Steam_trap www.iklimnet.com/expert_hvac/steam_traps.html
- $$d = \sqrt{\frac{4F_{\max}}{\pi v}}$$

  Where:

Answers

Exercise 3.9.2

b

This page titled 3.9: Valves - Types, Kinds, and Selection is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

3.10: Valves - Modeling Dynamics

Authors: Erin Knight, Matthew Russell, Dipti Sawalka, Spencer Yendell

Introduction

A valve acts as a control device in a larger system; it can be modeled to regulate the flow of material and energy within a process. There are several different kinds of valves (butterfly, ball, globe etc.), selection of which depends on the application and chemical process in consideration. The sizing of valves depends on the fluids processing unit (heat exchanger, pump etc.) which is in series with the valve. Sizing and selection of valves is discussed in the other wiki article on Valve Selection. Valves need to be modeled to perform effectively with respect to the process requirements. Important components for the modeling of control valves are:

1. Flow
2. Inherent Flow Characteristics
3. Valve Coefficient, C_V
4. Pressure Drop
5. Control Valve Gain
6. Rangeability
7. Installed Characteristics

Efficient modeling of the valves can optimize the performance and stability of a process as well as reduce development time and cost for valve manufacturers.

In the following sections we briefly define the various variables and equations involved in modeling valves. The purpose of the following sections is to give you an overview of the equations required to model the valves for a particular system. Example problems at the end of the article have been provided to aid in the qualitative and quantitative understanding of how valves are modeled for chemical engineering processes.

Flow through a Valve

The following equation is a general equation used to describe flow through a valve. This is the equation to start with when you want to model a valve and it can be modified for different situations. The unfamiliar components such as valve coefficient and flow characteristics will be explained further.

$$F = C_v f(x) \sqrt{\frac{\Delta P_v}{sg}}$$

with

- F is the volumetric flow rate
- C_v is the valve coefficient, the flow in gpm (gallons per minute) that flows through a valve that has a pressure drop of 1psi across the valve.
- ΔP_v is the pressure drop across the valve
- sg is the specific gravity of fluid
- x is the fraction of valve opening or valve "lift" (x=1 for max flow)
- $f(x)$ is the flow characteristic

Flow Characteristics

The inherent flow characteristic, f(x), is key to modeling the flow through a valve, and depends on the kind of valve you are using. A flow characteristic is defined as the relationship between valve capacity and fluid travel through the valve.

There are three flow characteristics to choose from:

1. $f(x) = x$ for linear valve control
2. $f(x) = \sqrt{x}$ for quick opening valve control
3. $f(x) = R^{x-1}$ for equal percentage valve control

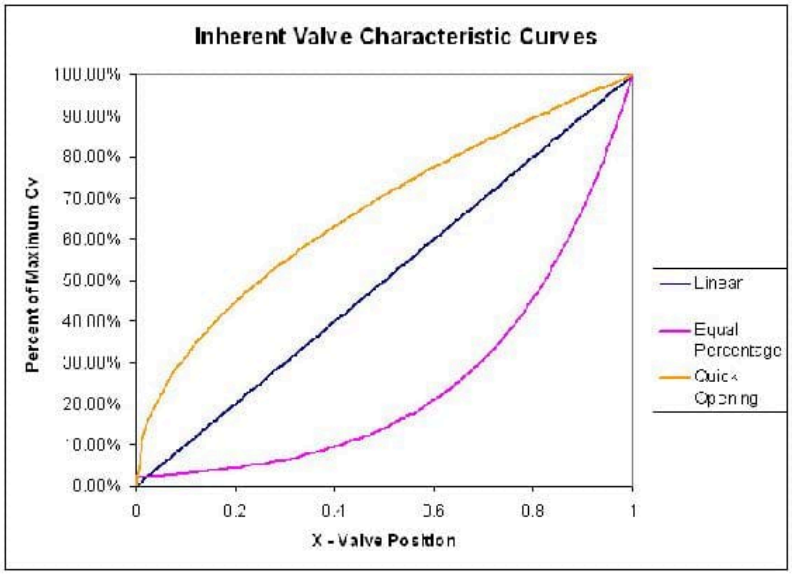
- R= valve design parameter (between 20 and 50)
- note these are for a fixed pressure drop across the valve

Whereas a valve TYPE (gate, globe or ball) describes the geometry and mechanical characteristics of the valve, the valve CONTROL refers to how the flow relates to the "openness" of the valve or "x."

1. Linear: flow is directly proportional to the valve lift (used in steady state systems with constant pressure drops over the valve and in liquid level or flow loops)
2. Equal Percentage - equal increments of valve lift (x) produce an equal percentage in flow change (used in processes where large drops in pressure are expected and in temperature and pressure control loops)
3. Quick opening: large increase in flow with a small change in valve lift (used for valves that need to be turned either on or off frequently or where instant maximum flow is required, for example, safety systems)

For the types of valves discussed in the valve selection article, the following valve characteristics are best suited:

1. Gate Valves - quick opening
2. Globe Valves - linear and equal percentage
3. Ball Valves - quick opening and linear
4. Butterfly Valves - linear and equal percentage



Cv_{max} depends on pipe characteristics and was chosen to be 110 gpm in this example. Constant pressure throughout the pipe line is assumed and the curves are accurate when the valve position is between 5% and 95% open.

Comparing the slopes of the graphs for the quick opening and equal percentage valves, we can see that a quick opening valve would experience greater change in flow with slight change in valve position in the lower flow range. The opposite is true for the higher range of flow. The equal percentage valve experiences slighter change in flow with respect to valve position in the lower range of flow.

When selecting the appropriate control valve, it is often the goal of the engineer to choose a valve that will exhibit a linear relationship between F and x over the normal operating position(s) of the valve. This linear relationship provides the most control for the operator. The flow characteristic observed through an installed valve, and all process factors considered (i.e. total pressure drop, etc.), is termed the installed flow characteristic. Therefore, it is not always the case that an inherently linear valve is desirable or even useful. An inherently linear valve is appropriate when there is a linear relationship between the valve position and the actual flow rate; however, consider the case where $\Delta P_L \neq 0$ and of significant value. In this case a valve with inherent equal

percentage flow characteristic would be appropriate. The inherently non-linear valve would compensate for ΔP_L and result in an installed linear flow characteristic.

Valve Coefficient, C_v

The valve coefficient, C_v, is defined as the flow in gpm that flows through a valve with a pressure drop of 1psi across the valve (ΔP_v = 1*psi*). Cv is an important parameter that comes up in other modeling equations. It is specific to the valve you are using.

$$C_v = \frac{29.9d^2}{\sqrt{K}}$$

- d = internal diameter of the pipe in inches
- K = resistance coefficient
- K is specific to the pipe shape, diameter and material. Table of typical K values

Pressure Drop

The pressure drop in the pipe line (pressure drop due to the pipe line and any other equipment in series with the valve), ΔP_L, is defined as:

$$\Delta P_L = k_L \times sg \times f^2$$

- *f* = flow through the pipe in gallons per minute [gpm]
- k_L = [$\frac{psi}{gpm^2}$] = constant friction coefficient for the pipe and any equipment in series with the valve
- *sg* = specific gravity of the liquid

The pressure drop across the valve is defined as:

$$\Delta P_v = sg\frac{f^2}{(C_v)^2}$$

So, the total pressure drop is described by the equation:

$$\Delta P_o = \Delta P_v + \Delta P_L = \left(\frac{1}{C_v^2} + k_L\right) sgf^2$$

If the line pressure drop is negligble (constant pressure in the pipe line) then ΔP_L = 0 and ΔP_o = ΔP_v. When ΔP_L = 0 a valve with a linear flow characteristic will be desirable. When $\Delta P_L \neq 0$ and of significant value, a valve with flow characteristics closer to an equal percentage or quick opening valve will be more desirable.

Control Valve Gain

The gain of a control valve (K_L) is defined as the steady-state change in output (flow through a valve, *f*) divided by the change in input (controller signal, *m*). The flow through a valve, *f*, can have units of gallons per minute (gpm), pounds per hour (lb/hr) or standard cubic feet per hour (scfh). The controller signal, *m*, usually has units of percent of controller output (%CO). The basic relationship for control valve gain is shown below.

$$K_v = \frac{df}{dm}$$

One objective when choosing a valve is to achieve "constant valve gain". The gain is a product of the dependence of valve position on controller output, the dependence of the flow on Cv, and the dependence of Cv on the valve position. The change in valve coefficient, Cv, with respect to valve position depends on the valve characteristics f(x).

For linear characteristics:

$$\frac{dC_v}{dx} = Cv_{\max}$$

For equal percentage:

$$\frac{dC_v}{dx} = (\ln R)C_v$$

Constant Pressure Drop

The dependence of flow on the Cv depends on the pressure drop, so the equation for gain is different when there is a constant pressure drop or a variable pressure drop. If the inlet and outlet pressures do not vary with flow, the gain for either liquid or gas flow in mass units is:

- %CO = percent controller output
- W = mass flow rate
- R = valve design parameter (usually between 20 and 50)

Note: the sign is positive if the valve fails closed (air-to-open) and negative if the valve fails open (air-to-close)

Variable Pressure Drop

The valve gain for variable pressure drop is more complicated. As an example, the gain for an equal percentage is

$$K_v = \pm\frac{\ln \alpha}{100}\frac{flow}{(1 + k_L C_v^2)}\frac{gpm}{\%CO}$$

k_L = constant friction coefficient for line, fittings, equipment, etc.

The flow term cancels some of the effect of the Cv term until the valve is fully opened, so this gain is less variable with valve opening. Therefore the installed characteristics are much more linear when compared to the inherent characteristics of an equal percentage valve.

Rangeability

Valve rangeability is defined as the ratio of the maximum to minimum controlable flow through the valve. Mathematically the maximum and minimum flows are taken to be the values when 95% (max) and 5% (min) of the valve is open.

Flow at 95% valve position Rangeability = ---------- Flow at 5% valve position

A smaller rangeablilty correlates to a valve that has a small range of controllable flowrates. Valves that exhibit quick opening characteristics have low rangeablilty values. Larger rangeability values correlate to valves that have a wider range of controllable flows. Linear and equal percentage valves fall into this category.

Another case to consider is when the pressure drop across the valve is independent of the flow through the valve. If this is true then the flow is proportional to C_V and the rangeability can be calculated from the valve's flow characteristics equation.

Modeling Installed Valve Characteristics

When a valve is installed in series with other pieces of equipment that produce a large pressure drop in the line compared to the pressure drop across the valve, the actual valve characteristics deviate from the inherent characteristics. At large in-line pressure drops the pressure drop, and consequently the valve coefficient, varies with flow through the valve. These changes can cause changes in the rangeabilty and distorts inherent valve characteristics.

In the following Microsoft® Excel model, the variation from inherent valve characteristics is illustrated. A number of parameters can be changed to match the flow conditions through a valve. This model simulates both linear and equal percentage valve characteristics. To more clearly demonstrate the deviation from inherent characteristics; simply change the C_{Vmax} value of the valve. Notice how the installed valve characteristics and valve rangeability change drastically.

Installed Valve Characteristcs Model

Special Considerations for the Equation describing Flow Through a Valve

Compressible Fluids

- Manufacturers such as Honeywell, DeZurik, Masoneilan and Fischer Controls have modified the flow equation to model compressible flows. The equations are derived from the equation for flow through a valve but include unit conversion factors and corrections for temperature and pressure, which affect the density of the gas. It is important to remember to account for these factors if you are working with a compressible fluid such as steam.

Accuracy

- This equation, and its modified forms, is most accurate for water, air or steam using conventional valves installed in straight pipes. If you are dealing with non-Newtonian, viscous or two phase systems the calculations will be less accurate.

Example 1: Verbal Model of a Control Valve

Problem Statement:

Verbally model a fail open control valve positioned as a safety measure on a reactor processing an exothermic reaction.

Solution

1. **Describe the Process:** In the fail-open control valve a quick opening valve opens with a failure signal. Open is its default position once the signal goes off.
2. **Identify Process Objectives and Constraints:** A fail-open control valve is a safety measure. For example, if your cooling heat exchanger fails and the reactor starts to heat up and produce excess gases, the fail-open control valve would release excess gasses before pressure builds up to produce an explosion. The size of the valve is a constraint because it limits how much fluid can escape. The valve size determines the maximum flow rate. The shape and angles of the valve are modeling constraints. Sudden and gradual contraction or enlargement of the pipe diameter inside the valve and the connecting piping, will change the resistance coefficient and therefore the maximum velocity.
3. **Identify Significant Disturbances:** Significant internal disturbances include the escalating pressure and temperature as the exothermic reaction gets out of control.
4. **Determine the Type and Location of Sensors:** A pressure sensor would be located in the tank with the control valve that would provide the signal to the fail-open control valve. To achieve redundancy, a temperature sensor would be located on the heat exchanger to signal failure of the cooling apparatus.
5. **Determine the Location of Control Valves:** A fail-open control valve (or multiple valves) would be placed on the top of the tank to allow exit of the gasses in the processing unit.
6. **Apply a Degree-of-Freedom Analysis:** The only manipulated variable is the valve coefficient. This depends on the valve's diameter and resistance coefficient K. The control objective is the maximum flow rate. The pressure drop will vary according to the failure. Therefore there is one degree of freedom.
7. **Implement Energy Management:** This doesn't apply to our confined example, but in a larger system we could install a backup cooler if this reaction were truly dangerous.
8. **Control Process Production Rate and Other Operating Parameters:** The exit flow rate can not exceed the maximum flow rate through the control valve.
9. **Handle Disturbances and Process Constraints:** If our first control valve fails to sufficiently lower the tank pressure, a signal would be sent to a second valve and depending on the reaction, a backup cooling system. A secondary cooling system would be too expensive for many cases, but if you were dealing with a nuclear reactor or something highly explosive it may be worth the investment.
10. **Check Component Balances:** Does not apply. Preventing accumulation is the point of this control valve.

11. **Apply Process Optimization:** Our manipulatable variable is choosing a valve with a specific C_v. The valve should be able to withstand extreme temperatures and high pressures. It would be a gate valve, which opens completely upon failure. For other sizing concerns refer to “Valve Sizing.”

Example 3.10.1: Modeling Installed Valve Characteristics

A new valve is being installed downstream from a water pump. The friction coefficient of the pump and associated piping that will be in series with the new valve is

$$k_L = 1.4 \times 10^{-4} \left(\frac{psi}{gpm^2} \right).$$

The flow through the line from the pump is 300 gpm. The desired pressure drop across the valve is 4 psi. A high level of control is desired for the flow through the new valve. Two valves are being considered, one has an inherent linear characteristic, the other is equal percentage (α=50). From the manufacturer’s literature, both have a C_{Vmax} value of 200. Use the Installed Valve Characteristics Model to determine which valve has a higher range of controllable flows values.

Solution

Note that the pressure drop across the pipe is 13.5psi, which is significantly larger than the pressure drop across the valve (4 psi). These conditions indicate that the characteristic flow through the valves may not match the inherent characteristics. This is verified by the plots and also by the calculated rangeability values shown in the valve model spreadsheet. The equal percentage valve has a higher rangeabilty value, corresponding to a higher range of controllable flows.

References

- Bequette, B. Wayne. *Process Control Modeling, Design, and Simulation*, Upper Saddle River, New Jersey: Prentice Hall.
- Crane Co. *Flow of Fluids Through Valves, Fittings, and Pipe*, Joliet, IL: CRANE.
- "Friction Losses in Pipe Fittings"(PDF), Western Dynamics, LLC., retrieved September 11, 2006.
- Perry, R. H., and D. Green (ed). *Perry's Chemical Engineering Handbook*, 7th ed. New York: McGraw-Hill.
- Seborg, Dale E., Thomas F. Edgar, Duncan A Mellichamp. *Process Dynamics and Control*, New York: John Wiley & Sons.
- Smith, Carlos A., Armando B. Corripio. *Principles and Practice of Automatic Process Control*, 3rd ed. New York: John Wiley & Sons.
- "Valve Sizing and Selection." The Chemical Engineers' Resource Page. 1442 Goswick Ridge Road, Midlothian, VA 23114. retrieved Sept 24, 2006.

This page titled 3.10: Valves - Modeling Dynamics is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

4: PIPING AND INSTRUMENTATION DIAGRAMS

4.1: PandID General Information

4.2: Piping and Instrumentation Diagram Standard Notation

4.3: Piping and Instrumentation Diagrams - Location of Controls and Standard Control Structures

4.4: Piping and Instrumentation Diagrams - Standard Pitfalls

4.5: Safety Features in Piping and Instrumentation Diagrams

4.6: Regulatory Agencies and Compliance

This page titled 4: Piping and Instrumentation Diagrams is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

4.1: PandID General Information

Piping and Instrumentation Diagrams (P&IDs) use specific symbols to show the connectivity of equipment, sensors, and valves in a control system. The following sections will outline general information about P&IDs that is necessary to to know before trying to draw one.

P&ID vs. PFD

P&IDs may often be confused with PFDs, or process flow diagram. P&IDs and PFDs generally utilize the same notation for equipment. However, they serve different purposes and provide different information. The purpose of a PFD is to show exactly what a process does during operation, and a P&ID shows all controllers, valve types and the materials that are used in construction. A PFD shows the connectivity and relationships between the major equipment and materials in a process. However, it also includes tabulated design values such as normal, minimum, and maximum operating conditions that a P&ID does not contain. A PFD does not include minor piping systems or other minor components that a P&ID typically includes. The difference between P&IDs and PFDs is that P&IDs typically include more information regarding piping and safety relief valves than process flow diagrams. P&IDs do not contain operating specifications that PFDs contain, such as stream flows and compositions. It is important to note that differences between PFDs and P&IDs will vary between institutions. Most corporations maintain designated standards to create and modify the documents. Both PFDs and P&IDs are controlled documents and need to be maintained with a document control procedure. To see an industrial example of PFD vs P&ID refer to, PFD/PID industry example Wiki Page.

Information Incorporated in P&IDs

The following information is given on a P&ID that is not explicit on a PFD:

- ALL valves and valve types
- Controllers present
- Controller architectures
- Pipe diameters, materials of construction, and insulation properties (including minor piping systems)
- Equipment materials of construction

Uses of P&IDs

- Develop operational methodology
- Develop safety philosophy and safeguards
- Develop control philosophy
- Serve as a basis for control programming
- Serve as a communication document for how the process works
- Serve as a basis for equipment design, pipe design, estimating cost, purchasing
- Use for evaluation of construction process
- Train employees
- Serve as a conceptual layout of a chemical plant
- Provide a common language for discussing plant operations

Characteristics of P&IDs

- Grouped by specific section of plant
- Show the connections between all the sensors and actuators
- A general schematic - NOT a layout and NOT to scale. It should be noted that P&IDs do not specifically imply the following: same elevation of equipment, relative sizes, where valves are located on the equipment, how close equipment is to each other, and impeller types/location. They are also not the same as control or incidence diagrams. This type of information can be seen in either a plant layout drawing (can be either satellite view, showing distance between units, or a slice of building, showing height of units) or construction drawings, such as plant blueprints.

- Must be clear and uncluttered
- Must be systematic and uniform. P&IDs are used extensively in industry to document plant information. These documents need to be easily read by anyone working within the company, and easily explained to anyone else. OSHA audits can occur anytime and it is imperative that operational information can be provided to the auditor when requested. Without standard notation, it would be very difficult to go from plant to plant within your company and understand the P&IDs.
- Are generally highly confidential and have restricted access, as they can be used to replicate a process

What A P&ID Is Not

- Not an architectural diagram of a process (should show the flow of material across the plant floor between sensors and actuators, not necessarily corresponding to a 3D location)
- Does not need to be drawn perfectly to scale
- Does not imply any relative elevations
- Do not need manual switches
- No actual temperature, pressure, or flow data
- Leave out any extensive explanations

What A P&ID should include

- Instrumentation and designations
- Mechanical equipment with names and numbers, and their specifications such as material, insulation, maximum flow rate, working pressure and temperature, maximum power etc.
- All valves and their identifications
- Process piping, sizes and identification
- Miscellaneous - vents, drains, special fittings, sampling lines, reducers, increasers and swagers
- Permanent start-up and flush lines
- Flow directions
- Interconnections references
- Control inputs and outputs
- Interfaces for class changes
- Vendor and contractor interfaces
- Identification of components and subsystems delivered by others
- Intended physical sequence of the equipment

P&ID Revisions

- Revisions should be clearly identified
- Regularly issued to all related employees at each significant change to the process, as well as at benchmark points
- If small changes are made that don't warrant a completely new revision, "red pencil" additions are generally accepted between issues
- 15-20 revisions are typical during process design
- All revisions need to be communicated to EVERYONE so that only the latest revision is used. This is critical in order to avoid serious (not to mention expensive) construction mistakes. (Typically outdated P&ID is discarded to avoid confusion)

How do you generate a P&ID?

P&IDs can be created by hand or computer. Common programs, for both PC and Mac, that create P&IDs include Microsoft Visio (PC) and OmniGraffle (Mac). For creating a P&ID by hand or on a computer, please refer to the P&ID Standard Notation Wiki Page for standard equipment notation as well as computer templates.

This page titled 4.1: PandID General Information is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

4.2: Piping and Instrumentation Diagram Standard Notation

Authors: Halley Crast, Andrew Laskowski, Maurice Telesford, Emily Yatc

By: Stephanus Oscar, Kaitlin Harrington , Suhendra Lie

Introduction

Piping and Instrumentation Diagrams (P&IDs) use specific symbols to show the connectivity of equipment, sensors, and valves in a control system. These symbols can represent actuators, sensors, and controllers and may be apparent in most, if not all, system diagrams. P&IDs provide more detail than a process flow diagram with the exception of the parameters, i.e. temperature, pressure, and flow values. "Process equipment, valves, instruments and pipe lines are tagged with unique identification codes, set up according to their size, material fluid contents, method of connection (screwed, flanged, etc.) and the status (Valves - Normally Closed, Normally Open)."[1] These two diagrams can be used to connect the parameters with the control system to develop a complete working process. The standard notation, varying from letters to figures, is important for engineers to understand because it a common language used for discussing plants in the industrial world.

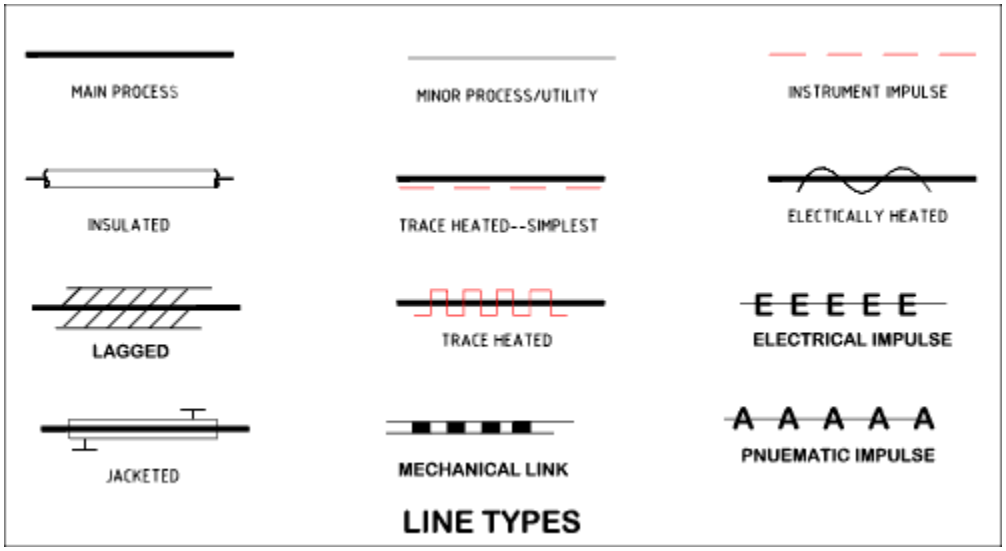
P&IDs can be created by hand or computer. Common programs, for both PC and Mac, that create P&IDs include Microsoft Visio (PC) and OmniGraffle (Mac). As with other P&IDs, these programs do not show the actual size and position of the equipment, sensors and valves, but rather provide a relative positions. These programs are beneficial to produce clean and neat P&IDs that can be stored and viewed electronically. See below for P&ID templates for these programs.

This section covers four main types of nomenclature. The first section describes the use of lines to describe process connectivity. The second section describes letters used to identify control devices in a process. The third section describes actuators, which are devices that directly control the process. The final section describes the sensors/transmitters that measure parameters in a system.

Line Symbols

Line symbols are used to describe connectivity between different units in a controlled system. The table describes the most common lines.

Table 1: Line Symbols



In Table 1, the "main process" refers to a pipe carrying a chemical. "Insulated" is straightforward, showing that the pipe has

insulation. "Trace heated" shows that the pipe has wiring wrapped around it to keep the contents heated. "Lagged" indicates on a P&ID that the pipe is wrapped in a cloth or fiberglass wrap as an alternative to painting to improve the appearance of the pipe see here for more information. The last column in Table 1 shows pipes that are controlled by a controller. "Electrical impulse" shows that the manner in which information is sent from the controller to the the pipe is by an electrical signal, whereas "pneumatic impulse" indicates information sent by a gas.

In addition to line symbols, there are also line labels that are short codes that convey further properties of that line. These short codes consist of: diameter of pipe, service, material, and insulation. The diameter of the pipe is presented in inches. The service is what is being carried in the pipe, and is usually the major component in the stream. The material tells you what the that section of pipe is made out of. Examples are CS for carbon steel or SS for stainless steel. Finally a 'Y' designates a line with insulation and an 'N' designates one without it. Examples of line short codes on a P&ID are found below in Figure A.

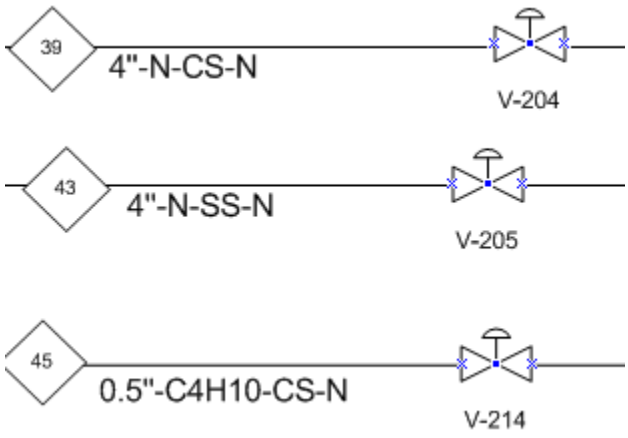


Figure A: Line Labels

This is useful for providing you more practical information on a given pipe segment.

For example in stream 39 in Figure A, the pipe has a 4" diameter, services/carries the chemical denoted 'N', is made of carbon steel, and has no insulation.

Identification Letters

The following letters are used to describe the control devices involved in a process. Each device is labeled with two letters. The first letter describes the parameter the device is intended to control. The second letter describes the type of control device.

Table 2: First Identification Letter

First Letter	Parameter Controlled
A	Analysis
C	Conductivity
D	Density
E	Voltage
F	Flow rate
I	Current
L	Level
M	Moisture (Humidity)
P	Pressure or Vacuum
T	Temperature
V	Viscosity

Table 3: Second Identification Letter

Second Letter	Type of Control Device
A	Alarm
C	Control
I	Indicate
T	Transmit
V	Valve

For example, the symbol “PI,” is a “pressure indicator.”

Valve Symbols

The following symbols are used to represent valves and valve actuators in a chemical engineering process. Actuators are the mechanisms that activate process control equipment.

Table 4: Valve Symbols

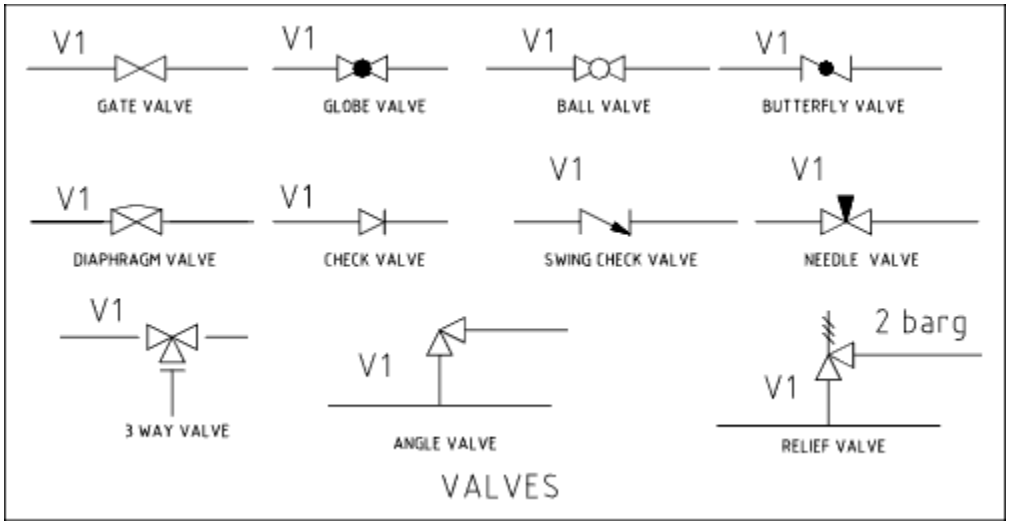
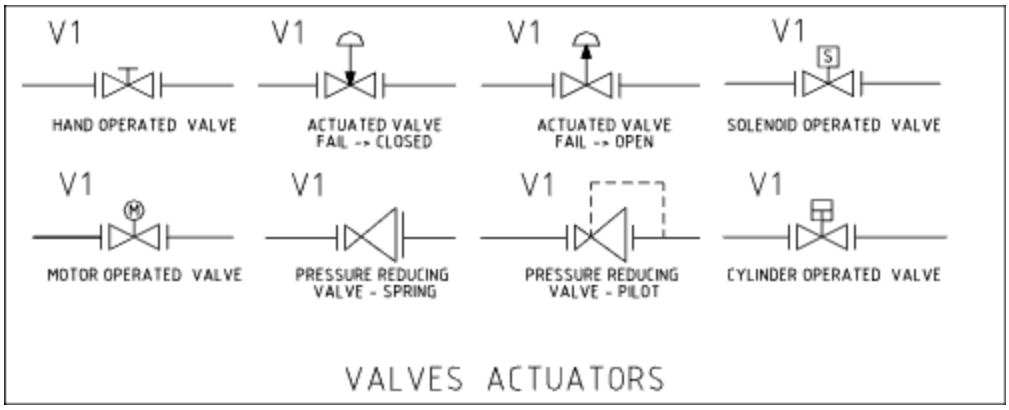


Table 5: Valve Actuator Symbols



General Instrument or Function Symbols

Instruments can have various locations, accessibilities, and functionalities in the field for certain processes. It is important to describe this clearly in a P&ID. Below is a table of these symbols commonly used in P&IDs.

General instrument or function symbols

	Primary location accessible to operator	Field mounted	Auxiliary location accessible to operator
Discrete instruments	1	2	3
Shared display, shared control	4	5	6
Computer function	7	8	9
Programmable logic control	10	11	12

1. Symbol size may vary according to the user's needs and the type of document.
2. Abbreviations of the user's choice may be used when necessary to specify location.
3. Inaccessible (behind the panel) devices may be depicted using the same symbol but with a dashed horizontal bar.

Source: Control Engineering with data from ISA S5.1 standard

Discrete instruments are instruments separate or detached from other instruments in a process. Shared display, shared control instruments share functions with other instruments. Instruments that are controlled by computers are under the "computer function" category. Instruments that compute, relay, or convert information from data gathered from other instruments are under the "Programmable logic control" section.

For example, a discrete instrument for a certain process measures the flow through a pipe. The discrete instrument, a flow transmitter, transmits the flow to a shared display shared control instrument that indicates the flow to the operator. A computer function instrument would tell the valve to close or open depending on the flow. An instrument under the "Programmable logic control" category would control the valve in the field if it was pneumatically controlled, for instance. The instrument would gather information from discrete instruments measuring the position of the actuator on the valve, and would then adjust the valve accordingly.

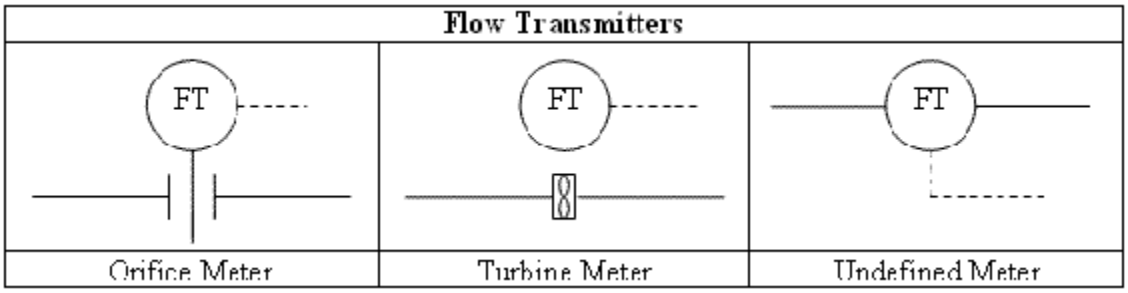
In the chart above, it is necessary to know where the instrument is located and its function in order to draw it correctly on a P&ID. A primary instrument is an instrument that functions by itself and doesn't depend on another instrument. A field mounted instrument is an instrument that is physically in the field, or the plant. Field mounted instruments are not accessible to an operator in a control room. An auxiliary instrument is an instrument that aids another primary or auxiliary instrument. Primary and auxiliary instruments are accessible to operators in a control room.

Transmitter Symbols

Transmitters play an important role in P&IDs by allowing the control objectives to be accomplished in a process. The following are commonly used symbols to represent transmitters.

Below are three examples of flow transmitters. The first is using an orifice meter, the second is using a turbine meter, and the third is using an undefined type of meter.

Table 6: Transmitter Symbols

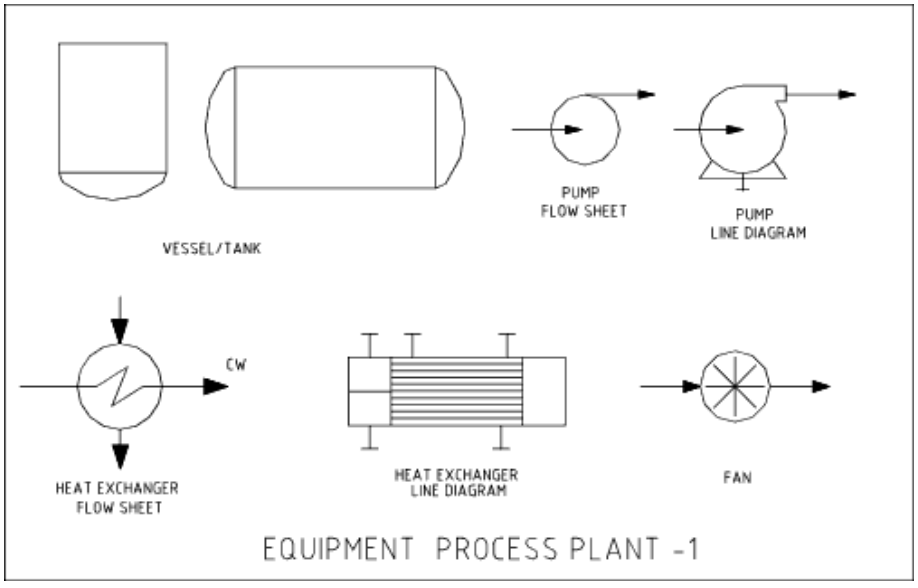


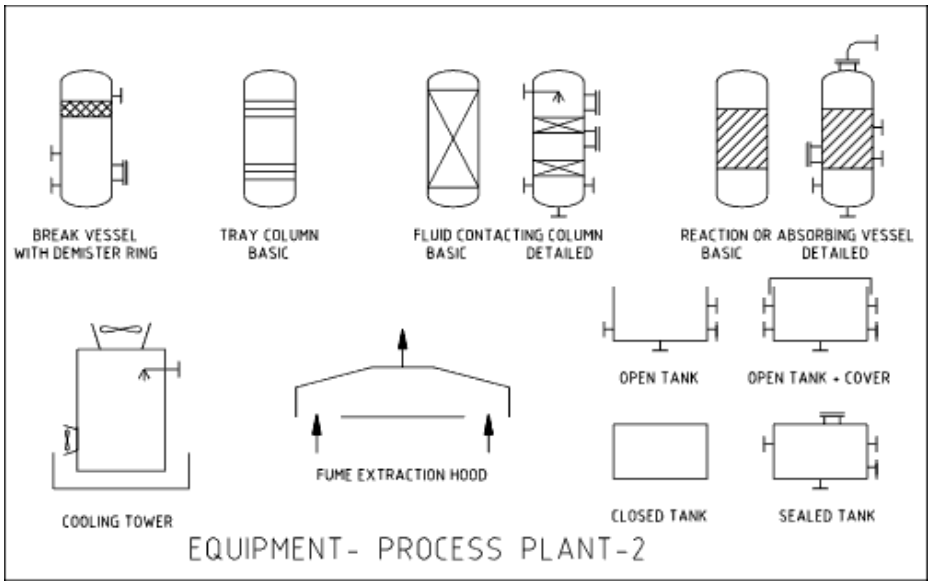
The location of the transmitter depends on the application. The level transmitter in a storage tank is a good example. For instance, if a company is interested in when a tank is full, it would be important for the level transmitter to be placed at the top of the tank rather than the middle. If the transmitter was misplaced in the middle because a P&ID was misinterpreted then the tank would not be properly filled. If it is necessary for the transmitter to be in a specific location, then it will be clearly labeled.

Miscellaneous Symbols

The following symbols are used to represent other miscellaneous pieces of process and piping equipment.

Table 7: Process Equipment





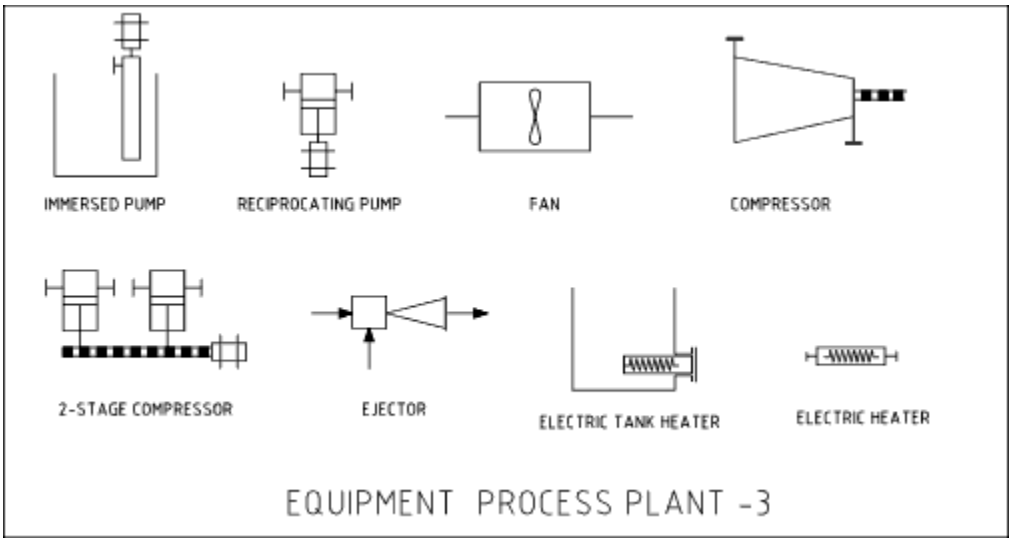


Table 8: Line Fittings

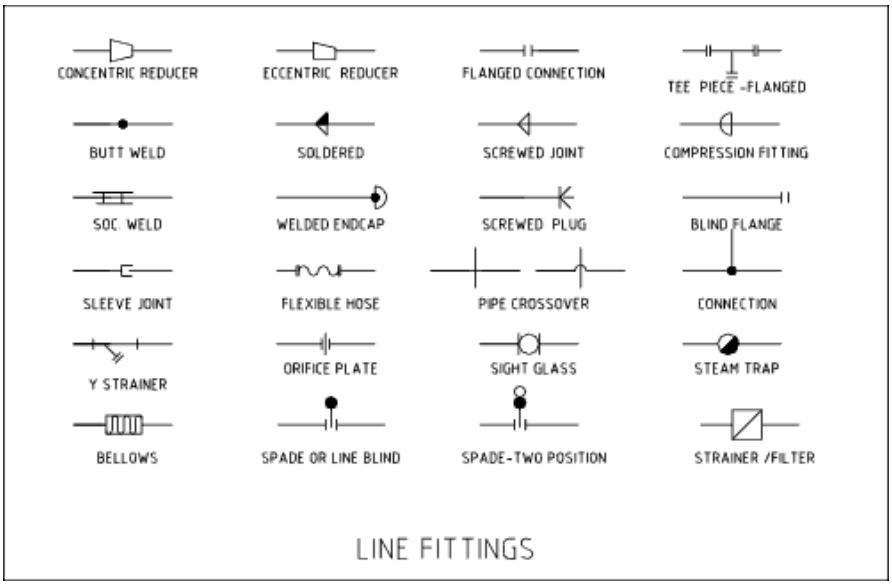
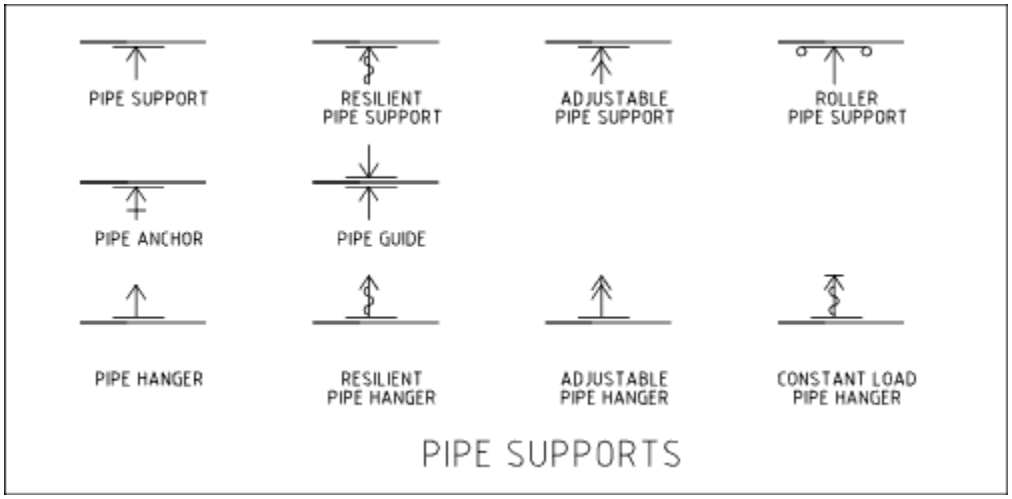


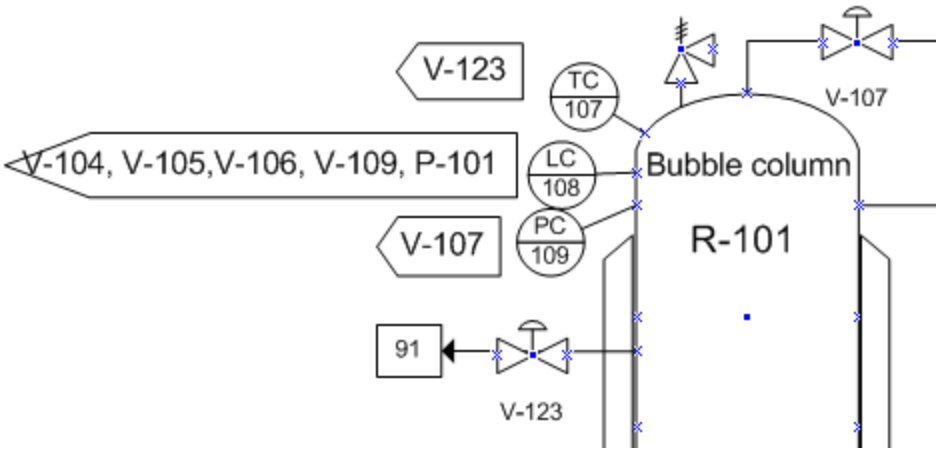
Table 9: Pipe Supports



Crafting a P&ID

In order to greatly simplify P&ID diagrams for the purposes of this class, a standard convention must be employed. This convention simplifies the many control devices that need to be used. For the sake of brevity, sensors, transmitters, indicators, and controllers will all be labeled on a P&ID as a controller. The type of controller specified (i.e. temperature or level) will depend on the variable one wished to control and not on the action needed to control it.

For instance, consider if one must control the temperature of fluid leaving a heat exchanger by changing the flow rate of cooling water. The actual variable to be controlled in this case is temperature, and the action taken to control this variable is changing a flow rate. In this case, a temperature controller will be represented schematically on the P&ID, not a flow controller. Adding this temperature controller to the P&ID also assumes that there is a temperature sensor, transmitter, and indicator also included in the process.



As you can see on the P&ID above, these controllers are represented as circles. Furthermore, each controller is defined by what it controls, which is listed within arrow boxes next to each controller. This simplifies the P&ID by allowing everyone the ability to interpret what each controller affects. Such P&IDs can be constructed in Microsoft Office Visio.

Sample Diagram

Below is a sample P&ID Diagram that is actually used in an industrial application. It is clearly more complicated than what has been detailed above, however, the symbols used throughout remain the same.

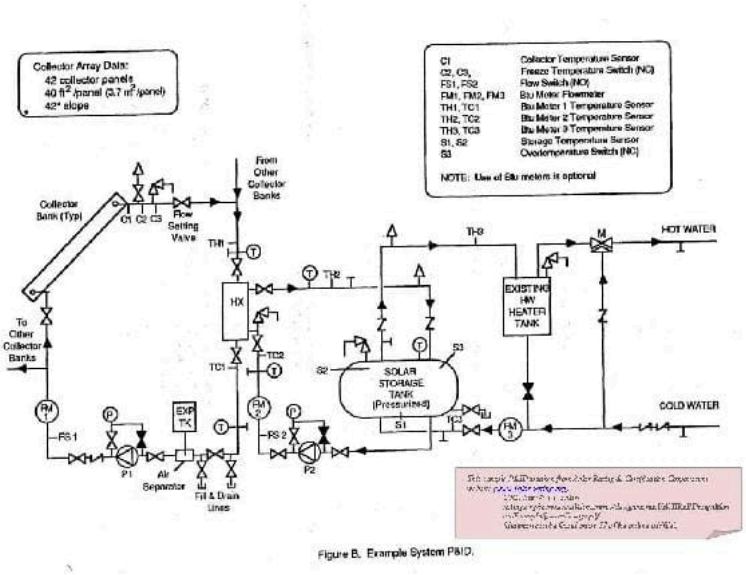
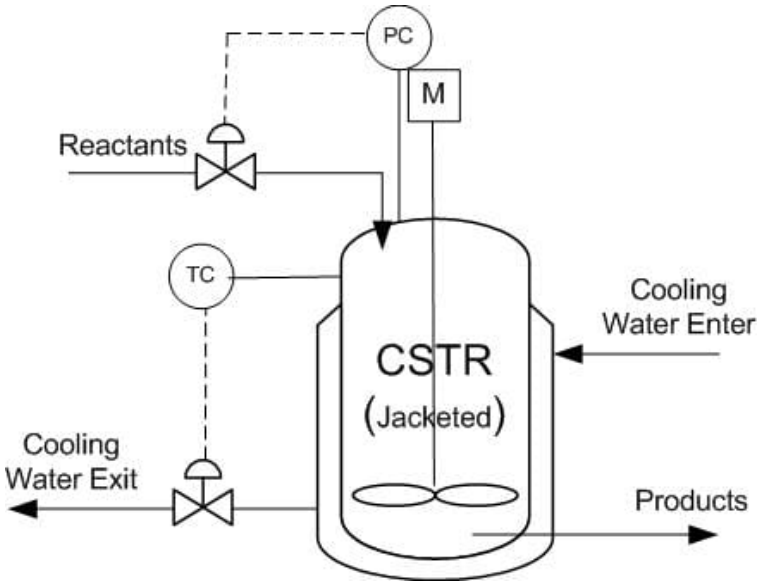


Table 10: Sample P&ID Diagram

Example 1

Describe the following controlled process in words:

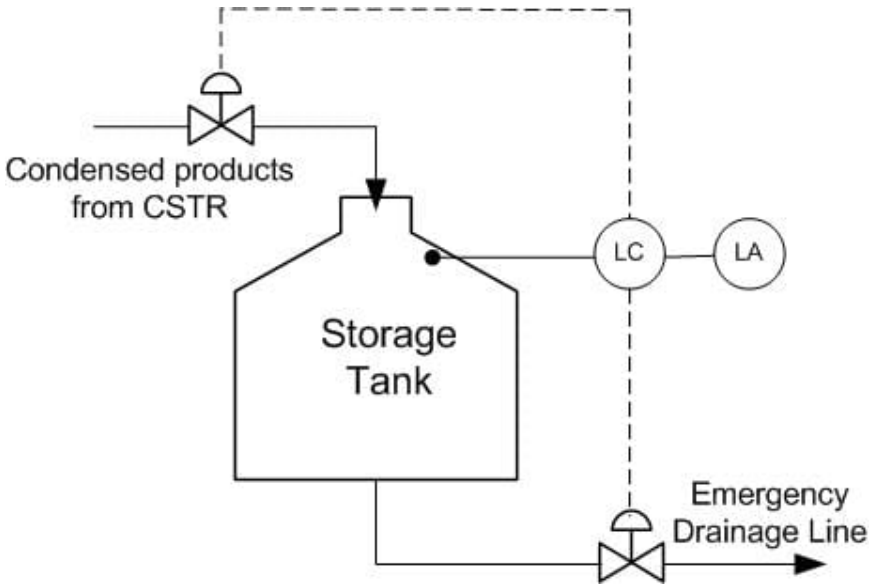


Answer: Reactants enter a jacketed CSTR where a reaction takes place and the products exit. The reactor is cooled via a coolant water stream. The temperature inside the reactor vessel is monitored with a temperature controller (also contained in the controller is a sensor, indicator, and transmitter), which electrically controls a valve. The valve can alter the flowrate of the coolant water stream, thereby controlling the temperature inside the reactor. A pressure controller is also present which feeds back to an inlet valve. Therefore, we can deduce that this reaction is most likely gas phase and if the CSTR becomes too full (high pressure) the inlet valve will close.

Example 2

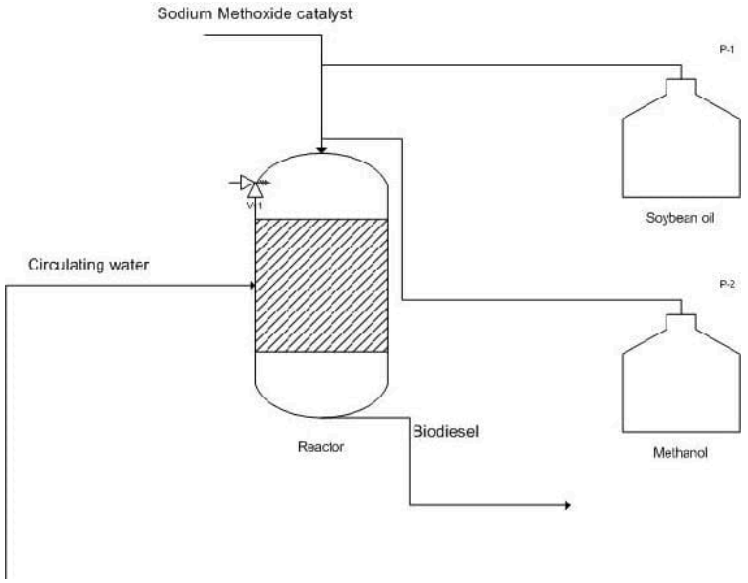
Draw a proper P&ID diagram of the following process:

A storage tank is filled with condensed products formed via the CSTR in Example 1. The tank contains a level controller at a set point on the top of the tank. If this tank were to fill, materials would get clogged up in the reactor. Therefore, if the tank reaches 90% of its total capacity, the level controller will send an electric signal, which opens an emergency drainage line located at the bottom of the tank. The level controller will also activate an alarm alerting plant engineers that there is a problem with the storage tank. Finally, the level controller will also close the inlet valve to the storage tank.

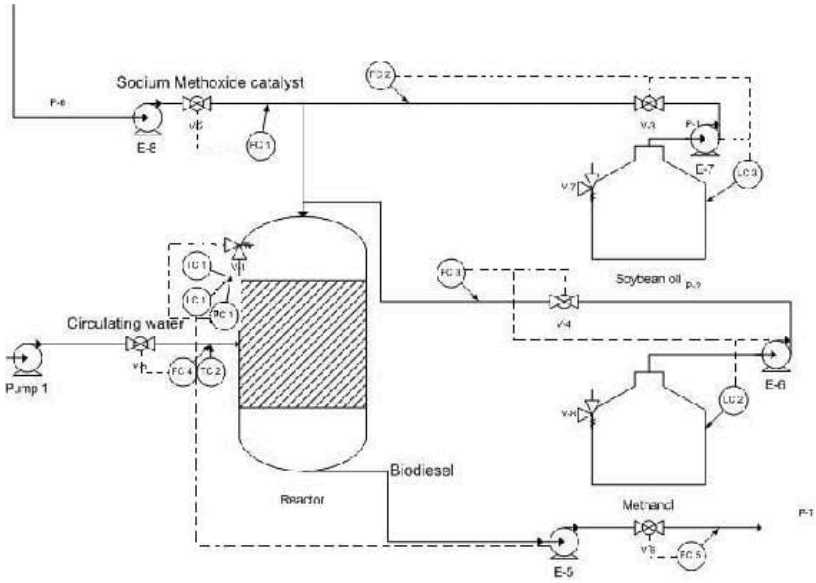


Example 3

Below is a P&ID diagram of the transesterification process to produce biodiesel. Soybean oil, methanol, and the sodium methoxide catalyst are pumped in to the reactor. The temperature of the reactor is regulated by the circulation water. The resulting biodiesel is then pumped out of the reactor and goes on to other processes so that it can be sold. Below is a P&ID of the process that is missing the valves, pumps, and sensors. Add the pumps, sensors, and valves that are needed to successfully control the process.

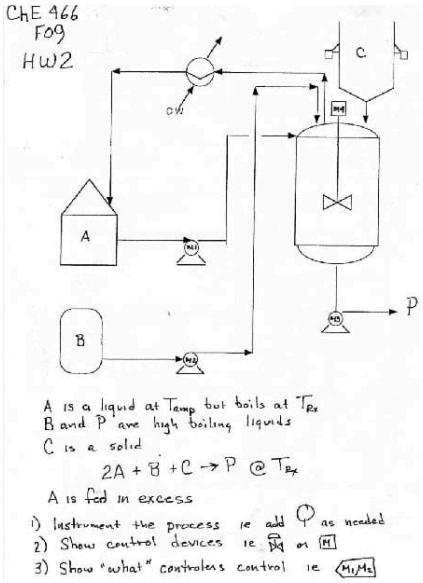


Solution:

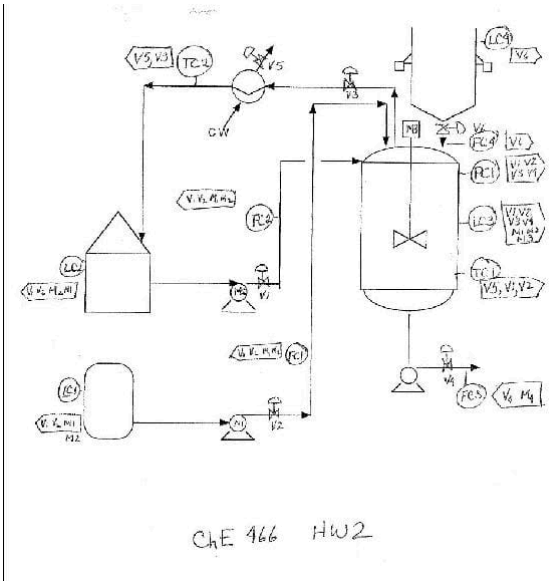


Example 4

Below is a example problem of a typical P&ID problem. A is a liquid at Tamp but boils at Trx. B and P are high boiling point liquids and C is a solid. The reaction for the process is 2A+B+C-->P at Trx. Ais fed in excess.



Below is the solution to the problem above.



Sage's Corner

A Brief Guide To Piping and Instrumentation Diagram Standard Notation

video.google.com/googleplayer...40157432698049

slides for this talk

Additional Reading

1. Wikipedia ChemE P&ID symbols entry
2. Wikipedia P&ID entry
3. Control Engineering Online Article: how to read P&IDs

Here are templates for creating P&IDs in Visio (PCs) and OmniGraffle (Mac).

References

- Mineral Engineering Technical Services; "Process Engineering.". Last Updated 05/09/2007.
- Solar Rating and Certification Corporation; "Example System Design.". Last Updated 05/09/2007.
- Karim, Nazmul M.; Riggs, James B. "Chemical and Bio-Process Control." 3rd Edition. Ferret Publishing.
- Ogunnaike, Babatunde A.; Ray, W. Harmon. "Process Dynamics, Modeling, and Control." New York Oxford: Oxford University Press, 1994.
- Beardmore, Roy; "Flow Sheet Symbols.". Last Updated 12/02/2006.

This page titled 4.2: Piping and Instrumentation Diagram Standard Notation is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

4.3: Piping and Instrumentation Diagrams - Location of Controls and Standard Control Structures

Introduction

A Piping & Instrumentation Diagram (P&ID) is a schematic layout of a plant that displays the units to be used, the pipes connecting these units, and the sensors and control valves. Standard structures located on a P&ID include storage tanks, surge tanks, pumps, heat exchangers, reactors, and distillation columns. The latter three comprise most of the chemical process industry single-unit control problems.

P&IDs have a number of important uses in the design and successful operation of chemical process plants. Once a process flow diagram is made, P&IDs help engineers develop control strategies that ensure production targets are met while meeting all safety and environmental standards. These diagrams are also very important for locating valves and process components during maintenance and troubleshooting.

Each page of a P&ID should be easy to read and correspond to a specific action of the plant. Also, symbols used in the P&ID are uniform throughout. Diagrams for heat exchangers, continuously stirred tank reactors (CSTRs), and distillation columns shown in subsequent sections of this wiki represent a small sample of standard symbols for these structures. For a list and explanation of standard controls-related symbols, please refer to PID Standard Notation.

General Strategies for Implementation of Good Control Systems

There are several guidelines to follow when designing P&IDs for a plant. The first consideration is stable operation of the plant so that all safety and environmental standards are met. Maintaining product quality should also be a primary design objective. Additionally, systems should be designed to respond quickly to rapid changes in rate and product quality. Usually control systems should be run in an automatic mode so the system will correct itself, as opposed to a manual mode, which requires operator supervision.

When setting up a control system, it is useful to first focus on mass balance control, which can be monitored by level and pressure control loops that use gas or liquid flow rates. Additionally, a product control structure should be set up to ensure efficient process operation. To prevent controls from conflicting with one another, a final examination of the controlled streams should be performed. Further instrumentation can then be installed to prevent conflicts between controllers. Finally, appropriate tolerances should be established for controllers that directly affect the action of other controllers. Determination of the optimal control placement is essential for successful plant operation.

Standard Structures and Location of Control Features

As mentioned above, control of heat exchangers, reactors, and distillation columns represent the majority of single-unit control problems. The table below summarizes the control schemes derived in the following examples. The most common control schemes are listed in the table below.

Table 1: Summary of common P&ID structures

Structure	Control Variable	Manipulated Variable	Location of Control Valve	Comments	Typical Disturbances
Steam heated Heat Exchanger	Temperature of outgoing stream	Steam pressure	Steam line or condensate stream	Placement of a control valve on the steam line gives tighter control, but this type of valve is bigger and more expensive than one placed on a condensate line	Fluctuation of the inlet flow rate
Liquid/Liquid Heat exchanger	Temperature of outgoing stream	Bypass flow rate	Bypass stream	This allows high coolant flow rates and decreases fouling. Changes in bypass stream's flow rate result in temperature control loop disturbances that are easily regulated	Fluctuations in inlet flow rates
Jacketed CSTR w/ Endothermic RXN	Product stream temperature	Steam flow rate or steam pressure in the jacket	Steam line	Changing steam flow rates allows direct measure of the amount of heat entering the system and therefore a measure of conversion. It also reduces reactor sensitivity to heat load changes and enthalpy changes. Manipulating jacket pressure is better at absorbing reactor heat duty and steam enthalpy upsets but it does not directly measure heat load on the reactor	Fluctuations in feed flow rate, temperature, and enthalpy of heating medium
Jacketed CSTR w/ Exothermic RXN	Inlet temperature of coolant	Coolant flow rate	Coolant feed stream	There is a slow response to fouling on heat transfer surfaces, but there is a fast response to changes in inlet coolant temperature and coolant supply pressure	Changes in feed flow rate, temperature; Fluctuation of coolant enthalpy

For most equipment, a degrees of freedom analysis is first performed and then control scheme is designed based on the degrees of freedom.

Heat Exchanger

To monitor the performance of a heat exchanger, the product stream is important. Usually, the product stream must be within some temperature range before it continues to downstream process units. The outlet temperature of this stream can be used to calculate the heat transfer. The steam is controlled in order to obtain the desired product stream temperature. One way to influence the product temperature is by controlling the flow of the heated steam. This flow-based control may take some time to implement and therefore cause fluctuations in the process. Depending on the process, these fluctuations may or may not be acceptable. Instead of monitoring the flow-rate, the steam pressure may be monitored, achieving tighter control of temperature. A change in pressure is much easier to monitor and correlates directly with a change in temperature of the steam. This offers an effective way to control process temperatures. Pressure control also enables the physical condition of the piping to be monitored, since pressure changes occur as fouling progresses. The pressure, combined with the flow-rate and temperature of the heated steam, can be used to calculate this fouling occurring inside a piece of equipment. It is important to remember that the steam should always be controlled on the inlet side of the heat exchanger for better pressure control and safety reasons, as shown in figure 1 below.

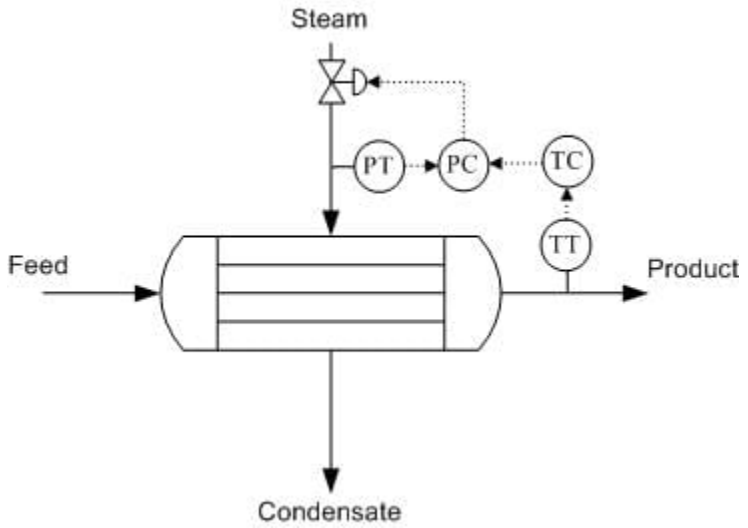


Figure 1. Heat Exchanger with pressure control on steam inlet and temperature control on the product stream.

Distillation Column

Because the economic viability of an overall process is based significantly on product purity, it is important that distillation columns maintain stable operation. Changes in composition and flow-rate of the feed stream are common disturbances in distillation column operation. Improper functioning of controllers can undermine the effectiveness of the product composition. A degrees of freedom analysis can help place sensors and actuators in appropriate places, while not including too many sensors and actuators, in order to obtain an efficient control system.

Degrees of Freedom Analysis

A simple degrees of freedom analysis can make design of a control scheme easier as well as improve the control scheme overall. A process control approach to degrees of freedom, adapted from ECOSSE Module 3.1, requires tabulation of streams and extra phases.

$DOF = Streams - ExtraPhases + 1$

A typical distillation column contains the following streams; feed (1), bottoms (2), distillate (3), reflux (4), product (5), vapor-liquid mix (6), cooling water (7), and steam (8). There are also three locations where there are two phases, vapor and liquid, present in equilibrium. These are denoted as one "extra phase", since temperature and pressure are not independent in a two-phase system.

$DOF = 8 - 3 + 1 = 6$

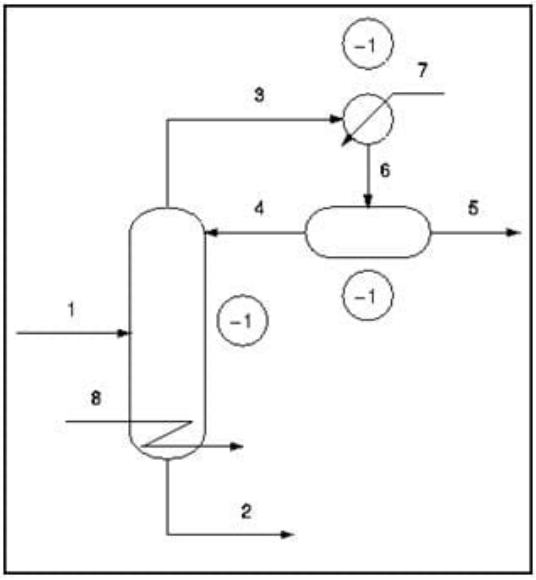


Figure 2. Degrees of freedom analysis on a typical distillation column (adapted from ECOSSE).

In typical practice, the condenser and reboiler pressure are specified as atmospheric pressure. Once the pressure is specified in the condenser and reboiler (both two-phase systems), the temperature is specified and therefore the degrees of freedom are reduced by two. The total degrees of freedom, in practice, is six.

$DOF = 8 - 5 + 1 = 4$

A simple degrees of freedom analysis, in any system, can help define the variables and reveal where the critical sensors for process control should be located. This type of analysis also minimizes the design of too many sensors, actuators, and valves, which may in theory control and regulate the same variable.

Sensor and Actuator Options

Product streams, reboiler steam, and the reflux stream should have flow sensor/transmitters so that each flow can be adjusted to meet the column's control objectives. A flow sensor should also be placed on the feed stream because disturbance to the feed flow rate are common. Each flow sensor should be connected to a flow valve to control the corresponding stream.

Differential pressure level sensor/transmitters should also be used for the accumulator (the vessel that collects condensed distillate) and bottom of the column because maintaining these levels is essential for reliable operation of the column. If flooding is an issue, a pressure differential across the column should indicate the onset of flooding. A pressure indicator should be installed at the top of the accumulator to monitor column pressure. Temperature in a distillation column is typically controlled by manipulating steam flow to the reboiler.

Control Schemes

There are a number of common control schemes for distillation columns. Optimally, a distillation column should be run with *dual composition* control because it saves energy. In dual compostion control, the temperature of both chemicals in a binary distillation is controlled. The system is more complex to setup and measurements required for control may be difficult. As a result, many distillation columns use single composition control instead. Common control configurations for distillation columns include reflux-boilup and distillate-boilup.

In the reflux-boilup configuration, the distillate composition and bottoms composition are the control variables. The reflux flow and the heat input control (vapor boilup) are the manipulated variables, which allow control of the liquid and vapor flow-rates in the column. With this control system, a quick response to changes in the feed composition is possible. In the distillate-boilup control configuration, the distillate flow and the vapor boilup are used to control composition. This configuration is a better choice for columns where the reflux ratio is high.

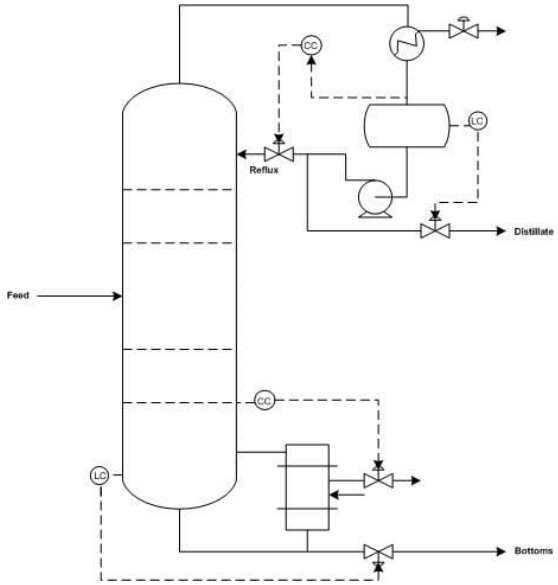


Figure 3. Distillation column with reflux-boilup control scheme (adapted from Luyben).

Reactor (Exothermic Reaction in CSTR)

In implementing controls in reactive systems, temperature is a good indicator for unit performance. Temperature is often related to reaction rate and varies with time in most reactors. However, a reactor exit stream with a constant temperature is often desired downstream in the process.

In an exothermic reaction in a CSTR, lower initial temperatures result in lower reaction rates and low heat generation. As the reaction progresses, heat generation increases rapidly due to higher reaction rates and high concentrations of reactants. As reactant concentration decreases, the heat generation once again becomes low. Exothermic reactor temperatures must be controlled to assure stable reactor operation.

To remove heat from an exothermic reaction, basic heat transfer principles are employed. A coolant is pumped through a shell outside of the reactor. Since the heat removal is linear, the temperature of the coolant should be controlled. In doing this, it is possible to increase the driving force for heat transfer to slow a reaction, or conversely, to allow the reaction to further progress by decreasing the driving force for heat transfer.

Degrees of Freedom Analysis

In order to determine where to place controls, sensors, and valves in a exothermic reaction operation, a degrees of freedom analysis, similar to that carried out in the distillation section 3.2, can be helpful.

A typical CSTR contains the following streams; reactant A (1), reactant B (2), product (3), and coolant to the jacket (4). There are no locations where there are two phases, vapor and liquid, present in equilibrium - assuming the reaction is liquid phase with no simultaneous evaporation or sublimation.

$$DOF = Streams - ExtraPhases + 1$$

$$DOF = 4 - 0 + 1 = 5$$

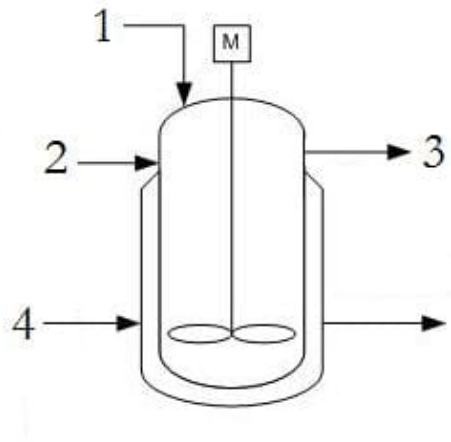


Figure 4. Degrees of freedom analysis on a typical exothermic reaction in a CSTR.

Since in most practical applications the reactants (1 and 2) as well as the product (3) flow rates are defined by demand for the product, there are only 2 degrees of freedom which allow for placement of a 2 valves or controllers on the coolant stream (4) as well as one of the reactant streams (1 or 2). Only one of the reactant streams need to be specified as the other can be determined by a ratio controller using stoichiometric coefficients.

Control Schemes

As discussed previously, a valve can be placed on the coolant inlet stream to ensure proper temperature control of the reactor. By controlling the coolant stream based on inlet conditions of the reactant streams, the control can respond quickly. But, if the coolant is controlled based on outlet conditions of the product stream, there is a lag in response, but it is easy to monitor heat transfer performance.

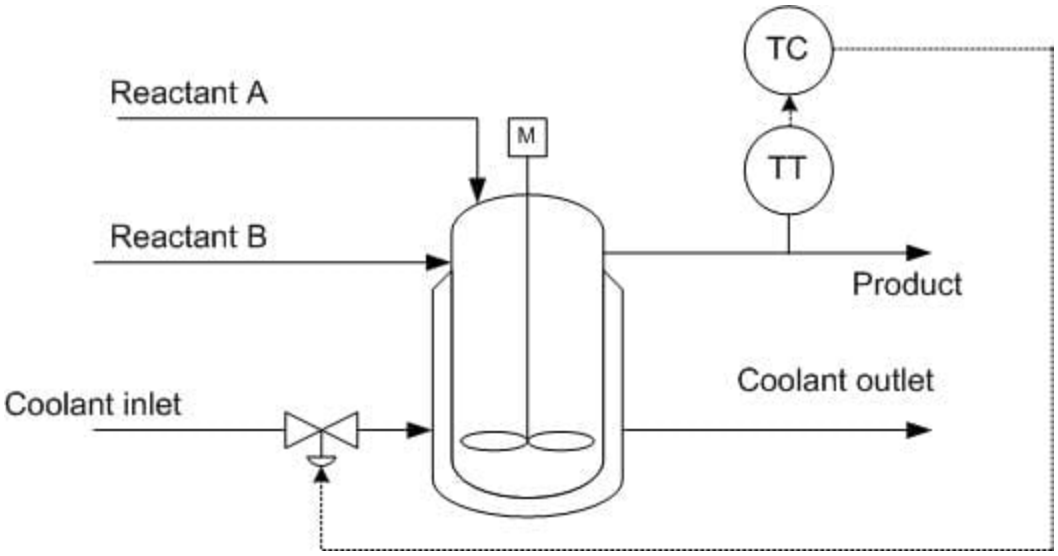


Figure 5. Jacketed CSTR with coolant control based on the outlet temperature conditions.

To translate this example to an endothermic reaction, the coolant flow would simply be translated to steam flow in order to provide heat to the reaction through the CSTR jacket.

Other Common Process Equipment

Beyond the heat exchangers, reactors, and distillation columns, many other pieces of process equipment, including furnaces, compressors, decanters, refrigerators, liquid-liquid extractors (LLEs), and evaporators, are subject to disturbances and require careful control.

Furnaces

For example, furnaces may be subject to frequent load changes as a process or customer requires more energy. To cope with these demands, the temperature of the outlet stream must be monitored and manipulated. Information from a temperature controller at the outlet stream can be used to effect changes in valves that control the flow-rate of air and fuel to the furnace. At this point, the best setup of the control system must be considered in light of factors such as safety, lag time for changes to occur, and equipment wear. For a furnace, the controls should generally be set up so that an excess of air is more likely than an excess of fuel. This reduces the risk of explosion due to the furnace filling with uncombusted fuel.

Liquid-Liquid Extractors

In liquid-liquid extractors, the interface level and pressure are the controlled variables. Disturbances in flow rate of the entering stream can affect interface level and prevent complete separation of the heavy and light components. From this, it is obvious that there should be valve controls on both exit streams. The best control scheme depends on the operation of the process. When the heavy phase is continuous (light phase flows upward through heavy phase), changes in interface level should be controlled by adjusting the flow-rate of the light product, while the pressure is controlled by adjusting the flow-rate of the heavy product out of the column. Figure A is a representation of what occurs in a single stage extractor. Generally, single stage extractors are using in chemical labs, where as multistage extractors are used in industry. A multistage extractor uses the immiscible liquid stream from the previous stage as the feed in the following stage. Figure 6 depicts the control scheme previously described.

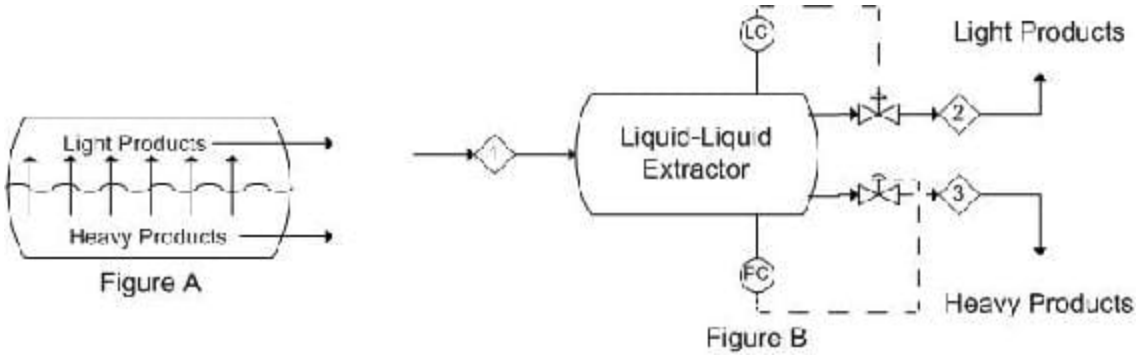


Figure 6. Liquid-liquid extractor control scheme with interface level and pressure control.

When the light phase is continuous, the control system must be set up in the opposite manner. Figure C is again a depiction of what occurs in a single stage extractor. Figure D is a representation of a control scheme that could be implemented. This is the reverse of the control scheme in Figure 6.

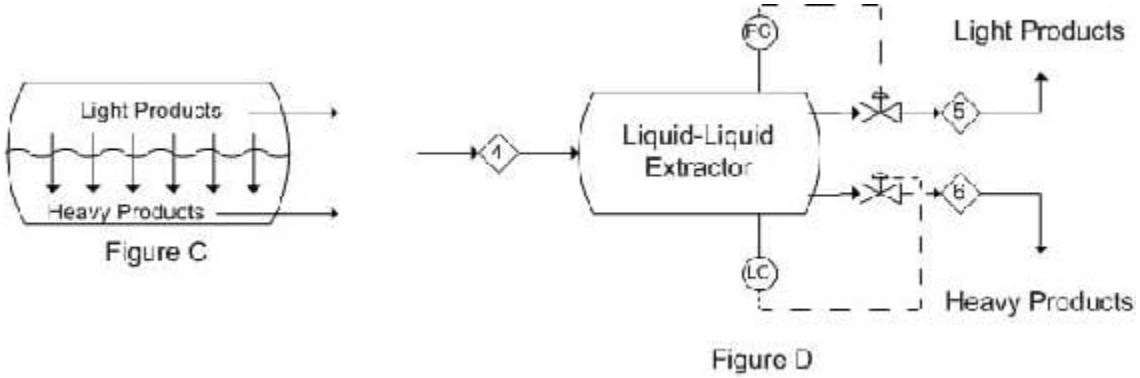


Figure 7. Liquid-liquid extractor control scheme with light-phase control.

These representations are only two possibilities for basic control schemes that can be implemented in a process as there are several controllers and aspects of the specific processes that may need to be controlled or monitored.

Compressors

Compressors are another valuable component in process design. They allow for the reduction of the volume of an incoming stream by increasing the pressure at which the stream is maintained. They can also be used to pump liquids through the process, as liquids are highly incompressible compressors cannot be used for volume reduction. For this, there must be a specific control system as to prevent adverse effects due to extremely high or low pressures. There are several types of compressors, including: dynamic, axial, and rotary [1] to name a few. Because the increase in pressure is governed by the ideal gas law, there is most often and increase in temperature as well. This can be left as is, or sent to a heat exchanger for temperature reduction. Heat exchangers were discussed above.

One such example is the use of a centrifugal compressor to reduce the volume of a fuel stream for storage. Using a compressor will allow for volume reduction as gasses are easily compressed, this is also economically friendly as it reduces the size of tank necessary to store the fuel stream. The tank should also be equipped with a pressure reducing valve, to bring the stream back to a desired pressure, depending on the process. A diagram of this scheme is as follows:

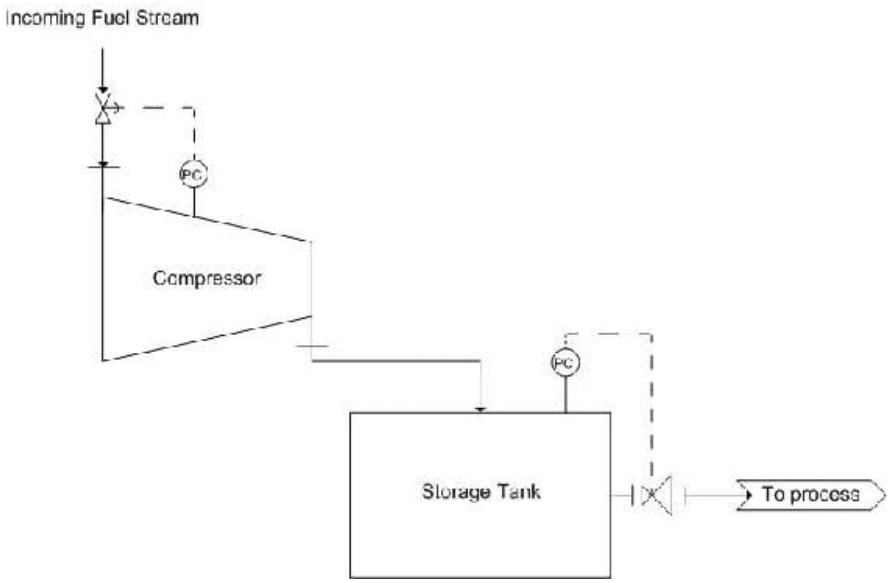


Figure 8. Compressor control scheme with pressure control.

The pressure controller on the compressor controls the valve on the incoming fuel stream. This ensures that if there is a build up in pressure, the flow into the system will be stopped in time. Also, a pressure controller should be placed on the storage tank. This is controlled by the pressure reducing valve mentioned earlier.

Decanters

Decanters, much like Liquid-liquid extractors, use solubility as their principle of separation. Unlike Liquid-liquid extraction, these require some time for the separation to occur. Generally, the separation is a liquid-liquid or liquid-solid separation. Decanters are widely used in the wine industry to separate out sediment from the wine. Utilizing the wine separation example, a possible control scheme is as follows. Here, there is only a level sensor on the decanter as it is a liquid-solid separation. Note that there is a release stream, used to remove collected sediment. An analytical or pH sensor could also be used here to maintain the correct chemistry or pH, as wine is the final product. It is also important to note that the exact placement of the level sensor would vary depending on the desired level in the tank. Here, the level sensor is not shown in its exact placement.

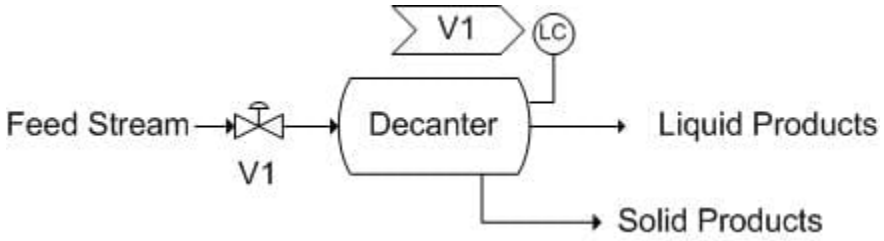


Figure 9. Decanter control scheme with level control and pH control.

Another vessel that is very similar to decanters are knockout drums. These vessels are generally located after heat exchangers or other pieces of equipment that result in a multiphase system. These vessels are used to separate the two phases, generally gas- and liquid separation. A possible control scheme is depicted below. The incoming stream is a liquid-gas mixture coming from a heat exchanger. Thus, there is a pressure sensor on the knockout drum. A level controller could be used, but this is effectively measuring the same thing, so it has been omitted. Also, because it is coming from a heat exchanger, a temperature controller has been included to control the amount of cooling taking place in the heat exchanger.

The pressure controller (PC) controls V2, the stream into the tank. The temperature controller (TC) controls V1, the valve on the coolant stream to the heat exchanger.

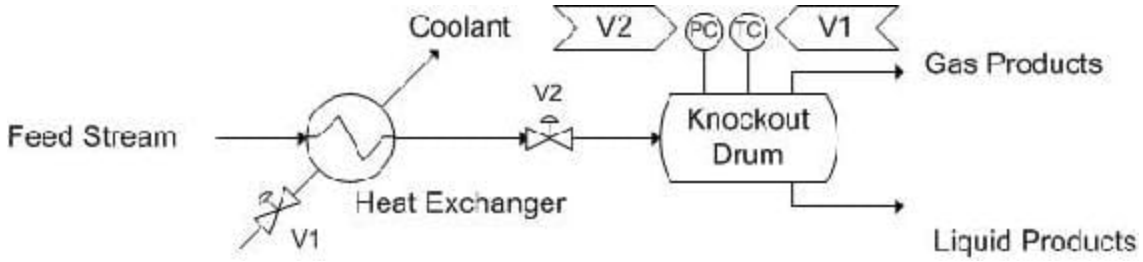


Figure 10. Knock-out drum control scheme with pressure and temperature control on different streams.

These examples illustrate the typical method of locating control systems for process equipment. After choosing the location for valves based on process constraints, there still remain a number of possibilities for the actual manner in which to control the valves. The lag-time for changes may be longer for certain configurations, while the danger of damaging equipment is smaller. The controls configuration will depend strongly on the safety concerns associated with a specific process.

Selecting Controls and their Locations for a Multi-Unit Process

The following steps should be followed when setting up controls for multi-unit processes. For more in-depth descriptions, refer to the Step-by-Step Method For Describing Controls and Their Purpose.

1. **Determine process objectives**, taking into consideration product specifications, economic constraints, environmental and safety regulations, etc.
2. **Identify boundaries for normal operation**. These can be based on equipment limitations, safety concerns, environmental regulations, and the economic objectives of the processes.
3. **Identify units and streams in the process that are susceptible to significant disturbances**. These disturbances commonly occur in feed streams, product streams, and reactor vessels, but can be present anywhere that temperature or pressure or other variables are changing.
4. **Select the types and locations of sensors** in order to properly measure and monitor critical process variables.
5. **Determine the appropriate types and locations for control valves** in order to appropriately adjust process variables so that they remain within the normal operating boundaries. Controls should be set up to minimize response time between sensing a change and taking corrective actions. The ideal location for any given control depends on the process unit or units that it affects.
6. **Perform a degree of freedom analysis**.
7. **Energy Considerations**. An energy balance should be performed for the process. This step involves transporting energy to and from process units. This may include removing heat generated by a reactor and using it elsewhere in the process. Control valves will help regulate the flow of such streams.
8. **Control Process Production Rate and Other Operating Parameters**. Adjusting process inputs, such as reactant feed rates, can alter other variables in the process. Process controls must be able to respond to these adjustments to keep the system within operating boundaries.
9. **Set up control system to handle disturbances and minimize their effects**. (See Chemical Process Controls > PID Control)
10. **Monitor Component Balances**. Accumulation of materials within a system is not desirable and can lead to inefficiency in the process or catastrophic failure.
11. **Control individual unit operations**. Each unit of a multi-unit process needs to be individually controlled in order for control of the entire system to be possible.
12. **Optimize the process**. If the system has degrees of freedom, process variables can be manipulated in order to more efficiently or economically create product.

Example 4.3.1: Completing a Control System

Consider the following reaction in a semi-batch reactor:

$$(CH_3CO)_2O + H_2O \longrightarrow 2\,CH_3COOH$$

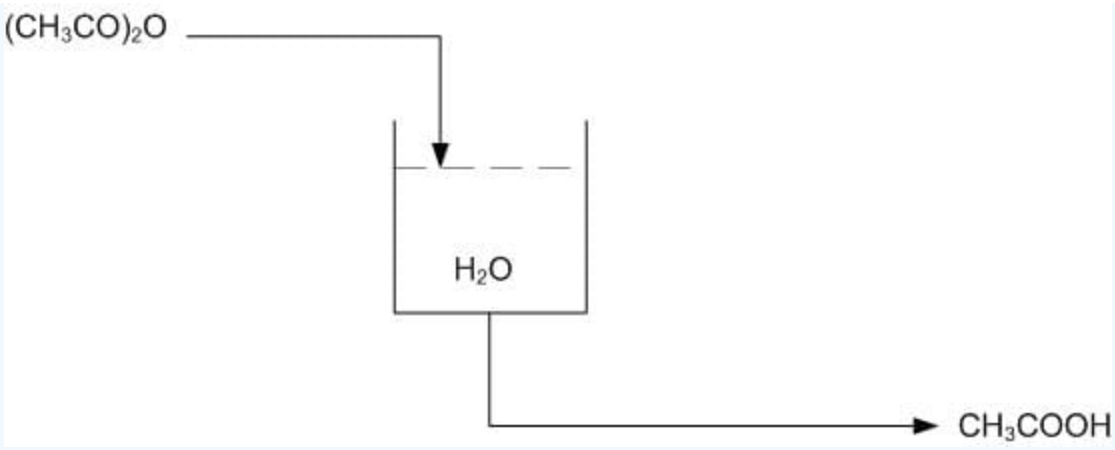


Figure E-1. Basic Flow Diagram of reaction of Acetic Anhydride with water to produce Acetic Acid.

The objective is to produce CH_3COOH at a given concentration, i.e. 15-20%. The tank should also be drained once the given concentration is achieved and the reaction has gone to completion. The tank can only hold a specified volume and 80% is filled with water. The reaction should be run in excess water. Once a given amount of acetic anhydride has entered the reactor, it should be left in the tank until the reaction has gone to completion. The completion of the reaction can be determined by temperature, as the reaction is exothermic.

Using the multi-unit controller design from above, determine the location of needed sensors and valves on the flow diagram.

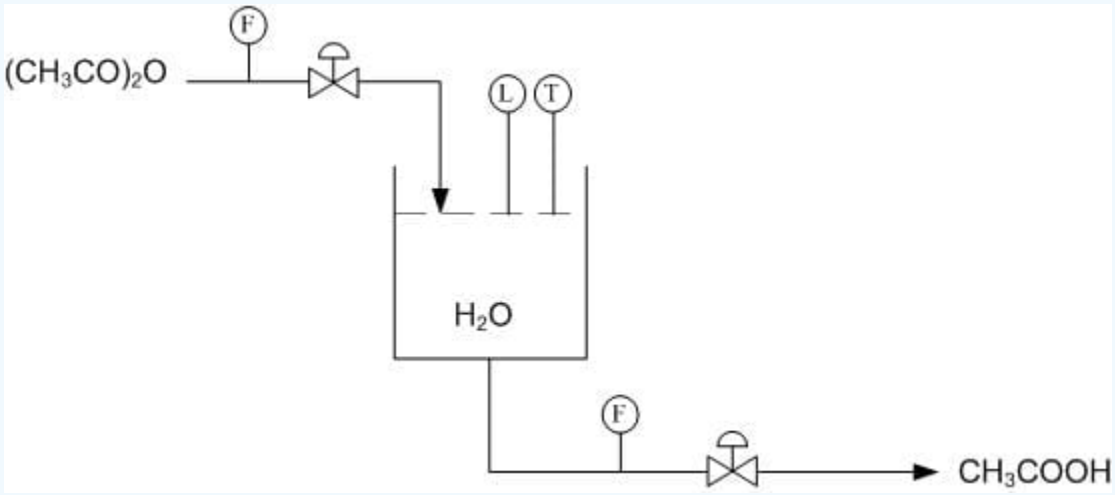


Figure E-2. Location of sensors and valves.

Using Figure E-2 determine how to use the level and temperature sensors to limit the amount of acetic anhydride entering the reactor and also how to drain the system once the reaction has gone to completion.

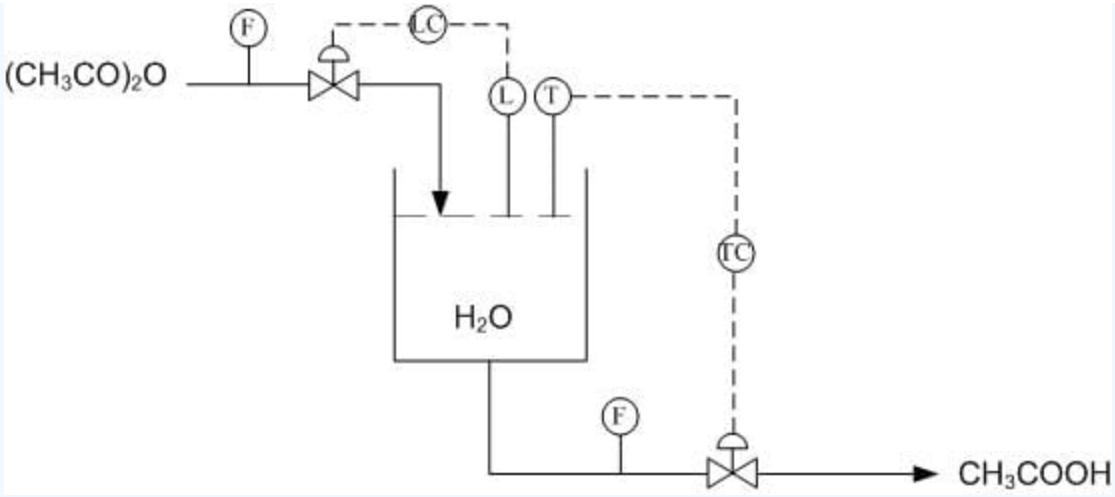


Figure E-3. Control system for production of acetic acid.

Using the above system, when the right amount of acetic anhydride has entered the reactor the level sensor will stop the flow into the reactor. Once the reaction has gone to completion and the temperature sensor has been triggered, due to a peak in temperatures from the exothermic reaction, the tank will drain.

Example 4.3.2: Control of a Heat Exchanger

Examine the figure below, which shows two different control schemes for a heat exchanger. The outlet stream temperature must remain constant. How is this goal accomplished using the control systems shown below? Which stream is the manipulated stream? What are the advantages and disadvantages of each setup?

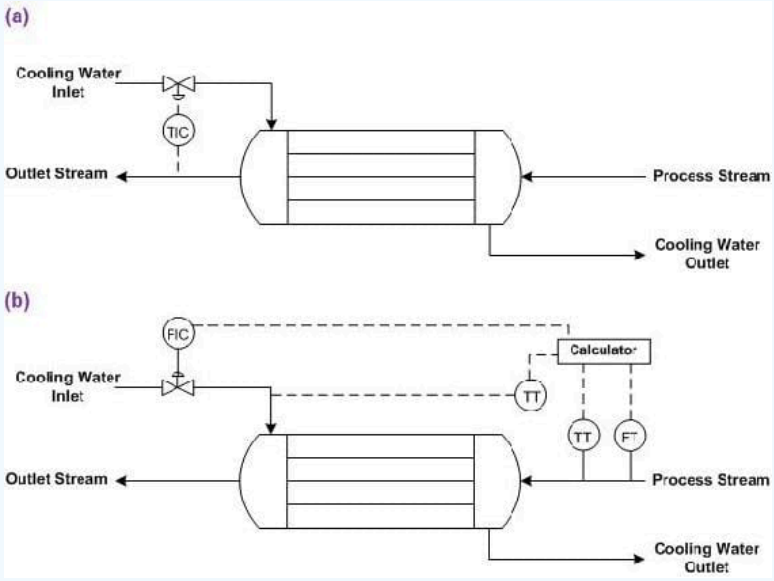


Figure E-4. Two different control strategies for a heat exchanger (adapted from Turton)

Solution

In both setups, the cooling water flow is manipulated to control the temperature of the outlet stream. The control strategy for each heat exchanger is different.

Figure E-4 (a) shows a heat exchanger where the flow of cooling water is adjusted based on the temperature of the outlet stream. The advantage to this setup is that it is quite straightforward and the cause of the change in the temperature of the outlet stream does not have to be identified. The disadvantage is that a temperature change has already occurred before action is taken, and if the lag-time for the additional cooling water to cool the process stream is long, the outlet stream temperature may be inconsistent.

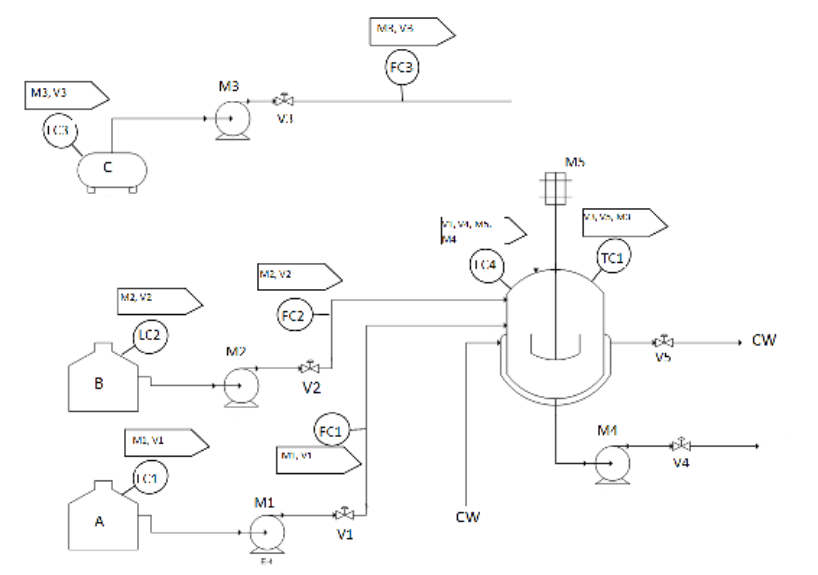
Figure E-4 (b) shows a heat exchanger were the flow of cooling water is adjusted based on the flow-rate and temperature of the inlet stream and the temperature of the coolant stream. Based on a set of calculations using the heat transfer coefficient of the heat exchanger, the outlet stream temperature can be predicted. The advantage of this system is that action is taken before changes occur in the outlet stream temperature. The disadvantage is that if the actual heat exchange capacity differs from the expected value due to fouling or mechanical problems, the outlet temperature will not be correct.

Example $4.3.2B$: An alternate notation, with explanation of controller location and function

This presents an alternate notation of controls. The arrows above the controller indicate what parts of the process that the controller controls.

$$\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{AB}$$

1. B is added slowly into a full charge of A
2. The temperature is maintained at TR (RXN temperature)
3. Upon completion of reaction, AB is cooled to TP
4. C is added to AB to make final product, which leaves the tank



Solution

- **LC1**: This controls M1 and V1.

This controller is used to keep the level of A in the A tank above a preset minimum.

- **LC2**: This controls M2 and V2.

This controller is used to keep the level of B in the B tank above a preset minimum.

- **LC3**: This controls M3 and V3.

This controller is used to keep the level of C in the C tank above a preset minimum.

- **FC1**: This also controls M1 and V1.

This controller totalizes the amount of A that has flowed out of the A tank, and turns off flow when the flow amount has reached a preset value.

- **FC2**: This also controls M2 and V2.

This controller totalizes the amount of B that has flowed out of the B tank, and turns off flow when the flow amount has reached a preset value.

- **FC3**: This also controls M3 and V3.

This controller totalizes the amount of C that has flowed out of the C tank, and turns off flow when the flow amount has reached a preset value.

- **LC4**: This controls V1, V4, M4, and M5.

V1 - If level of liquid in the reaction tank goes over a preset maximum, V1 is closed.

V4 - If level of liquid in the reaction tank goes under a preset minimum, V4 is closed.

M4 - If level of liquid in the reaction tank goes under a preset minimum, M4 is turned off.

M5 - If level level of liquid in the reaction tank goes under the level of the agitator, M5 is turned off.

- **TC1**: This controls V3, V5, and M3.

M3, V3 - If the temperature in the reaction tank is at TP, M3 is turned on and V3 is opened.

V5 - To maintain the temperature in the reaction tank at TP, V5 is opened and closed accordingly.

References

- Bequette, W.B. *Process Control Modeling Design and Simultion.*, New Jersey: Prentice Hall, pp 466-477
- Luyben, W.L. *Plantwide Process Control.*, New York: McGraw-Hill, pp 53-69, pp 201-203, pp 235-247
- *Module 3.1: Control of Distillation Columns.* The ECOSSE IGDS Controls HyperCourse. University of Edinburgh, Scotland.
- Ogunnaike, Babatunde A. *Process dynamics, modeling, and control*, New York: Oxford University Press 1994, pp 1149-1153
- Riggs, James B. *Chemical and Bio Process Control.*, Ferret Publishing, pp 491-504
- Turton, Richard. *Analysis, Synthesis, and Design of a Chemical Process*, New Jersey: Prentice Hall, pp 408-409
- U.S. Chemical Safety and Hazard Investigation Board; "Morton International Explosion." www.csb.gov/completed_investigations/docs/mortondigest.pdf. Last Updated 07/17/2003.
- Tompkins, Bill. "Reaction and action." www.ehponline.org/members/2004/112-6/focus.html. Last Updated 09/09/2007
- Wikipedia. "Gas Compressors" http://en.Wikipedia.org/wiki/Gas_compressor. Last Updated 11/27/2008
- Wikipedia. "Liquid-Liquid Extraction" en.Wikipedia.org/wiki/Liquid-liquid_extraction. Last Updated 11/20/2008

Contributors and Attributions

- **Authors:** Christopher Bauman, Jennifer DeHeck, Merrick Miranda, Evan Leonard
- **Stewards:** Robert Chockley, Chloe Sweet, Irene Brockman, Scott Dombrowski
- **Anchors:** Megan Boekeloo, Stacy Young

This page titled 4.3: Piping and Instrumentation Diagrams - Location of Controls and Standard Control Structures is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

4.4: Piping and Instrumentation Diagrams - Standard Pitfalls

Introduction

Piping and Instrumentation Diagrams (P&ID) are standardized in many ways, and there are some fundamental safety features that are absolute requirements for all P&IDs. Unfortunately, many people forget these features in their designs unintentionally. Lacking these safety features could lead to serious engineering problems. It is important to eliminate these pitfalls when designing a P&ID.

In the following sections, different pitfalls of P&IDs will be discussed. The equipment design section will concentrate on how equipment might be accidentally neglected and misplaced while designing a safe and functional process. The design process section will describe how a lack of consideration of process conditions would lead to serious design pitfalls. Overspecification and underspecification of equipment and design will be discussed. There are also some miscellanous pitfalls associated with interpretation and inherent problems of the P&ID itself. Examples are introduced at the end to illustrate common errors and pitfalls of P&IDs.

P&ID Production Pitfalls

Most of the common pitfalls of P&ID production result from the engineer forgetting that the design on paper represents a real physical process and that there are practical and physical limitations that need to be considered in the process design. The following are some errors to be particularly careful of when drawing up a P&ID:

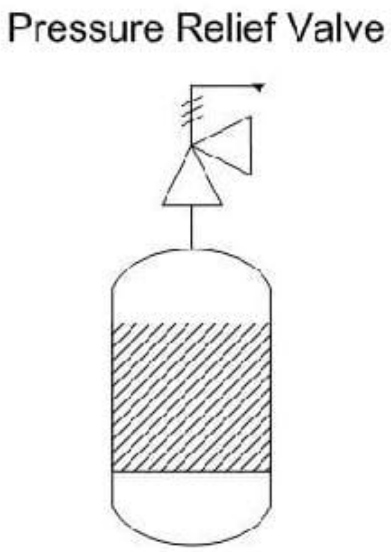
Equipment Design

Safety valves

Safety valves are part of the essential valves system for P&IDs. Together with isolation valves, they are an absolute requirement for instrument design. Safety valves are required to install for all gas, steam, air and liquid tanks regardless of the tank's function for pressure relief purposes. Engineers should be aware of their system's set pressure, relief pressure, percent overpressure, maximum allowable pressures etc. when selecting a safety valve. The US law requires all tanks of pressure greater than 3 psig to have safety valves installed.

Different pressure tanks require different safety valves to best fit their safety design. Therefore, engineers must be very careful in selecting the right safety valves for their systems. However, when constructing a P&ID, engineers sometimes forget adding safety valves to their design, and this could cause serious problems. For example, if one forgets to add a pressure relief valve or safety valve on a reaction tank of gas and liquid, the extra pressure accumulating would exceed the preset pressure limits for safety design. This could lead to a serious explosion!

A pressure relief valve is symbolized by two triangles orientated at 90 degrees to each other, as shown below.

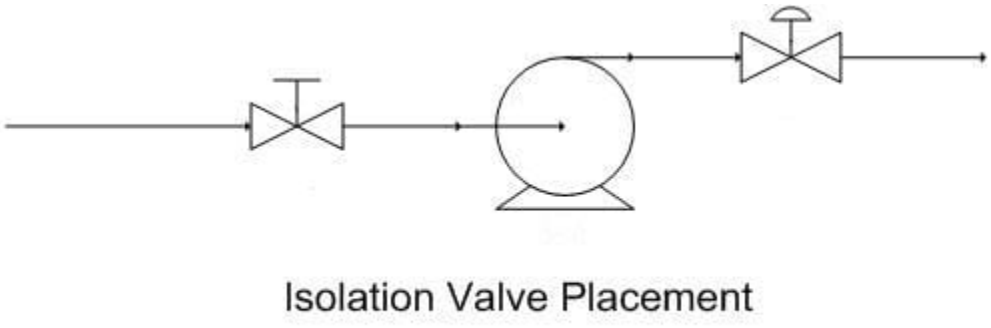


This eliminated excess pressure that might build up inside a reactor.

Isolation valves

The isolation valve is used to isolate a portion from the system when inspection, repair or maintenance is required. Isolation valves are placed around the junctions in the distribution system. They are also part of the absolute requirement for P&ID construction. Engineers should be aware of their system's pressure, voltage, process medium, pipe sizing and flow rates when selecting the right isolation valves that work under the conditions for their system.

If engineers forget to add isolation valves on their P&ID design, serious problems would occur when the system needs to be partially shut down for maintance or other reasons. There would be no way to control the unit's operation other than existing flow valves. Some upstream production problems could affect downstream production since appropriate isolation valves are not installed for safety and production purposes.



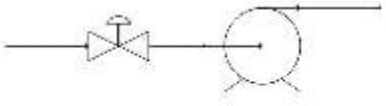
Notice that the automatic valve after the pump is a flow valve, while the manual valve before the pump is an isolation valve. One example of this use is for cleaning. If some parts of the pump cannot be cleaned in place, the pump will need to be taken apart. The isolation valve can cut off the flow to the pump, allowing it to be safely taken apart.

Valves and pumps

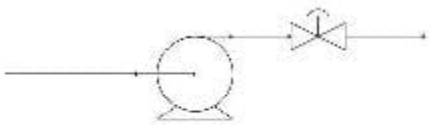
Some of the most commonly used pieces of process equipment that show up on P&IDs are valves. Valves control the flow of fluid through pipes by opening to allow flow or closing to stop flow. One of the problems associated with valves on P&IDs is the sheer number of them and deciding where to place them in the process.

A common mistake with valve placement has to do with the interaction between valves and pumps. It is important that valves be

placed **after** pumps in a given pipe. When a valve is placed before a pump, should the valve close and the pump has not been shut off, there will not be a constant supply of fluid to pump. This is known as *starving* the pump. Starving the pump can create a disturbance known as cavitation, and it is very hard on the equipment and can cause the pump to break. Placing valves after the pump ensure that even if the pump does not shut off, it is still filled with the proper fluid.



Improper Valve Placement

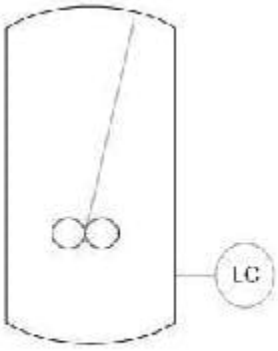


Proper Valve Placement

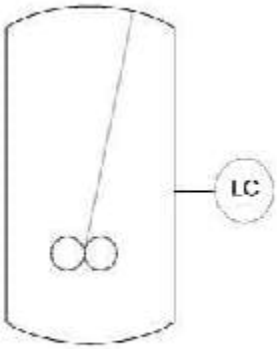
These same principles apply to valve placement with respect to compressors. Placing a control valve upstream of a compressor eliminates your ability to control pressure in the pipeline downstream of the compressor and introduces a risk of starving the compressor.

Agitators

A point that is very easy to miss, and very expensive if it is missed, is that if a vessel is equipped with an agitator, the vessel **must** be filled enough to submerge the agitator before the motor is turned on. Agitators are designed to operate against the resistance of fluid. Without this resistance, the agitator will rotate much faster than the equipment is designed for. This rotation may result in hitting the harmonic frequency of the structure, compromising its integrity and causing the agitator to rip from its foundation. This creates not only a fiscal predicament (not only ruining the agitator but also the vessel), but a safety nightmare. When designing a process, one must make sure he or she knows and accounts for how much fluid must be in an agitated vessel to avoid this situation. This can easily be solved by adding a level sensor to the tank that the agitator is being used in.



Improper Placement



Proper Placement

When placing the level sensor on the tank, make sure to place the sensor above the level of the agitator. This will ensure that the agitator is submerged in the fluid. It would be incorrect to place the level sensor below the agitator.

Instrument selection and placement

Instruments are designed to operate properly under specific conditions. Every instrument has a range over which it functions properly, and instruments must be selected that are appropriate for their applications. For example, a pressure gauge might have a working range of 5 psig - 50 psig. You would not want to use this gauge for sensitive measurements in the range 3 - 6 psig. Instrument material must also be considered during the selection process. If the substance being monitored is corrosive, for example, the instrument must be made of a corrosion-resistant material.

Once an appropriate instrument has been selected, it must be appropriately placed. For example, a level control is not useful in a pipe because there is no need to measure any water level inside of a pipe, much like a flow controller is not useful in a storage tank because there is no flow. Similarly, a flow controller should not be placed on a valve, but instead downsteam from the valve. However, level controls **are** useful in storage tanks and reactors, while flow controllers are useful in pipelines. Instruments must

be selected and placed to reliably provide useful information and to accurately control the process.

Equipment Selection

When creating a P&ID, the equipment that is selected to be used is very important, not only to maintain a smooth process but also for safety purposes. Each and every piece of equipment from 100,000 liter storage tanks to temperature sensors has *Operational Limitations*. These are the conditions under which a given piece of equipment operates as expected, with safe, consistent, and reproducible results. For example, when storing a highly pressurized gas at 2,000 psig, one wouldn't want to use a storage tank that has been pressure tested up to 3 psig. The process conditions are way outside the operational limitations and would pose a serious safety hazard. The same goes for sensors and gauges of all types. On a pressurized vessel at 2,000 psig, it would be no good to use a pressure control system that has a sensor that is meant to measure up to 100 psig, the results would not be accurate.

Operational limitations can usually be found in the equipment manual sent by the manufacturer or possibly on the manufacturer's website. If it is not found in either place, an engineer is obligated to contact the manufacturer and find the operational limitations before using a piece of equipment in a process. More can be found on equipment selection for some (but by no means all) equipment below:

- Temperature Sensors
- Pressure Sensors
- Level Sensors
- Flow Sensors
- Composition Sensors
- pH and Viscosity Sensors
- Biosensors
- Valves

Process Design

Unit operation input/output
Providing clearly specified inputs and outputs to and from process units is vital. For the safety of the system, proper control and the prevention of disaster, it is important to show where each substance came from and where it is going. The P&ID must show all material streams to and from separation units, heat exchangers, and reactors.

For example, if the P&ID is not clearly denoted in a reaction to create construction explosive materials, then a large and perhaps fatal calamity can occur at the chemical plant. The exiting streams of a plant may pour into a river which may violate environmental regulations. Therefore, specifying inputs and outputs is imperative in a P&ID.

Pressure and flow
The movement of fluid is essential in many production lines. Transporting material from a tank to a reactor, a reactor to a distillation column or from a column into a tanker truck all involve the movement of fluids. When designing P&IDs, a process engineer must decide how they are going to attack this problem, namely whether or not a pump is needed to move the fluid.

The main issue here is pressure. Fluid, both gaseous and liquid, moves down the pressure gradient from high pressure to low pressure. The rule of thumb is that if the source of the fluid has a much higher pressure than the destination of the fluid, a pump is not needed. Even if the source has a pressure only a fraction higher than that of the destination, a pump may yet be needed for the fluid to flow through the pipes, from the source to the destination. However, the flow and pressure from the source must also be compared to the needs of the process. For example, if a inlet flow has highly fluctuating pressure, a pump or valve should be used to regulate the pressure for the process. Also, if the diameter of the pipe does not remain constant and is fluctuating as well, perhaps due to outside pressure or force, a pump should be used to control the pressure inside the system.

A common mistake young process engineers make is when charging a pressurized vessel, they do not use a pump powerful enough to overcome that pressure. This causes *backflow* and can ruin process equipment. This problem is especially prevalent in recycle streams. In some situations it may be appropriate to use a check valve (a valve allowing fluid to flow through into it from only one direction), to add an additional barrier against backflow. An example of such a situation would be when a fuel is to be mixed with air for combustion: if any air were to backflow to the fuel source, a dangerous, explosive situation would arise. In this case the use of a check valve would be appropriate.

An example of when a pump is never needed is when liquid from one vessel at ambient pressure is being transported to another

vessel at ambient pressure that is at a lower elevation. This is known as a *gravity feed* and utilizing gravity feeds where possible can significantly decrease the cost of a process. As long as the pressure at the exit of the pipe is lower than the pressure created by gravity, a gravity feed can be used. An example of such a gravity feed is the flushing of a toilet, by which the water from the cistern at a higher elevation falls to the water closet at the lower elevation.

Underspecification

For safety and control purposes, **redundancy** is desirable in control systems. A process lacking in suitable redundancy would be **underspecified**. For example, a reactor containing an exothermic reaction would be underspecified if it only contained a temperature control. The potential danger posed by an exothermic reaction warrents a high degree of safety: if one controller were to malfunction, a backup must be in place. A pressure control should also be included, and perhaps a level control as well depending on the specific process.

Overspecification

On the flipside of underspecification is **overspecification**. Adding too many controllers and valves on vessels and lines is unnecessary and costly. Since process control is highly dependent upon the situation, it is difficult to set specific limits on the necessary amount of specification. However, a controller should not be used unless there is a specific need for it. Misused redundancy is a common example of overspecification. It is unlikely that a water storage tank needs level, temperature, and pressure controllers. A degree of freedom analysis, as shown in a distillation column example, can be helpful in determining the exact number of controllers, sensors, actuators, and valves that will be necessary to adequately control a process.

Problems When Using a P&ID

Other than equipment and production pitfalls, there are also some general P&ID interpration and inherent problems.

Interpretation of P&IDs

Although it is essential for P&IDs to represent the right instruments with the right references, many P&IDs do not support scaling or do not require scaling as part of the system. P&IDs are drawn in a way that equipment and piping are displayed for ease of interpretation. They do not show the physical placement and location of different systems or the actual sizes and length of equipment and pipelines.

Experience tells us that many engineers have overlooked certain pieces of equipment or over-estimated the size of equipment and piping, leading to calculation and construction errors. For example, an engineer could be designing new equipment on the existing P&ID and not realize that existing equipment is so densely populated that extra equipment will not fit. It is important to check the physical space of an area before adding equipment to a process.

Inherent Problems of P&IDs

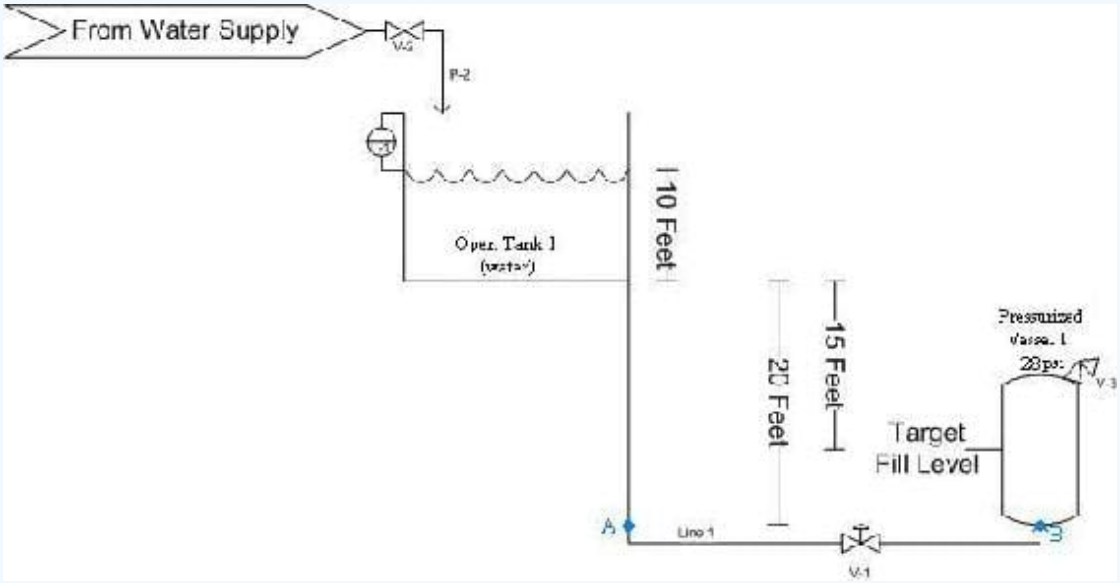
There are a few problems with P&IDs that are only solved by being conscious of them. One problem is that P&IDs are constantly being updated, revised, changed, and added to. When reading a P&ID it is always important to check the date it was last revised and if there is a later revision available. Making adjustments to equipment that isn't there anymore not only causes confusion and frustration, but is a waste of everyone's time. Moreover, making sure you are using the most recent edition to the P&ID will ensure that you do not purchase equipment to install that has already been installed.

Usually, companies will have a computer database with the most current P&IDs. Before modifying or working with an old paper copy of a P&ID, check to make sure it is up to date with the most current revisions.

Another problem with P&IDs is that even if they are well made and technically perfect, complex processes often appear cluttered. There is an inordinate amount of information contained in a P&ID, and as such P&IDs are next to impossible to take in at a glance. It is important to carefully study each document to fully understand the process. The mistake of assuming you gained all the information off of a P&ID from simply scanning it can leave you with mental holes that you don't even realize you have. This can be avoided by reading the P&ID carefully on the first pass and constantly referring back to it when questions arise.

Example 4.4.1

A pipe connects a water tank (open to the atmosphere) and a vessel pressurized to 28psi. The open tank has 10 feet of water in it. A level control system ensures that the 10 foot level is maintained in the open tank. The bottom of the pressurized vessel is 20 feet below the bottom of the open tank and starts with no water in it. The goal is to fill the pressurized vessel up to 5 feet. Due to a poor design by the project engineer, the water is fed into the bottom of the pressurized vessel. Given that the density of water is 62.4 lbm/ft^3 and the gravitational constant is 32.2 ft/s^2, is a pump needed? If so, where should it be placed? Assume that there is no pressure drop due to friction against the pipe and that the air pressure of the pressurized tank remains at a constant 15psi. The figure below may be helpful in visualizing the process. (Hint: Remember in English units, a "fudge factor" is needed when using P = ρ*g*h)



Solution

First, we must determine the pressure at the stream exit (point B). When the vessel is full, we know that the pressure created by the water is:

$$P = \frac{\rho g h}{32.2}$$

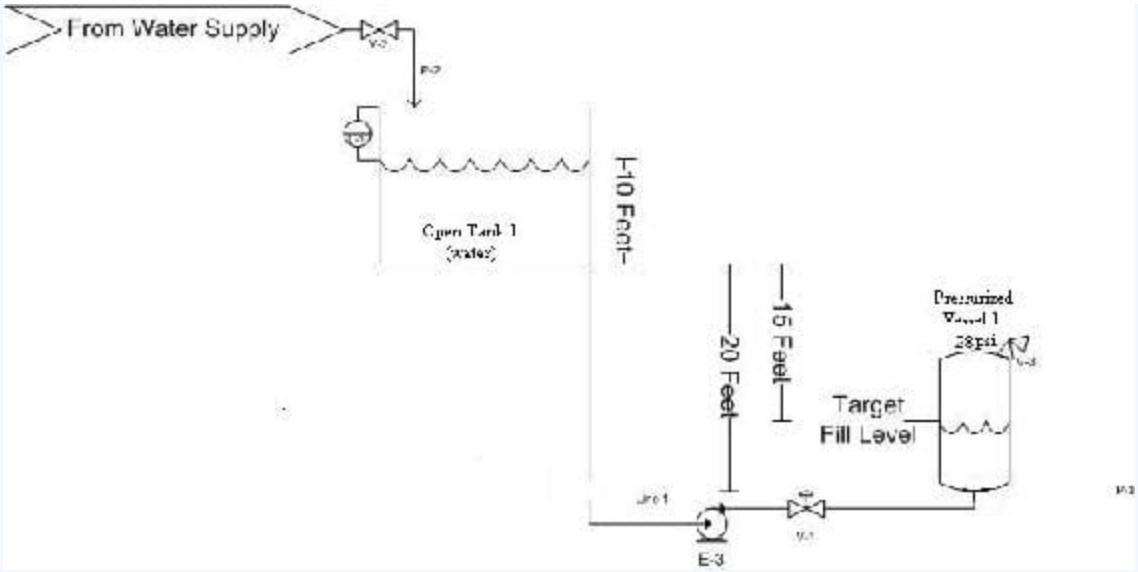
Plugging in 62.4 for the density, 32.2 for g and 5 for h, we find that the pressure caused by the water is 312 lbf/ft^2. Dividing by 144 to convert to psi, we see this is equivalent to 2.2 psi. When we add the 28psi from the air pressure in the vessel, we get a grand total of 30.2 psi at the pipe exit.

The pressure in the feed pipe at the height of the exit to the pressurized vessel (point A) is again given by:

$$P = \frac{\rho g h}{32.2}$$

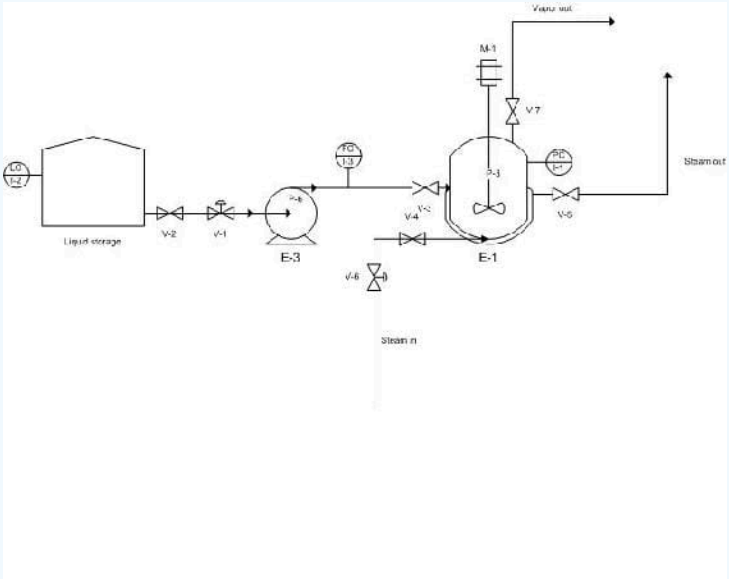
Only this time, the top of the water is 30 feet above the pipe exit instead of 5 feet. Plugging in this value, we find that the pressure is 1872 lbf/ft2. After converting, we find this is equivalent to 13psi. Since the water tank is open, we must also add 14.7psi for atmospheric pressure. This gives us a total pressure of 27.7psi.

The pressure of the feed at point A in the diagram is LESS than the pressure at point B, so a pump is needed. The pump should be placed BEFORE the valve labelled V-1 on the pipe as shown below. This is to prevent starving the pump when the valve closes.



Example 4.4.2

There are several errors in the following P&ID, which shows a vessel where a liquid is vaporized using heat transfered from steam passing through a surrounding jacket. Identify the errors and explain why they are mistakes.



Solution

1. Control valve V-1 should be after the pump, not before. Placing a pump after a control valve takes away your control of the flow rate and can also damage the pump if the valve is shut off while the pump is running.
2. Vessel E-1 is only being monitored with a pressure control. This is not adequate redundancy for process control. There should also be a temperature control (to make sure the vessel is at an appropriate temperature for vaporization) and a level control (to make sure the agitator is always submerged in liquid).
3. Although the vaporization takes place at ambient pressure, there should be a safety valve on the vessel. What if there is a malfunction and too much vapor is produced?
4. The input and output streams for vessel E-1 are not sufficiently labeled. The source of input streams and destination of output streams should be specified to ensure adequate control of a process.

5. There should be a flow meter and control valve on the vapor stream exiting the vaporization vessel so that you can monitor and control the amount of vapor exiting.

References

- Center for Chemical Process Safety. (March 2006). *Process Safety Beacon: Pressure Relief Systems - Do You See Any Hazards?* CEP
- Melhem, G.A. and Howell, P. (September 2005). *Designing Emergency Relief Systems for Runaway Reactions.* CEP
- Ulrich, G.D. and Vasudevan, P.T. (July 2006). *Predesign With Safety in Mind.* CEP
- Williams, J.P. (September 2004). *Reliability for Safety Instrumented Systems.* CEP

Contributors and Attributions

- Authors: Angela Yeung, Curt Longcore, Ben Van Kuiken, Jeffrey Carey
- Stewards: Eric Black, Megan Boekeloo, Daniel Carter, Stacy Young

This page titled 4.4: Piping and Instrumentation Diagrams - Standard Pitfalls is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

4.5: Safety Features in Piping and Instrumentation Diagrams

Introduction

Safety has become integral to the manufacturing world. The implementation of proper safety techniques and accident prevention can not only save time and money, but prevent personal injury as well. P&IDs, when properly utilized, are a powerful resource to identify safety hazards within the plant operations. The following sections provide an overview for the safety hazards that exist within a process, and illustrate the importance of P&IDs in a chemical plant.

Safety in Design

During the early stages of plant design it is critical to determine important safety features that remove potential hazards from effecting the facility environment. Regulations require that plant designers play a major role in minimizing the risks associated with these hazards. However, in order to do so, designers need to be aware of the hazards that exist during plant activity. The facility design team must develop a detailed drawing (P&IDs) including specifications of the plant process and environment to ensure that every aspect with regards to safety is covered.

Hazard vs, Risk

When discussing safety, the terms hazard and risk are often used interchangeably. However, the difference in definition between the two terms is critical in utilizing the information they provide in increasing the safety within a plant. Hazard is defined as a potential source of danger and risk is defined as the level of threat associated with the hazard. A risk of a hazard occurring can be represented mathematically by the following equation:

$$\text{Risk} = \text{Frequency} \times \text{Consequences}$$

Frequency represents the probability that a hazard will occur, and consequence represents the impact of that hazard. Values for each parameter of the equation above are assigned by using either experimental information or educated judgment based on engineering models and experience. Each plant design process will have specific safety hazards and risks associated with it. Therefore, there is no predetermined value that can be assigned to each variable. For example, in one situation, a water tower may be well supported inside the plant facility, and in another situation a water tower of similar structure may be hoisted against a rusted frame work outside of the plant facility. Both of these situations have different levels of risk associated with them. Needless to say, the risks associated with a plant setup can be reduced substantially by minimizing the probability that a hazard will occur.

Hazard Locations and Risk Hotspots

Common hazard locations exist in any place containing large amounts of energy. The degree of danger is proportional to the amount of energy stored at that location. Risk can be directly linked to kinetic energy, potential energy, work, heat, enthalpy and internal energy sources. Kinetic energy, otherwise known as energy in motion, is present in any moving component. The component may be vibrating, rotating, or translating, and this motion causes the kinetic energy of that part to greatly increase. Within industry, personal injuries and fatalities are common hazards associated with moving parts. Reactors and cooling towers placed in high locations contain potential energy. If a structural failure were to occur within a plant, these structural units and their contents could fall from a large elevation onto another processing system or a human being, releasing all the chemical contents stored inside. Similarly, the stored work in springs and other devices can cause fatigue and wear on the mechanical system over time, and result in eventual machine failure. Heat released from a reaction within a chemical reactor can be rapid and fatal if not accounted for. The buildup of heat can cause serious consequences with runaway reactions and boiler explosions. The enthalpy and internal energy of a reaction typically are the cause of most runaway reactions and destructive fires in a plant.

Risk Hotspots

Uncontrolled chemical hotspots created within a chemical plant are a common source of hazard, besides the hazard locations pertaining to energy sources such as kinetic and potential energy. Risk hotspots mostly occurs in the piping system and associated valves of the system, joints, traps, and various other piping elements. Possible malfunction of the system due to

structural corrosion can be triggered by the failure to maintain the piping systems efficiently and periodically. Even if one of the valves in the system has corroded and is unable to function properly, the fluid flowing through the pipe might get trapped, and the resultant buildup of pressure in the pipe may cause major safety hazards including fatal ones such as an explosion. If the system is not shut off before the pressure gets out of control, the pressure buildup in the pipe will cause it to burst, releasing all of its internal contents to the surrounding environment.

Storage vessels are other pieces of equipment that must be cared for properly. Since the plant operators and engineers do not usually interact with the storage vessels, as compared to other parts of the plant process such as the piping system, they are considered to be of secondary importance and commonly overlooked. Storage vessels have much more content inside them than pipes, so a leak or a burst would take longer to get back under control than pipes, which can by plugged more easily. Problems which may arise from storage vessels are not associated with their design, but in fact from not thoroughly and periodically maintaining them. Possible complications that can arise from neglecting the storage tanks are over-pressurization, overfilling, heating element malfunction, or simply equipment malfunction.

Chemical reactors are another common location where risk hotspots occur. The nature and design of commercial chemical reactors is to handle a controlled explosion. However, if the control element is removed, disaster is bound to occur, just as in any other part of the chemical plant. The most common type of hazard is a runaway reaction inside a batch reactor. When the plant facility looses electricity or cooling water, a runaway reaction will strike inside the reactor. Once a runaway reaction has spun up inside the reactor, many other hazards may follow, such as flow reversal in the pipes, incorrect reagent charging, heat exchanger failures, and external fires. Other hazards may perhaps be even more serious, such as engineering errors that could potentially cause a runaway reaction to occur, including inappropriate material selection, inadequate equipment inspection or failure to fully understand the chemistry or exothermic nature of a reaction. A runaway reaction originating inside the chemical reactor can easily cause a chain reaction across the rest of the equipment at the facility, and can result in the entire system malfunctioning.

Other process equipment that may be hazardous and where risk hotspots commonly arise, are vacuum operators, furnaces, pumps, gas movers, compressors, and heat exchangers. The location and type of specific piping and unit operations are available on the process P&ID. A responsible process engineer should use the P&ID to identify all risk hotspots, and act accordingly to monitor and maintain a safe working environment. In addition, a standardized plan should be constructed so that in the event of a malfunction, the correct steps can be taken to bring the faulty part back under control. Supplementary precautions should be taken to prevent a comparatively minor malfunction in the system from becoming a disaster which may violate environmental regulations and even endanger human lives.

Safe Design Principles

The ISD or Inherently Safer Design movement was a doctrine striving for safer chemical processing procedures. This movement was pioneered by Trevor Kletz in 1976, and promotes the design of processes so safe, that no catastrophic failure can occur within the plant. The following principles apply to initial process design:

1. Use the fewest number of hazardous substances in the smallest quantities and still maintaining plant productivity
2. When possible, substitute hazardous chemicals with chemicals that are less dangerous
3. Practice moderate operating conditions in the plant
4. Use the simplest plant design possible
5. Design equipment in the plant to minimize the effects of a hazardous incident

The specifications determined by the process designers are communicated through the P&ID.

Hazards in Construction

In order to eliminate hazards, a operations personnel must be able to identify that a hazard exists. Hazards that may be encountered on plant sites may be categorized into three main types:

1. Hazards harmful to health: When workers are exposed to or come in contact with asbestos, corrosives, irritants, toxins, or noxious gases try to avoid by specifying the processes, which lead to this exposure
2. Hazards likely to cause personal injury: Hazard awareness is increased when people have to work in situations likely to expose them to the risk of personal injury, including moving plant machinery or working in areas where objects are likely to fall. Situations where there are live electrical circuits overhead, buried power lines, and confined working conditions are likely to cause personal injury.
3. Hazards likely to lead to a catastrophic event: These hazards have consequences beyond the site boundary. They include fire

outbreaks, explosions, flooding, or premature collapse of structures, cranes, tunnels and excavations.

Fail-Safe Design

The fail-safe design of a unit operation (such as a reaction vessel) requires a complete understanding of the operation at hand, and the knowledge of all the worst-case conditions. A fail-safe system is a unit operation such that, if any or all of the worst-case conditions were to occur, the operation would shut itself down automatically and in a safe fashion. In the case of a run-away reaction, if reagent feed limits, interlocking controls, and integrated heat balances are all properly maintained, the reaction cannot "run away." Other precautions such as purges, vents, dump tanks and quenches are available for reaction vessels, and should be visible on the P&ID.

Inherent Safer Predesign

The table below provides a guideline for identifying and minimizing hazards partly based upon Kletz's rules for ISD in the Safe Design Principles section above. The step/rule column describes the action taken. The tools column describes the mechanism by which the actions are taken. The experimental and analytical resources column describes the knowledge by which the mechanisms are created. The Literature References column describes where the knowledge can be found.

Step/Rule	Tools	Experimental and Analytical Resources	Literature References
1. Identify	Fire triangle, Flammability-limit chart, Chemical reactivity chart, Safety compatibility chart, Safety stream chart, Toxicity ratings	Thermodynamic calculations, Reactor design equations, calorimetry, flammability charts	Perry's Handbook (10) Table 26-10 and associated pages
2. Eradicate (ISD Rule 2)	Inert-gas blanketing, Failsafe design	see above	see above
3. Minimize, Simplify, Moderate, Attenuate (ISD Rules 1, 3, 5)	Moderate ignition sources, Keep reactive volumes small	Logical chemical engineering analytical and computational skills	Merk Index, MSDSs, Chemical Engineering textbooks and monographs
4. Isolate (ISD Rule 4)	Separate hazardous operations - Surround in an impenetrable structure	Logical chemical engineering judgment	

Relief Systems

Organizations such as the American Society of Mechanical Engineers, American Petroleum Institute and National Fire Protection Association layout recommendations and design standards so that most engineers with proper training can setup proper emergency relief systems for single phase flow. Unfortunately, this is often not good enough for reactive systems. These systems are much more complex and include multiple phase flow, runaway reaction potential and self reactant material. When designing an emergency relief system (ERS) it is necessary to understand all aspects of the chemicals and processes that will be in play. This includes but is not limited to: kinetics of the possible reactions, contamination, interactions with air, rust, piping, or water, phase changes and runaway reactions. The follwing topics deal with the hazards that require designing relief systems and the prevention of runaway reactions.

Planning and Design

When designing an ERS, it is important to consider the worst-case scenario. This is based upon a thorough knowledge of the reactions, materials, and environment of the process. Some of the most critical scenarios are over-pressurization in a reactor.

All disaster scenarios can be analyzed using a hazard and operability (HAZOP) study. This HAZOP study will analyze a process based upon human, equipment, and environmental factors.

The HAZOP team must take the following steps to ensure that all potential scenarios are taken into account:

- Review the potential hazards of all chemicals. This includes non-operational conditions and interactions with contaminants

- Study the chemical process including all possible reactions, rearrangements, decompositions, etc.
- Review the P&ID's for the process
- Study the specific reactor and storage vessels for material composition compatibility, size, surface area, instrument ranges, and set points

With this information, the HAZOP team can determine most of the potential disaster scenarios. Most likely the worst case scenario will involve fire induced runaway reactions.

Design Strategies

There are various techniques often used to prevent failure from over-pressurization, fire, runaway reaction or other disaster scenarios. The following are some design suggestions that will decrease this risk:

- The use of insulation in case of fire. Unfortunately, insulation will also minimize heat loss from the reactor during a runaway reaction. It is important to consider the ability of a vessel to drain when using insulation. If the contents can be drained and cooled before the reaction starts to runaway, catastrophe can be prevented.
- Design to avoid fire damage on sensitive equipment. The possibility of fire damage on electrical equipment or sensors will cause control difficulties when trying to slow down a runaway reaction, or monitor another emergency situation.
- Consider the structural integrity of the entire system due to fire damage.
- Use multiple purge streams and valves to separate materials. Separating reaction components into a storage vessel with a large surface area will allow for faster cooling.
- Install two separate relief devices in case one has been compromised due to fouling or solid particle blockage.

Overall safety relief plans must be made for all possible scenarios.

Reaction Kinetics

Reaction kinetics usually determine the potential for safety disaster. For the average exothermic reaction, the reaction rate doubles with every 10°C increase. This will lead to an exponential increase in energy which will force both the temperature and pressure in a system to uncontrollable levels. Since reaction rates are sensitive not only to temperature, but also pressure, contaminants, concentrations and phases, all possibilities must be adequately analyzed for plant operation to be deemed safe.

Relief devices should be designed to handle chemicals and/or mixtures in any phase. For example, if a reactor temperature increases suddenly and a safety relief valve is activated the material must be able to flow through the valve quickly and safely. If the chemical has changed phases, or has a higher pressure that the valve cannot accommodate, the chemical will not be able to escape and build up pressure in a reactor. This would greatly increase the risk of a disaster. Because of the extent of possibilities when considering reaction rates and kinetics, it is usually necessary to consult some sort of computer simulation or dynamic simulation tool to plan for every possible scenario.

Runaway Reactions

Runaway reactions are caused when exothermic reactions are fed more energy due to malfunctioning cooling systems. This causes an exponential increase in temperature, which in turn causes an increase in pressure, and finally damage to the reactor and/or plant. The possibility of malfunctioning cooling systems must always be considered for flammable materials. If materials are being stored at temperatures above their flash point, fire is always a possibility. Fire will cause a reaction to reach runaway conditions with very little reactant. Relief systems for fire induced runaway reactions must be larger than conventional runaway reactions.

While relief systems are often designed for over-pressurization, this might not be enough when considering flammable materials. Constant or prolonged exposure to flame will cause most normal reactant or storage vessels to fail causing chemical leaks or plant fires. Fire proof insulation must be used together with normal pressure relief systems to prevent system failure. An important design variable when considering fire induced runaway reactions is fire flux. This variable considers heat impact on a reactor due to fire. Formulas for calculating fire flux can be obtained from the National Fire Protection Association. The elevation of flames must also be considered when designing reactor vessels and safety relief systems. Pool fires can produce flames that are hundreds of feet high; using the P&ID will be important in ensuring that process components sensitive to fire are sufficiently protected from something like pool fires.

Two Phase Flow

For systems in which multiple phase flow is possible, all phases must be planned for. For vapor hybrid systems, all-vapor flow should be considered. Likewise, when foam flow is possible, all-foam flow should be planned for. This is all part of the mindset of planning for the worst case scenario. The most conservative design should always be used when faced with uncertainties of phase flow.

Often times, a runaway reaction will cause high-viscosity two phase flow. Relief valves and safety features must be ready for viscous flow. Many resources suggest averaging the viscosities of the two phases to plan for two phase flow. It is important to keep in mind however, that two phase flow discharge will separate in the discharge line. This will lead to higher pressure drop. Piping sizes are often underestimated due to this unplanned scenario. Undersized piping will lead to valve rupture and back pressure and could cause venting disturbances during the emergencies where venting is critical to the safe shutdown of a process.

Regulations

There are often conflicts and contradictions in federal regulations and recommended practices for safety design guidelines. When faced with such confusion the designer should at least design to the level of the Occupational Safety and Health Administration's (OSHA) requirements, as well as meeting regulations set by the Environmental Protection Agency (EPA), the Process Safety Management(PSM), and the Risk Management Program (RMP). While these guidelines and regulations produce a minimum standard to follow, when dealing with reactive systems, this does not always meet process needs. For more details on the various regulatory agencies that govern chemical processes, see Regulatory Agencies and Compliance.

Additional Safety Hazards in Chemical Plants

This section outlines potential safety hazards commonly found in places where chemicals are stored or chemical processes are taking place. While these hazards are rather easily prevented through attention to detail and general awareness, neglecting them can have catastrophic consequences.

Time Sensitive Chemicals

Some chemicals have a "shelf life," or an expiration date provided by the manufacturer. The chemical must be used by this date or properly discarded. These are typically reactive chemicals, which can become unstable after a certain period of time, possibly rupturing the vessel in which they are stored. An example of this is a monomer that begins to polymerize unless an inhibitor is present. This inhibitor is completely consumed after a certain period of time, allowing polymerization to occur, and therefore must be used or discarded by that time. Another example deals with the formation of peroxide, which can be a severe fire and explosion hazard. It may also be a health concern, causing severe mucous membrane, respiratory tract, skin, and/or eye burns. Peroxide-forming materials should be stored carefully, labeled with the date received and the date first opened on the container. Chemicals should be disposed of or checked for peroxide formation after six months; do not open any container with solid formation around the lid. There are several ways to prevent this from happening. Material Safety and Data Sheets (MSDS) should be available for any chemical the plant uses. These can inform you if a particular chemical becomes unstable after a certain period of time. It is also important to make sure there are procedures in place for handling time-sensitive materials. Lastly, investigating near-miss occurrences can help to ensure that future incidents do not occur.

Pressure Relief Systems

Any open pipe in a chemical plant is a potential discharge site. While operator convenience and maintenance remains a concern, safety takes precedence when dealing with relief systems. It is essential to operator safety that relief valve discharge sites are located in areas that pose a low risk of exposing personnel to chemical hazards and are directed away from all access platforms. It is common in industry to tie multiple pressure relief devices to an emergency vent header that releases on the roof of the plant. Long stretches of unsupported pipes also pose a potential threat. Force generated by material flow could bend or break the pipe impeding on plant operations, and more importantly, injuring personnel.

Dust Explosion Hazards

Most flammable solids can form an explosive dust cloud if the particles are small enough. Materials such as wood, grain, sugar, plastics, and many metals can all form these explosive dust clouds. Dust explosions occur when a combustible material accumulates in a confined area and is exposed to an ignition source. High risks areas are usually those that are neglected such as tops of vessels and tanks, on pipes, storage bins, bucket elevators, and dust collectors.

To prevent dust explosions, it is essential to implement good housekeeping practices. Care must be taken so that the cleaning process, such as sweeping, doesn't turn a dust layer into a more dangerous cloud. Operations and process engineers should be aware of all the mechanical and safety control equipment associated with preventing fires. Equipment like electric vacuums must be appropriate for use in an area where an explosive dust cloud could form. This entails no holes or cracks in the cord, sufficient grounding, and receiving site approval for use of the equipment.

Overfilling Tanks

The overfilling of vessels has long been a leading cause of serious incidents in chemical and petroleum industries. When a level sensor or high-level alarm fails, reactive material can spill over the tank and accumulate. If this material is exposed to an ignition source, there is the potential for an explosion, leading to serious property damage, environmental issues, and injury to operations personnel. When filling or draining a tank with material, operators should be aware of all relevant level, pressure and temperature controls in place, and watch for abnormal trends. Also, all safety critical alarms surrounding a vessel should be tested regularly at frequencies recommended in plant process-safety-management procedures. Conducting regular process maintenance on safety critical alarms have numerous benefits, which include reducing the risk of operating a plant with faulty equipment, and increasing operator awareness with the location and function of critical safety devices.

Containing Storage Tank Spills

Engineering controls are implemented into plant designs to account for potential disasters, such as a spill, leak, or complete emptying of a storage tank. Industrial-sized chemical plants store large amounts of raw materials, products, and byproducts on-site. The amount of each varies, but is typically between three to thirty (or more) days of the required supply or amount produced. The stored chemicals can be hazardous, flammable, explosive, and/or reactive with each other. In the event of a tank spilling, dikes are built around tanks to contain the spill and protect the surrounding community from the spill. The regulations of the dikes include the following: dike volume must be 1.5X the largest storage tank contained by the dike, reactive materials cannot be stored in the same dike, and scuba gear must be present on-site if any dike is deeper than four feet. Since the dikes cannot contain any reactive materials, the implementation of dikes affects the overall plant layout. Typically, dikes are designed to be like speed bumps and have a height less than one foot so fork-lifts and tankers can easily maneuver through the plant. Dikes are an effective engineering control that greatly improve the safety of a chemical plant.

Temperature and Pressure Ratings

Before a vessel is put into plant operation, it is rated and stamped by the manufacturer with temperature and pressure limits. Problems arise when personnel overlook the inverse relationship between temperature and pressure for gases. Be aware that equipment rated for a specific temperature and pressure, cannot be operated at the same pressure if the temperature is increased. While it may seem extremely intuitive (think ideal gas law), this relationship is too often overlooked, usually with serious consequences. When operating any process, pay attention to the temperature and pressure ratings. If they aren't readily known, review equipment files or contact the manufacturer before making any changes to the process. Also, operation and maintenance should always be performed according to strict standards laid out by a plant standard operating procedure (SOP). Any changes must be reviewed and approved by a cross functional team.

The above hazards and preventive measures have been outlined to illustrate how important safety is in a chemical plant, and the importance of being aware of your surroundings and all possible safety hazards. Many of the aforementioned topics may seem like common sense, but it is very easy to overlook small details in the scope of a large-scale chemical process. Keeping safety in mind at all times as the paramount of any process can ensure that people leave their shift the same way they arrived.

Alarms in Processes

Alarm configuration and specification is an important part in the design and operation of any chemical process. Alarms are

implemented in a process design to aid in the control of the process. Federal and industrial documents only specifically reference alarms in the context of processes exceeding regulatory compliance limits. In this sense, alarms are used to control safety and environmental hazards. Other important uses of alarms are to control product yield, product quality, and operational limits of process equipment. This section will discuss the steps taken to implement alarms in processes, common different levels of alarms, and common instances in which alarms are useful to comply with regulations.

Alarm Lifecycle

When it has been determined that an alarm is needed to aid in the control of a process, the alarm must be specified. The following are basic steps in implementing an alarm.

1. First, the process designer needs to know what category the alarm fits into. This is important because responses from different categories of alarms are usually managed differently. They may be prioritized in case multiple alarms occur at once, so that the proper follow-up reports and procedures may be taken accordingly. Common categories include product quality parameters, safety, environmental considerations, and equipment protection.
2. After determining the specific use of the alarm, the limits must be set. If the process variable exceeds the set limits, the alarm will be triggered. (More information on limits will be discussed in Alarm Levels.)
3. Next, the computer system for the alarm must be configured. The computer system may contain logic loops that automatically change control parameters to offset the problem and merely inform the operator that it has done so. Alternatively, an alarm may trigger horns, flashing lights, or send a page to an operator, alerting them that there is a problem that needs attention.
4. The correct user response and interaction must then be defined for the alarm. This includes providing proper training for the particular process, procedures and operator manuals describing how the event should be investigated, guidelines on when action needs to be taken, and guidelines on when to escalate the situation to a more serious event.

Proper communication from the automated alarm system is critical. Care needs to be taken that alarms only signify abnormal conditions that require a response. For example, the successful completion of a batch operation is an important piece of information and should generate a computer message so that the operator knows. However, this does not represent an abnormal situation and should therefore not show up on the computer as an alarm. While this seems quite obvious, most industrial plants struggle with maintaining alarm systems free of "nuisance alarms." It is also important that alarms have proper descriptions so that the reason an alarm appears is clear. For example, an alarm could appear in a large plant that says, “LI-501 exceeds limits” by default. Although this might be useful to the engineer designing the alarm system as an indication that tank 501 has a high level of material, an operator that sees the alarm or other engineers working on the system may not know what the alarm is communicating. A more universally meaningful alarm indicator might say, “Material level in tank 501 is high.” With this simple change in the computer system, the alarm would be more effective for personnel to locate the problem quickly and act accordingly.

Alarm Levels

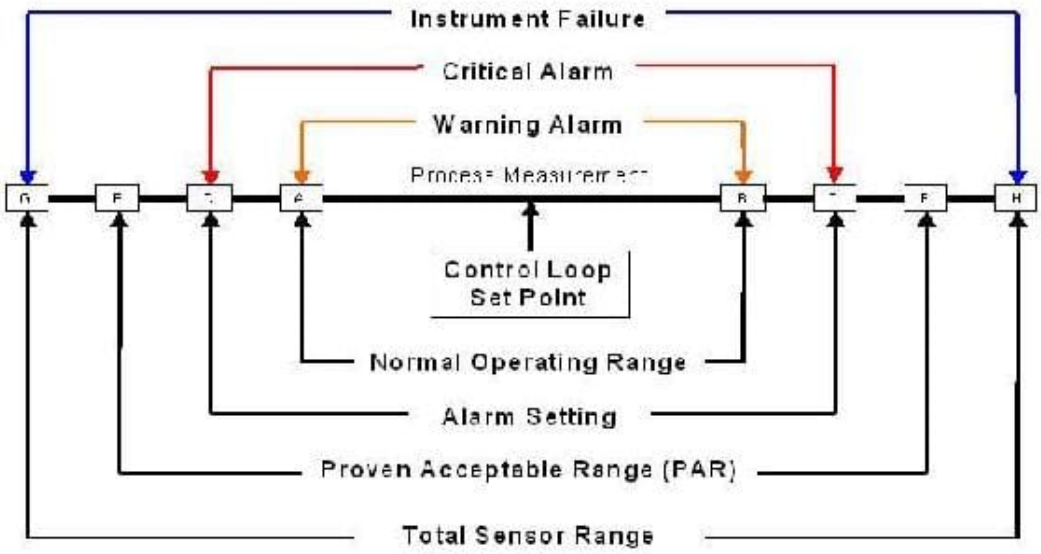
Alarms are available in a wide variety of types, with multiple levels of alarm. In all processes, disturbances occur that can shift a plant's operation away from normal. When this happens, measures are usually taken by computers, such as with the use of P&ID control loops, to keep the process under control. With these control systems, processes are designed to fall within a range of acceptable normal operating limits. When a process deviates beyond these normal limits, an alarm should be triggered.

For most processes, the minimum for safe operation is two levels of alarms: warning and critical. The warning alarm tells plant operators that the process has deviated beyond the acceptable limits and provides them with the time and ability to take corrective action so that the product quality is not affected and environmental and safety regulations are not exceeded. If the right actions are not taken or are not taken quickly enough to correct the problem, a critical alarm may then be triggered. The critical alarm tells the plant operators that conditions are dangerously close to breaching what is allowed. In many cases, the critical alarm will call for a systematic shut-down of the operation until the problems can be addressed.

The conditions at which warning and critical alarms are triggered are those conditions that exceed the limits determined for the process. Measurement uncertainty must always be considered because all devices in the control system will be subject to some possible error, even if it is small as +/- 1%. Setting an alarm at exactly the proven acceptable range for the process could allow a measured value to fall within this range, even though the actual value lies outside. This is called a "false acceptance." By performing error analysis and statistical distribution theory, the alarm limits can be adjusted as needed. This is a process called "guard banding," and it prevents real disturbances in the process from being ignored by the alarm system. Information about alarm limits should be well documented so if changes to the system are proposed, designers know how the limits were originally

determined.

The figure below provides a visual representation of alarm ranges.



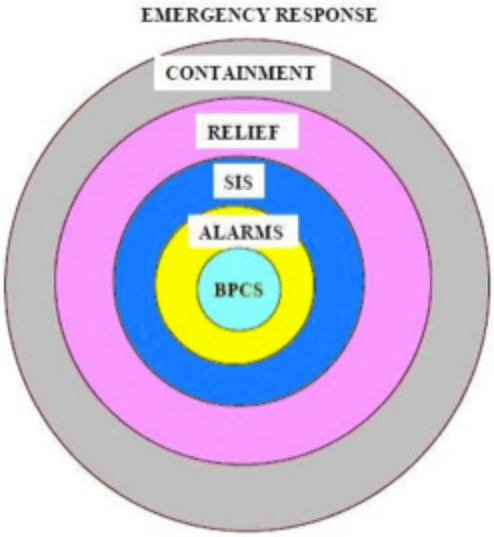
As seen in the center of the above figure, the control loop set point is the optimum point of control for the process (e.g. the optimum temperature and concentrations of reactants for a reaction). It is impossible to maintain the process at exactly this point, so there is a range of "normal" operation, inside which the process is still considered to be running in an acceptable way. The warning alarm would be triggered when the process goes outside of the limits of this range (lower than A or higher than B), allowing time for the process to be brought back under control. The critical alarm would then go off if the process goes beyond the alarm setting (lower than C or higher than D). This setting is determined by guard banding the process acceptable range (PAR) for uncertainty, seen in the figure by the fact that the alarm setting lies well inside the PAR. The spaces between E and C and between D and F are determined by the uncertainty. Lastly, this PAR must be inside the total range of the sensor that determines the instrument failure.

Alarms must be analyzed based on their priority:

- HIGH	= Hazard to people or equip., action required
- MEDIUM	= Loss of SS, close monitoring required
- LOW	= investigate when time available

SAFETY STEPS TO HIGH RELIABILITY

A safe system involves many layers of responses when an incident occurs.

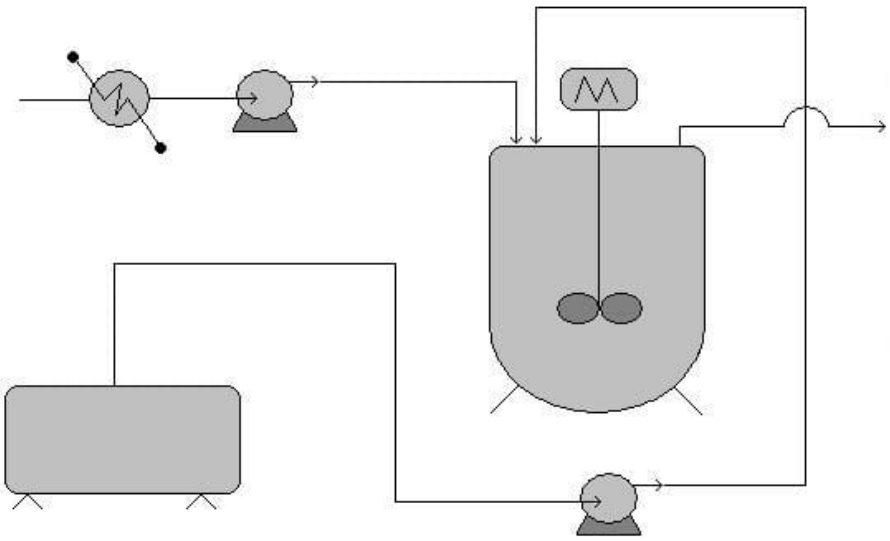


- The center of the ring is the basic process control system.
- The first layer of response is the alarm system which draws attention.
- The second layer is the Safety Interlock System which can stop/start the equipment.
- The third layer is the Relief system which leases pressure build-up in the sytem.
- The fourth layer is containment which prevents material from reaching workers, community, or the environment.
- The last layer to the ring is the emergency response system which involves evacuation, fire fighting, etc.

Worked out Example 1

A reagent recovery unit for a chemical process plant is being designed. The goal is to recover tin from a tin-plating waste stream through binary extraction with carbon tetrachloride. The equipment used in the process and P&ID are as follows:

- Carbon tetrachloride storage tank
- Mixing Vessel
- Pumps
- Heat Exchanger
- Associated piping



With knowledge of hazardous locations and risk hotspots, and safe design principles, analyze the flow diagram. Identify areas of risk and specific improvements that should be made to the process design before implementation begins. (note: no piping, valves, or controllers exist on the P&ID for the sake of clarity. Ignore these in your analysis) .

Areas of risk

Any piece of equipment could potentially become dangerous if the right situation were to arise. Each piece of equipment used in this process is no exception.

- Storage tank - These process units that are of secondary importance to the process goal, don't draw as much attention from plant operators and engineers, and tend to receive less maintenance.
- Mixing Vessel - The potential for heat buildup due to heat of mixing and the kinetic energy of the spinning motor makes this vessel a risk hotspot. Improper control of temperature within the vessel or motor speed (especially during periods when the tank may be empty) can lead to equipment malfunction or even explosions
- Pumps and Heat exchangers - The energy stored through pressurization by the pumps and potentially high temperature solutions handled by the heat exchanger present the possibility for danger. Wherever energy is stored, danger is associated with its potential release.
- Associated piping - Corrosion or failure to maintain pipes and associated elements are the main cause of the malfunction leading to danger. Leaks or total failures can release reactive materials on other equipment causing corrosion or malfunction

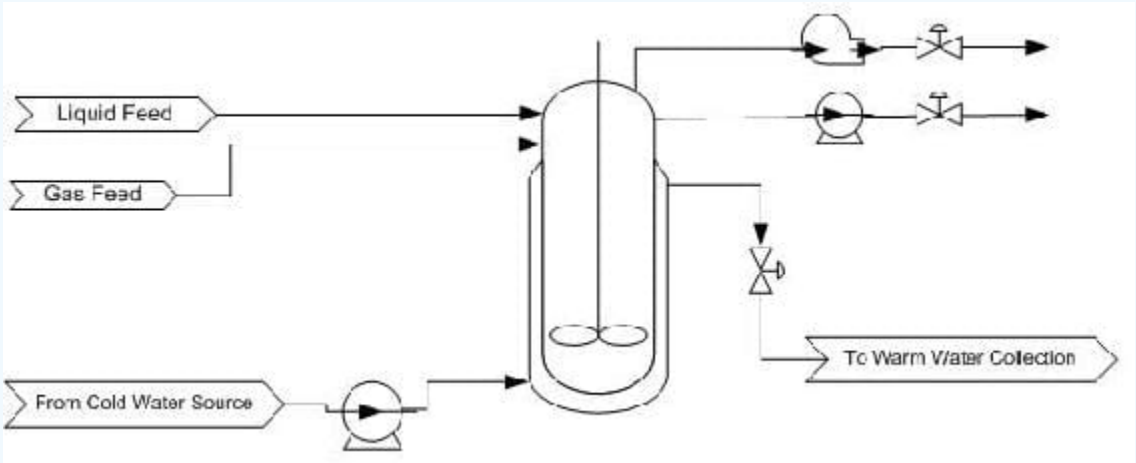
Specific Improvements

This process could be improved and made safer in the following ways.

- Simplification of the piping used to connect the storage tank to the mixing vessel will reduce the risk of leaks and malfunctions due to piping complications.
- The addition of a emergency relief valve and temperature controlling/insulating element to the mixing vessel will greatly reduce the risk of equipment failure due to overheating or overpressureization due to heats of mixing or outside heating influences.
- Use a reactor with a rounded top, as this type can withstand much higher pressures.
- The substitution of carbon tetrachloride with a less hazardous solvent, if the specific solubility required to extract the tin still exists, like cyclohexane, isopropyl alcohol, or 1,1,1-trichloroethane reduces the risk of health complications due to exposure, and possible explosions.

Example 4.5.1

A P&ID appears below for the production of a solution containing a pharmaceutically active compound. The reaction taking place in the CSTR is highly exothermic. After examining the P&ID for this part of the process, describe a possible alarm system.



Solution

The CSTR for the exothermic reaction is jacketed with a cooling water stream. An alarm should be in place to monitor the reactor temperature. A warning alarm can notify the operator that the temperature is too high and corrective action needs to be taken. A critical alarm should be in place to warn that the reactor is nearing runaway conditions and an immediate response is needed. If the necessary action is not taken, systematic shutdown of the reactor could occur. This would involve closing the valves, flooding the jacket with cooling water, and having the impeller on. Another possibility for an alarm, although we do not know how the products are being used specifically, is in a composition measurement of the product containing the pharmaceutically active compound. Depending on where this stream is going and how it is being used, too high a concentration could be dangerous if no other concentration-altering steps occur before the finished product goes out to consumers.

Actual Case

The bulletin below is taken for the US Chemical Safety and Hazard Invesitgation Board

Washington, DC, July 15, 2004 - *The U.S. Chemical Safety and Hazard Investigation Board (CSB) today released a Safety Bulletin calling on chemical plant and refinery operators to exercise appropriate caution when performing work on piping and equipment that could contain hazardous materials.*

The bulletin, set to be considered by the Board at a public meeting today, results from the CSB's investigation into a January 13, 2004, explosion and fire at the Huntsman Petrochemical facility in Port Neches, Texas, where two employees were seriously burned and significant damage occurred to nearby equipment. The explosion and fire occurred as workers attempted to purge a thousand-foot-long chemical process pipe in preparation for a cutting and welding operation.

Huntsman managers and workers were aware of the importance of completely removing hazardous material before cutting into the piping, and relevant warnings were contained in the company's written operating procedures. In this incident, workers first purged the piping with nitrogen to force out residual chemicals, including a hazardous mixture of peroxide and alcohol that reacts violently when heated. But unknown to the workers, the piping included a 300-foot-long section that was three feet lower than the rest of the piping, and despite the nitrogen purge, a significant amount of the hazardous mixture remained trapped. The next step in the operation was to use high-temperature steam to purge the piping of what workers believed would be a small amount of residual flammable hydrocarbon vapor. But the steam heated the peroxide that was trapped in the low section of piping. The peroxide then began to decompose, releasing heat and creating intense pressure. The pressure blew out a valve gasket and violently ruptured the pipe. Flammable vapors shot out of the openings and ignited into a large fireball, injuring plant workers.

After the accident, Huntsman found two drains in the low section of the pipe, which could have been used to remove the trapped liquid. Had Huntsman's procedures called for reviewing plant pipe drawings and physically walking the entire line within the work boundaries, the accident would likely have been avoided, CSB said.

Lessons Learned for this incident:

- *Physically examine all piping and components between isolation devices such as valves, and be sure piping drawings are current*
- *Use the drawings to identify key components, such as low-point drains that can be used to remove dangerous chemicals*
- *Prepare a specific written procedure for removing hazardous material and consider the consequences of working on piping that is not completely purged.*''

References

- AIChE (January 2006). Process Safety Beacon, CEP Magazine
- AIChE (March 2006). Process Safety Beacon, CEP Magazine
- AIChE (May 2006). Process Safety Beacon, CEP Magazine
- AIChE (November 2003). Process Safety Beacon, CEP Magazine
- AIChE (September 2006). Process Safety Beacon, CEP Magazine
- Gael D. Ulrich, Palligarnai T. Vasudevan (2006). Predesign With Safety in Mind, Univ. Of New Hampshire
- Georges A. Melhem, Peter Howell (2005). Designing Emergency Relief Systems for Runaway Reactions, CEP Magazine

- Joseph S. Alford, John Kindervater, Robert Stankovich (April 2005). Alarm Management for Regulated Industries, CEP Magazine
- U.S. Chemical Safety and Hazard Investigation Board, 15 July 2004. CSB News Release
- www.csb.gov/index.cfm?folder=news_releases&page=news&NEWS_ID=171

Contributors and Attributions

- Authors:Bradley Anderson, Aaron Bennick, Michael Salciccioli Stewards: Jocelyn Anleitner, Stephanie Combs, Diane Feldkamp, Heeral Sheth

This page titled 4.5: Safety Features in Piping and Instrumentation Diagrams is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

4.6: Regulatory Agencies and Compliance

Authors: Jayshree Mahajan, Bradley Anderson, Aaron Bennick, Michael Salciccioli
Stewards: Jocelyn Anleitner, Stephanie Combs, Diane Feldkamp, Heeral Sheth

Introduction

Regulatory agencies govern various sections of chemical plants from emissions and water treatment to personnel behaviors and company policy. Understanding of the types of organizations that demand high performance from chemical plants on a national and local level are extremely important for the safety of the employees, community and environment.

Regulatory Compliance

Compliance is an integral part of ensuring the safety of all that work in close contact with the chemical plant as well as minimizing fines and fees that come with violating regulations. Compliance can be organized into two main categories; plant safety and environmental safety. Plant safety ensures employees and the surrounding community are adequately protected during full-operation of the plant. Environmental safety comprises protecting and treating the environmental carefully when it comes to various contaminated streams in the plant.

Plant Safety for Employees and the Community

Safety is paramount in any chemical process, and alarms are essential for compliance with safety constraints. OSHA has established guidelines that must be followed in any plant when dealing with chemicals defined as highly hazardous. Limits for temperature, pressure, flow rates, and compositions need to be regulated. Alarms should be in place to warn operators when a limit is near so that steps can be taken to ensure the safety of people in the plant. An example of this is a safe temperature limit for a CSTR. The alarm signals that the temperature is too high and action is needed to prevent a runaway scenario. If the corrective action is not taken, or not taken quickly enough, a critical alarm can signal a computer program to automatically shut down the entire process or specific unit operation.

Some industries may expect product quality to be closely regulated by the FDA or other government agencies. Typically, this will be in a process where the final product is directly used by people and the margins for error are small. These processes include food processing and manufacturing consumer products, especially pharmaceuticals. These industries usually require systems in place that frequently validate alarms, as well as documentation for all critical alarm events. Measurements such as the weight percent of a pharmaceutically active compound in a solution must be carefully monitored, with recorded uncertainty analysis.

Environmental Safety

Emissions of solids, liquids, and gases in a plant are heavily regulated by government agencies. Regulations apply for processes that emit chemicals to the atmosphere (either directly or following a scrubber), processes that discharge material into a body of water, or processes that require containment control devices like check valves and rupture disks. Alarms are frequently used to comply with these regulations by measuring things such as pH and organic solvent concentration. Typically, a warning alarm will alert personnel that a threshold may be breached if action is not taken, allowing time to avoid an incident requiring a formal report. Critical alarms can alert operators that a threshold has been passed and automatically trigger the appropriate action, such as a systematic shutdown of the process.

Regulatory Agencies

Federal and national agencies maintain smaller state subsections of various programs and administrations. In most cases, the state level requires stricter compliance and lower limits. Solely state-controlled programs usually handle the air and water quality since any regulation violation results in the consequences for the immediate community.

Federal and National Agencies or Programs

Plant safety and environmental safety programs regulated on the national or federal level are monitored by three main agencies; the Environmental Protection Agency (EPA), the U.S. Department of Labor, the Food and Drug Administration, and the Department of Homeland Security. These three governing bodies have created numerous acts, committees, administrations and policies that protect the welfare of the employee, community, and environment.

The EPA at the federal level, provides acts, laws, and regulations, that help maintain and improve the air and water quality. The risk management program (RMP) is a mandatory program that "*require facilities that produce, handle, process, distribute, or store certain chemicals to develop a Risk Management Program*". A Risk Management Plan (RMP) must be submitted to the EPA for approval. Overall, risk management is a large part of process control as control systems must adequately function and maintain compliance of an entire facility. Failure logic for instrumentation, redundant sensors, and critical alarms are essential in maintaining compliance, but a RMP is crucial for handling low-likelihood emergency situations.

The U.S. Department of Labor maintains the Occupational Safety and Health Administration (OSHA) which provides rules and regulations for employers and employees on safe workplace practices. Although individual states may maintain their own occupational health and safety plans, OSHA is the governing body and authority on those programs. Inspections are performed to ensure that all employees have a clean and safe working environment that is hazard-free and risk-mediated. OSHA also maintains the Process Safety Management Program (PSM) which regulates requirements for facilities that handle highly hazardous chemicals. A list of chemicals that qualify include chlorine, formaldehyde, and hydroflouric acid. Requirements of PSM include frequent process hazard analysis (PHA), pre-startup safety review (PSSR), and incident investigations. A cooperative program that OSHA maintains is The Voluntary Protection Programs (VPP). These programs aim to bring together management and labor to provide comprehensive safety and health guidelines and regulations that keep all employees safe while on the job. Each facility must submit application for entry into the program and upon acceptance will follow a set of standards to ensure continued safety. The following is a detailed description of why PSM should be important to chemical engineers, and also some highlights of the main aspects of PSM as required by OSHA.

Process Safety Management (PSM) -- Why is this important to chemical engineers?

See the following links to videos provided by the United States Chemical Safety Board describing a few catastrophic events that have occurred in the chemical engineering industry:

- 2005 Texas City Refinery Explosion
- Bhopal Disaster

Besides the catastrophic nature of events that can occur from neglecting Process Safety, large chemical facilities are granted a privilege and license to operate by the different federal regulatory agencies. If these regulatory agencies perform an audit on a specific facility, and find that their regulations are not being followed, then extremely large fines can be levied to the company, even to the extent of shutting the facility down permanently by removing that facility's privilege to operate. In the case of PSM, it is OSHA who deals out these fines. For example, in 2009 OSHA attempted to levy a record *87 million* dollar fine to an integrated oil company, which has not been finalized in the legal system yet, but gives a good example of how important it is for companies operating in the U.S., if they want to continue to operate safely and economically, to follow all government regulations as closely as possible.

Unexpected releases of toxic, reactive, or flammable liquids and gases in processes involving highly hazardous chemicals have been reported for many years in various industries that use chemicals with such properties. Regardless of the industry that uses these highly hazardous chemicals, there is a potential for an accidental release any time they are not properly controlled, creating the possibility of disaster. As a result of catastrophic incidents in the past, and to help ensure safe and healthful workplaces, OSHA has issued the Process Safety Management of Highly Hazardous Chemicals standard (29 CFR 1910.119), which contains requirements for the management of hazards associated with processes using highly hazardous chemicals. OSHA's standard 29CFR 1910.119 emphasizes the management of hazards associated with highly hazardous chemicals and establishes a comprehensive management program that integrates technologies, procedures, and management practices. A detailed list of these standards can be found on the United States Department of Labor website: www.osha.gov/SLTC/processsafetymanagement/standards.html An effective process safety management program requires a systematic approach to evaluating the whole process. Using this approach the process design, process technology, operational and maintenance activities and procedures, training programs, and other elements which impact the process are all considered in the evaluation. Process safety management is the

proactive identification, evaluation and mitigation or prevention of chemical releases that could occur as a result of failures in process, procedures or equipment. OSHA prescribes essential tools to the success of process safety management including:

- **Process Safety Information**
- **Process Hazard Analysis**
- **Operating Procedures and Practices**
- **Employee Training**
- **Pre-Startup Safety Review**
- **Mechanical Integrity**
- **Management of Change**
- **Incident Investigation**
- **Emergency Preparedness**
- **Compliance Audits**

The thought is, with the simultaneous implementation of all of these things at a facility dealing with large amounts of highly hazardous chemicals, the risk of a catastrophic incident resulting from an unplanned release will be minimized. Following is a detailed discussion of each of these tools prescribed by OSHA.

Process Safety Information (PSI)

Complete, accurate, and up-to-date written information concerning process chemicals, process technology, and process equipment is essential to an effective process safety management program. The compiled information will be a necessary resource to a variety of users including the team that will perform the process hazards analysis, those developing the training programs and operating procedures, contractors whose employees will be working with the process, those conducting the pre-startup safety reviews, local emergency preparedness planners, and insurance and enforcement officials. PSI includes, but is not limited to:

- Material and safety data sheets (MSDS)
- A block flow diagram showing the major process equipment and interconnecting process flow lines
- Process Flow Diagrams (PFDs)
- Piping and Instrument Diagrams (P&IDs)
- Process design information, including the codes and standards relied on to establish good engineering design

Process Hazards Analysis (PHA)

A process hazards analysis (PHA) is one of the most important elements of the process safety management program. A PHA is an organized and systematic effort to identify and analyze the significance of potential hazards associated with the processing and handling of highly hazardous chemicals. A PHA is directed toward analyzing potential causes and consequences of fires, explosions, releases of toxic or flammable chemicals, and major spills of hazardous chemicals. The PHA focuses on equipment, instrumentation, utilities, human actions, and external factors that might impact the process. These considerations assist in determining the hazards and potential failure points or failure modes in a process.

A team from each process unit in the facility will be tasked with conducting a PHA for their process unit at regularly scheduled intervals as defined by OSHA. One example is in an oil refinery, where a PHA has to be conducted and documented for each process unit every five calendar years. The competence of the team conducting the PHA is very important to its success. A PHA team can vary in size from two people to a number of people with varied operational and technical backgrounds. The team leader needs to be fully knowledgeable in the proper implementation of the PHA methodology that is to be used and should be impartial in the evaluation. The other full or part time team members need to provide the team with expertise in areas such as process technology, process design, operating procedures and practices, alarms, emergency procedures, instrumentation, maintenance procedures, safety and health, and any other relevant subject as the need dictates. The ideal team will have an intimate knowledge of the standards, codes, specifications and regulations applicable to the process being studied.

There are various methodologies for conducting a PHA. Choosing which one is right for each individual facility will be influenced by many factors, including the amount of existing knowledge about the process. For more information on the different methodologies for conducting a PHA, see Center for Chemical Process Safety of the American Institute of Chemical Engineers

Operating Procedures

Operating procedures provide specific instructions or details on what steps are to be taken or followed in carrying out the task at hand. The specific instructions should include the applicable safety precautions and appropriate information on safety implications. For example, the operating procedures addressing operating parameters will contain operating instructions about pressure limits, temperature ranges, flow rates, what to do when an upset condition occurs, what alarms and instruments are pertinent if an upset condition occurs, and other subjects. Another example of using operating instructions to properly implement operating procedures is in starting up or shutting down the process.

Operating procedures and instructions are important for training operating personnel. The operating procedures are often viewed as the standard operating practices (SOPs) for operations. Control room personnel and operating staff, in general, need to have a full understanding of operating procedures. In addition, operating procedures need to be changed when there is a change in the process. The consequences of operating procedure changes need to be fully evaluated and the information conveyed to the personnel. For example, mechanical changes to the process made by the maintenance department (like changing a valve from steel to brass or other subtle changes) need to be evaluated to determine whether operating procedures and practices also need to be changed. All management of change actions must be coordinated and integrated with current operating procedures, and operating personnel must be alerted to the changes in procedures before the change is made. When the process is shut down to make a change, the operating procedures must be updated before re-starting the process.

Employee Training

All employees, including maintenance and contractor employees involved with highly hazardous chemicals, need to fully understand the safety and health hazards of the chemicals and processes they work with so they can protect themselves, their fellow employees, and the citizens of nearby communities. Training conducted in compliance with the OSHA Hazard Communication standard (Title 29 Code of Federal Regulations (CFR) Part 1910.1200) will inform employees about the chemicals they work with and familiarize them with reading and understanding MSDSs. However, additional training in subjects such as operating procedures and safe work practices, emergency evacuation and response, safety procedures, routine and non-routine work authorization activities, and other areas pertinent to process safety and health need to be covered by the employer's training program.

In establishing their training programs, employers must clearly identify the employees to be trained, the subjects to be covered, and the goals and objectives they wish to achieve. The learning goals or objectives should be written in clear measurable terms before the training begins. These goals and objectives need to be tailored to each of the specific training modules or segments. Employers should describe the important actions and conditions under which the employee will demonstrate competence or knowledge as well as what is acceptable performance.

Careful consideration must be given to ensure that employees, including maintenance and contract employees, receive current and updated training. For example, if changes are made to a process, affected employees must be trained in the changes and understand the effects of the changes on their job tasks. Additionally, as already discussed, the evaluation of the employee's absorption of training will certainly determine the need for further training.

Pre-Startup Safety Review

For new processes, the employer will find a PHA helpful in improving the design and construction of the process from a reliability and quality point of view. The safe operation of the new process is enhanced by making use of the PHA recommendations before final installations are completed. P&IDs should be completed, the operating procedures put in place, and the operating staff trained to run the process, before startup. The initial startup procedures and normal operating procedures must be fully evaluated as part of the pre-startup review to ensure a safe transfer into the normal operating mode.

For existing processes that have been shut down for turnaround or modification, the employer must ensure that any changes other than "replacement in kind" made to the process during shutdown go through the management of change procedures. P&IDs will need to be updated, as necessary, as well as operating procedures and instructions. If the changes made to the process during shutdown are significant and affect the training program, then operating personnel as well as employees engaged in routine and non-routine work in the process area may need some refresher or additional training. Any incident investigation recommendations, compliance audits, or PHA recommendations need to be reviewed to see what affect they may have on the process before beginning the startup.

Mechanical Integrity

Employers must review their maintenance programs and schedules to see if there are areas where "breakdown" is used rather than the more preferable on-going mechanical integrity program. Equipment used to process, store, or handle highly hazardous chemicals has to be designed, constructed, installed, and maintained to minimize the risk of releases of such chemicals. This requires that a mechanical integrity program be in place to ensure the continued integrity of process equipment.

Elements of a mechanical integrity program include identifying and categorizing equipment and instrumentation, inspections and tests and their frequency; maintenance procedures; training of maintenance personnel; criteria for acceptable test results; documentation of test and inspection results; and documentation of manufacturer recommendations for equipment and instrumentation.

Management of Change

To properly manage changes to process chemicals, technology, equipment and facilities, one must define what is meant by change. In the process safety management standard, change includes all modifications to equipment, procedures, raw materials, and processing conditions other than "replacement in kind." These changes must be properly managed by identifying and reviewing them prior to implementing them. For example, the operating procedures contain the operating parameters (pressure limits, temperature ranges, flow rates, etc.) and the importance of operating within these limits. While the operator must have the flexibility to maintain safe operation within the established parameters, any operation outside of these parameters requires review and approval by a written management of change procedure. Management of change also covers changes in process technology and changes to equipment and instrumentation. Changes in process technology can result from changes in production rates, raw materials, experimentation, equipment unavailability, new equipment, new product development, change in catalysts, and changes in operating conditions to improve yield or quality. Equipment changes can be in materials of construction, equipment specifications, piping pre-arrangements, experimental equipment, computer program revisions, and alarms and interlocks. Employers must establish means and methods to detect both technical and mechanical changes.

Temporary changes have caused a number of catastrophes over the years, and employers must establish ways to detect both temporary and permanent changes. It is important that a time limit for temporary changes be established and monitored since otherwise, without control, these changes may tend to become permanent. Temporary changes are subject to the management of change provisions. In addition, the management of change procedures are used to ensure that the equipment and procedures are returned to their original or designed conditions at the end of the temporary change. Proper documentation and review of these changes are invaluable in ensuring that safety and health considerations are incorporated into operating procedures and processes. Employers may wish to develop a form or clearance sheet to facilitate the processing of changes through the management of change procedures. A typical change form may include a description and the purpose of the change, the technical basis for the change, safety and health considerations, documentation of changes for the operating procedures, maintenance procedures, inspection and testing, P&IDs, electrical classification, training and communications, pre-startup inspection, duration (if a temporary change), approvals, and authorization. Where the impact of the change is minor and well understood, a check list reviewed by an authorized person, with proper communication to others who are affected, may suffice. For a more complex or significant design change, however, a hazard evaluation procedure with approvals by operations, maintenance, and safety departments may be appropriate. Changes in documents such as P&IDs, raw materials, operating procedures, mechanical integrity programs, and electrical classifications should be noted so that these revisions can be made permanent when the drawings and procedure manuals are updated. Copies of process changes must be kept in an accessible location to ensure that design changes are available to operating personnel as well as to PHA team members when a PHA is being prepared or being updated.

Incident Investigation

Incident investigation is the process of identifying the underlying causes of incidents and implementing steps to prevent similar events from occurring. The intent of an incident investigation is for employers to learn from past experiences and thus avoid repeating past mistakes. The incidents OSHA expects employers to recognize and to investigate are the types of events that resulted in or could reasonably have resulted in a catastrophic release. These events are sometimes referred to as "near misses," meaning that a serious consequence did not occur, but could have.

Employers must develop in-house capability to investigate incidents that occur in their facilities. A team should be assembled by the employer and trained in the techniques of investigation including how to conduct interviews of witnesses, assemble needed documentation, and write reports. A multi-disciplinary team is better able to gather the facts of the event and to analyze them and develop plausible scenarios as to what happened, and why. Team members should be selected on the basis of their training,

knowledge and ability to contribute to a team effort to fully investigate the incident.

Emergency Preparedness

Each employer must address what actions employees are to take when there is an unwanted release of highly hazardous chemicals. Emergency preparedness is the employer's third line of defense that will be relied on along with the second line of defense, which is to control the release of chemical. Control releases and emergency preparedness will take place when the first line of defense to operate and maintain the process and contain the chemicals fails to stop the release.

Employers will need to select how many different emergency preparedness or third lines of defense they plan to have, develop the necessary emergency plans and procedures, appropriately train employees in their emergency duties and responsibilities, and then implement these lines of defense. Employers, at a minimum, must have an emergency action plan that will facilitate the prompt evacuation of employees when there is an unwanted release of a highly hazardous chemical. This means that the employer's plan will be activated by an alarm system to alert employees when to evacuate, and that employees who are physically impaired will have the necessary support and assistance to get them to a safe zone. The intent of these requirements is to alert and move employees quickly to a safe zone. Delaying alarms or confusing alarms are to be avoided. The use of process control centers or buildings as safe areas is discouraged. Recent catastrophes indicate that lives are lost in these structures because of their location and because they are not necessarily designed to withstand overpressures from shock waves resulting from explosions in the process area.

Preplanning for more serious releases is an important element in the employer's line of defense. When a serious release of a highly hazardous chemical occurs, the employer, through preplanning, will have determined in advance what actions employees are to take. The evacuation of the immediate release area and other areas, as necessary, would be accomplished under the emergency action plan. If the employer wishes to use plant personnel-such as a fire brigade, spill control team, a hazardous materials team-or employees to render aid to those in the immediate release area and to control or mitigate the incident, refer to OSHA's Hazardous Waste Operations and Emergency Response (HAZWOPER) standard (Title 79 CFR Part 1910.1 20). If outside assistance is necessary, such as through mutual aid agreements between employers and local government emergency response organizations, these emergency responders are also covered by HAZWOPER. The safety and health protection required for emergency responders is the responsibility of their employers and of the on-scene incident commander.

Compliance Audits

An audit is a technique used to gather sufficient facts and information, including statistical information, to verify compliance with standards. Employers must select a trained individual or assemble a trained team to audit the process safety management system and program. A small process or plant may need only one knowledgeable person to conduct an audit. The audit includes an evaluation of the design and effectiveness of the process safety management system and a field inspection of the safety and health conditions and practices to verify that the employer's systems are effectively implemented. The audit should be conducted or led by a person knowledgeable in audit techniques who is impartial towards the facility or area being audited. The essential elements of an audit program include planning, staffing, conducting the audit, evaluating hazards and deficiencies and taking corrective action, performing a follow-up review, and documenting actions taken.

Other Federal Entities

The Food and Drug Administration (FDA) is an agency of the United States Department of Health and Human Services. The FDA regulates food, drugs, cosmetics, biologics, medical devices, radiation-emitting devices and vetenary products manufactured in the United States. The main goal of the FDA is to maitain that the products they regulate are safe, effective, and secure. The FDA is also responsible that the products are accurately represented to the public. State and local governments also help regulate these products in cooperation with the FDA. The FDA does not regulate alcohol, illegal drugs, and meat and poultry.

The Department of Homeland Security has recently taken a role in regulating chemical plants because of 9/11. Chemical plants are seen by the government as targets for terrorists and security in and around the plant is a major concern. Although the laws are typically state run, the Department of Homeland Security has required mandatory national security standard to chemical plants throughout the nation. Although the mandates were fought by legislative for years, the Department of Homeland Security has influence in the security in chemical plants. The law requires the plant to prepare a vulnerability test and submit a site security plan. In order to validate these activities, audits and site visits are/will be performed by government officials.

State Regulations

State regulations vary greatly from state-to-state depending on the main concern. For instance, beach quality is more important in California than Nebraska due to the geographical location. Similarly state legislative acts and administrations can expect high performance from the chemical industry as it has the capability to affect all parts of a community for many generations.

References

- U.S. National Profile on the Management of Chemicals. www.epa.gov/oppfead1/profile/
- U.S. Department of Labor. Occupational Safety & Health Administration. Process Safety Management. www.osha.gov/SLTC/processsafetymanagement/index.html
- U.S. Department of Labor. Occupational Safety & Health Administration. State Occupational Health and Safety Plans. www.osha.gov/dcsp/osp/index.html
- U.S. Department of Labor. Voluntary Protection Programs. www.osha.gov/dcsp/vpp/index.html
- Environmental Protection Agency. Regulatory Information by Business Sector. www.epa.gov/lawsregs/bizsector/index.html
- Environmental Protection Agency. Regulatory Information by Environmental Topic. www.epa.gov/lawsregs/envtopics/index.html#air
- Environmental Protection Agency. State Specific Regulatory Information Search Engine. www.epa.gov/lawsregs/where/index.html
- Environmental Protection Agency. Risk Management Plan. http://www.epa.gov/oem/content/rmp/index.htm
- United States Chemical Safety and Hazard Regulation Board. Homepage. www.chemsafety.gov/
- U.S. Department of Health & Human Services. Food and Drug Administration. http://www.fda.gov/
- United States Chemical Safety Board (CSB). http://www.csb.gov
- Center for Chemical Process Safety (CCPS). http://www.aiche.org/ccps

This page titled 4.6: Regulatory Agencies and Compliance is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

5: LOGICAL MODELING

5.1: Boolean Models - Truth Tables and State Transition Diagrams
5.2: Logical Control Programs - IF... THEN... WHILE...

This page titled 5: Logical Modeling is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

5.1: Boolean Models - Truth Tables and State Transition Diagrams

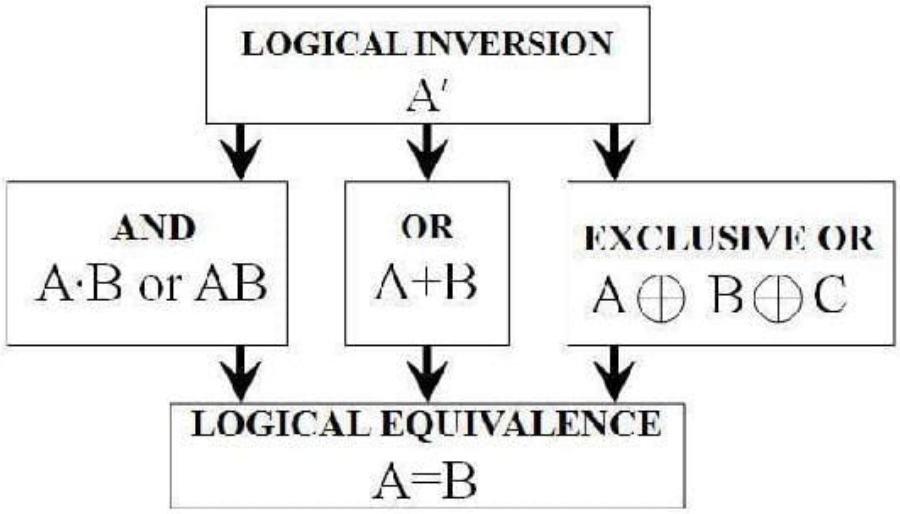
Introduction to Boolean Models

A Boolean is a variable that can only attain two values: True or False. In most applications, it is convenient to represent a True by the number 1, and a False by the number 0. A **Boolean model**, or **Boolean network**, is a collection of Boolean variables that are related by logical switching rules, or **Boolean functions**, that follow an If-Then format. This type of Boolean model is known as an autonomous model and will be the primary type of model discussed in this article.

In chemical engineering, Boolean models can be used to model simple control systems. Boolean functions can be used to model switches on pumps and valves that react to readings from sensors that help keep that system operating smoothly and safely. A simple application for level control of a CSTR is included in worked-out example 1. In this example, Boolean function is used to close the inlet stream and open the outlet stream when the level is higher than a specified point.

Boolean Functions

Boolean functions are logical operators that relate two or more Boolean variables within a system and return a true or false. A Boolean expression is a group of Boolean functions, which will be described individually below. When computing the value of a Boolean expression, Parentheses are used to indicate priority (working from inside out as in algebra). After that, LOGICAL INVERSION will always be first and LOGICAL EQUIVALENCE will be last, but the order of operation for the AND, OR, and EXCLUSIVE OR functions are specified with parenthesis.



Descriptions and examples of these functions are given below. A quick reference of each of the functions can be found after the examples.

LOGICAL INVERSION

LOGICAL INVERSION is a function that returns the opposite value of a variable. The function is denoted as a prime on the variable (e.g. A' or B') For example, if we say that A is true (A=1), then the function A' will return a false (A'=0). Similarly, if we say that A is false (A=0) then the function A' will return true (A'=1).

Example:

A=1, B=A' then B=0

AND

The AND function relates two or more Boolean variables and returns a true if-and-only-if both variables are true. A dot is used to denote the AND function, or it is simply omitted. For example "A and B" can be written as "A•B" or as "AB." In this example, the AND function will only return a true if-and-only-if both Boolean variables A and B have a value of 1.

Examples:

Variables	Results
A=1, B=1	AB = 1
A=1, B=0	AB = 0
A=1, B=1, C=1	ABC = 1
A=1, B=0, C=1	ABC = 0

OR

The OR function relates two or more Boolean variables and returns a true if any referenced variables are true. A plus is used to denote the OR function. For example "A or B" can be written as "A+B." In this example, the OR function will return true if either Boolean variable, A or B, has a value of 1.

Examples:

Variables	Results
A=1, B=1	A+B = 1
A=1, B=0	A+B = 1
A=0, B=0	A+B = 0
A=0, B=0, C=1	A+B+C = 1
A=0, B=0, C=0	A+B+C = 0

EXCLUSIVE OR

The EXCLUSIVE OR function relates two or more Boolean variables and returns **true** only when **one** of the variables is true and **all** other variables are false. It returns **false** when **more than one** of the variables are true, or **all** the variables are false. A circumscribed plus is used to denote the EXCLUSIVE OR function. For example "A EXCLUSIVE OR B" can be written as "A $\oplus$ B."

Examples:

Variables	Results
A=1, B=1	A $\oplus$ B = 0
A=1, B=0	A $\oplus$ B = 1
A=0, B=1	A $\oplus$ B = 1
A=0, B=0	A $\oplus$ B = 0
A=0, B=0, C=0	A $\oplus$ B $\oplus$ C = 0

Variables	Results
A=1, B=0, C=0	$A \oplus B \oplus C = 1$
A=1, B=0, C=1	$A \oplus B \oplus C = 0$
A=1, B=1, C=1	$A \oplus B \oplus C = 0$

LOGICAL EQUIVALENCE

The LOGICAL EQUIVALENCE function equates two Boolean variables or expressions. The LOGICAL EQUIVALENCE function, denoted as =, assigns a Boolean variable a true or false depending on the value of the variable or expression that it is being equated with. For example, A LOGICAL EQUIVALENCE B can be written as A = B. In this example, the value of A will be assigned the value of B.

Boolean Networks

As stated in the introduction, a Boolean network is a system of boolean equations. In chemical engineering, Boolean networks are likely to be dependant on external inputs as a means of controlling a physical system. However, the following sections pertain mostly to synchronous autonomous systems. An autonomous system is one that is completely independent of external inputs. Every Boolean variable is dependent on the state of other Boolean variables in the system and no variable is controlled by an external input. A synchronous system is one that logical switching (the changing of Boolean variables) occurs simultaneously for all variables based on the values prior to the incidence of change.

Here is an example of an autonomous boolean network:

Boolean Functions
A = B+C'
B = AC
C = A'

Truth Tables

A truth table is a tabulation of all the possible states of a Boolean Model at different time frames. A simple truth table shows the potential initial states at time, T_i, and the corresponding subsequent states at time T_{i+1}, of a Boolean network. Truth tables can provide one with a clearer picture of how the rules apply and how they affect each situation. Hence, they help to ensure that each output only has one control statement so that the Boolean rules do not conflict with each other.

Constructing Truth Tables

1. Draw up a table with the appropriate number of columns for each variable; one for each input and output.
2. The left side of the column should contain all possible permutations of the input variables at time T_i. One method to accomplish this might be to list all possible combinations in ascending binary order.
3. The right side of the column should contain the corresponding outcome of the output variables at the subsequent time T_{i+1}. A generic example of this with 2 variables can be seen below:

Inputs		Outputs	
T_i		T_{i+1}	
variable1	variable2	variable1	variable2

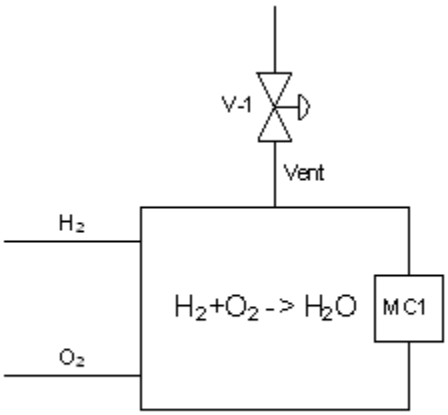
A quick way to check that you have all of the possible permutations is that there should be 2^X possible permutations for X input variables.

Example of a Truth Table

The sample system we will be using is based on hydrogen fuel cell technology. The equation for the operation of hydrogen fuel cells is

$$\mathrm{H_2 + O_2 \longrightarrow H_2O.}$$

One aspect of **Proton Exchange Membrane** (PEM) fuel cells (a type of fuel cell) is that the performance of the fuel cell is highly dependent on the relative humidity of the system (if humidity rises too high, the fuel cell will flood and H_2 and O_2 will be unable to reach the cell. If humidity falls too low, the fuel cell will dry up and the performance will drop.) The task is to create a Boolean model for this simplified water management system.



PEM Fuel Cell System

The system produces steam within the system, and there is a vent to release steam if the system becomes too saturated. In our system, we will assume that the inputs are stoichiometric and react completely. Also we will assume that pressure buildup from steam is negligible compared to the change in relative humidity. The only variable in question is the %relative humidity in the system.

- note: this is not how water management actually works in a fuel cell system, but it is a simple example.

A will represent the moisture controller response (0 indicates relative humidity or %RH < 80%, 1 indicates %RH >80%)
B will represent the valve status (0 is closed, 1 is open)

The corresponding Boolean functions for this model are given below (normally you would have to design these yourself to meet the criteria you desire):

A = A
B = A

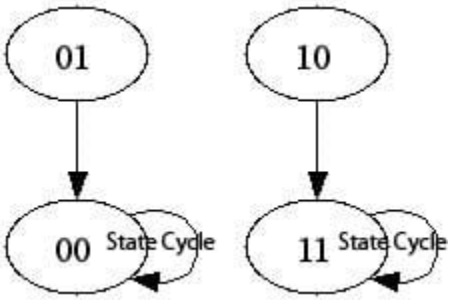
For this example with 2 input variables, there are 2^2 = 4 possible permutations and 2^2 = 4 rows. The resultant permutations for the outputs are: For A where Y=1, the number of 0s and 1s are $2^{(Y-1)}=2^{(1-1)}=1$. For B where Y=2, the number of 0s and 1s are $2^{(Y-1)}=2^{(2-1)}=2$.

The resultant truth table is below:

Inputs		Outputs	
T_i		T_{i+1}	
A	B	A	B
0	0	0	0
1	0	1	1
0	1	0	0
1	1	1	1

State Transition Diagrams

A state transition diagram is a graphical way of viewing truth tables. This is accomplished by looking at each individual initial state and its resultant state. The transition from one state to another is represented by an arrow. Then they are pieced together like a jigsaw puzzle until they fit in place. When one state leads to itself it simply points to itself. The following example is based on the truth table in the previous section.

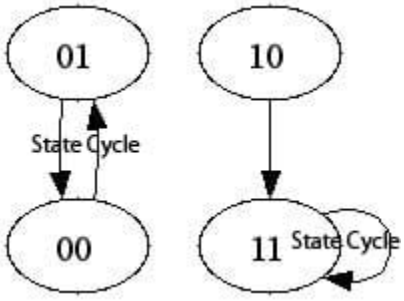


In this example, there are two state cycles. A state cycle is a combination of states around which the system continually enters and reenters. For a finite number of states, there will always exist at least one state cycle. A state cycle is also a pathway or a flowchart that shows the "decision making process" of a Boolean network. This feature is a direct result from two attributes of Boolean networks:

1. Finite number of states
2. Deterministic (there is a certain set of rules that determines the next state that will be entered)

In the example presented in the previous section, there were two state cycles. One advantage of state cycles is it easily allows you to see where your model will end up cycling and if there are any states that are not accounted for properly by your model. In the previous diagram, if the moisture controller indicated the humidity was below the set value, it would close the valve or hold the valve closed. If the moisture controller indicated that the humidity was above the set value, it would either open the valve or hold it open.

Consider this alternate system.



In this example, the state cycle says that if the meter says that the humidity is below the set point it would cycle the vent valve

open and closed. This would hurt the system and is not a desired outcome of the model.

For safety and functionality issues, a process control engineer would want to consider all possiblities in the design of any Boolean network modeling a real system.

Limitations of Boolean Networks

Advantages of Boolean Networks

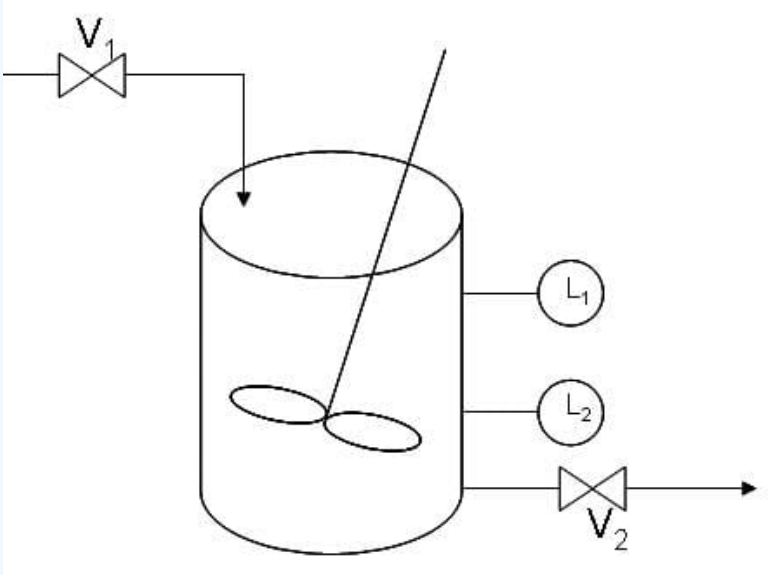
- Unlike ordinary differential equations and most other models, Boolean networks do not require an input of parameters.
- Boolean models are quick and easy to compute using computers.
- Boolean networks can be used to model a wide variety of activities and events.
- Boolean networks can be used to approximate ordinary differential equations when there are an infinite number of states.

Disadvantages of Boolean Networks

- Boolean networks are restrained to computing very simple math. They cannot be used for calculus and to calculate large quantities.
- Boolean models have relatively low resolution compared to other models.
- It is very time consuming and complicated to build Boolean networks by hand.

Example 5.1.1

A hypothetical CSTR needs to have its liquid level maintained below a safety mark by means of a sensor, L1, on the corresponding mark and a control valve placed on the inlet and outlet streams – V1 and V2 respectively. A typical application of the afore-mentioned system could involve heterogeneously catalyzed liquid reaction(s) with liquid product(s).



Solution

Conventions

Water level sensor

L1	0	1

water level	desirable	too high

Valve

V	0	1
position	closed	open

Initial State

Assume that the CSTR is empty and being filled up. CSTR, being empty, sets the value of L1 to zero. Filling up the CSTR could be done by opening valve 1 - V1 assuming a value of one - and closing valve 2 - V2 assuming a value of zero.

In coordinate form, the initial state is as such: (L1, V1, V2) = (0, 1, 0)

Problem Interpretation

Let h be the water level and WL1 be the safety mark defined in the CSTR. The system could assume one of the following states at any one time:

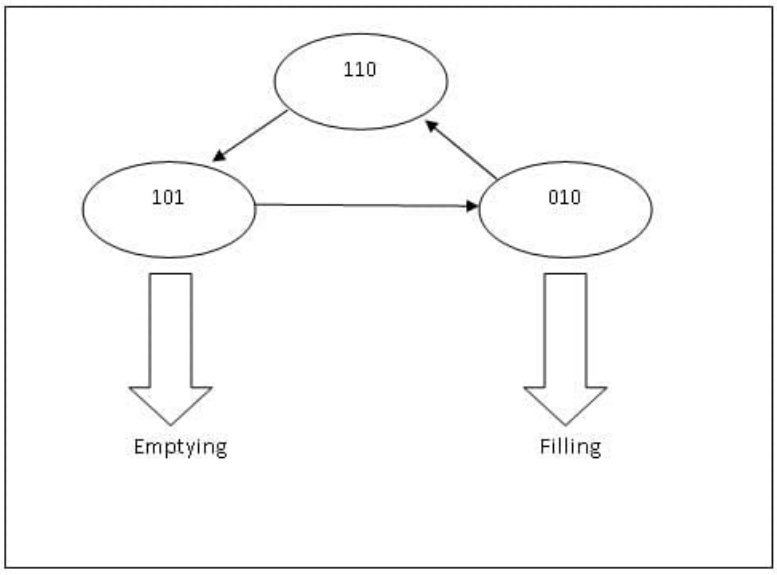
1) h < WL1 : desirable water level

Maximizing production of the chemical prompts the system to remain in its current state - that is, its initial state. (L1, V1, V2)final = (0, 1, 0) **final state**

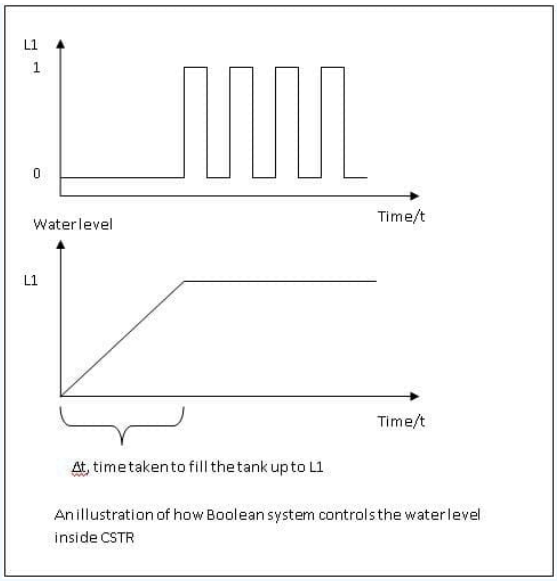
2) h > WL1 : water level too high

Prevention of flooding requires that the tank be emptied. As such, valve 1 (V1) should be closed to stop the input while valve 2 (V2) should be open to empty the extra water above the safety water mark. (L1, V1, V2)' = (1, 1, 0) **trigger to valve** (L1, V1, V2)final = (1, 0, 1) **final state**

State Cycle



Physical Significance



An illustration of how Boolean system controls the water level inside CSTR

References

- James E. Palmer and David E. Perlman (1993). *Schaum's Outline of Theory and Problems of Introduction to Digital Systems*, McGraw-Hill Professional. ISBN 0070484392
- Stuart A. Kauffman (1993). *The Origins of Order Self-Organization and Selection in Evolution*, Oxford University Press. ISBN 0195079515

Contributors and Attributions

- Joseph Casler, Andry Haryanto, Seth Kahle and Weiyin Xu, Adhi Paisoseputra, Andrew Kim, Hillary Kast, Stephanie Cleto

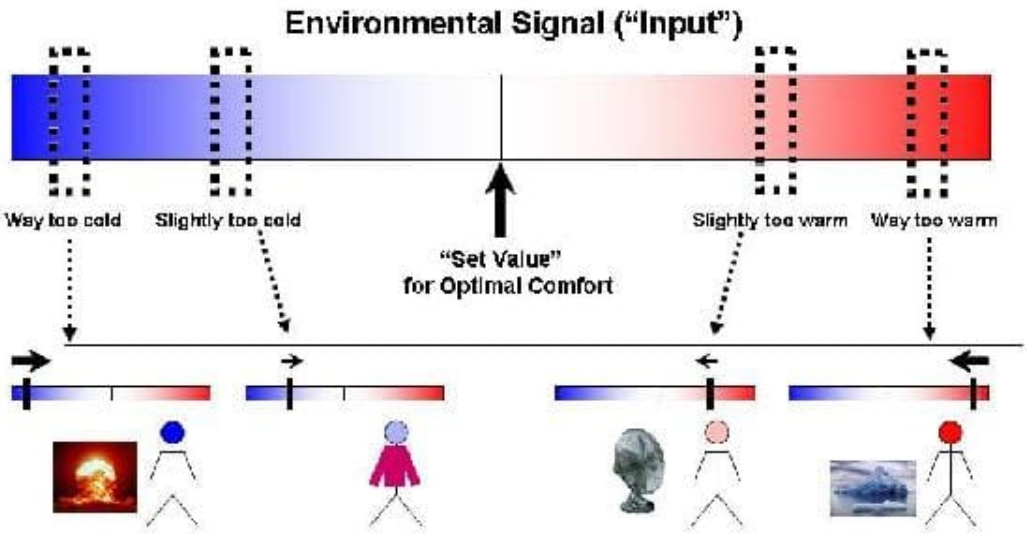
This page titled 5.1: Boolean Models - Truth Tables and State Transition Diagrams is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

5.2: Logical Control Programs - IF... THEN... WHILE...

Introduction

A logical control program is a set of conditional statements describing the response of a controller to different inputs. A controller is a computer used to automate industrial processes. Process engineers use control logic to tell the controller in a process how to react to all inputs from sensors with an appropriate response to maintain normal functioning of the process. Control logic (sometimes called process logic) is based on simple logic principles governed by statements such as IF X, THEN Y, ELSE Z yet can be used to describe a wide range of complex relationships in a process. Although different controllers and processes use different programming languages, the concepts of control logic apply and the conditions expressed in a logical control program can be adapted to any language.

The concepts behind logical control programs are not only found in chemical processes; in fact, control logic is used in everyday life. For example, a person may regulate his/her own body temperature and comfort level using the following conditional logic statements: IF the temperature is slightly too warm, THEN turn on a fan; IF the temperature is way too warm, THEN turn on the air conditioning; IF the temperature is slightly too cold, THEN put on a sweatshirt; IF the temperature is way too cold, THEN turn on the fireplace. The person takes an input from the environment (temperature) and if it meets a certain prescribed condition, she executes an action to keep herself comfortable. Similarly, chemical processes evaluate input values from the process against set values to determine the necessary actions to keep the process running smoothly and safely (aforementioned example illustrated below).



The following sections elaborate on the construction of conditional logic statements and give examples of developing logical control programs for chemical processes.

Logic Controls

Logical controls (IF, THEN, ELSE, and WHILE) compare a value from a sensor to a set standard for the value to evaluate the variable as True/False in order to dictate an appropriate response for the physical system. The control program for a chemical process contains many statements describing the responses of valves, pumps, and other equipment to sensors such as flow and temperature sensors. The responses described by the system can be discrete, such as an on/off switch, or can be continuous, such as opening a valve between 0 and 100%. The goal of a control program is to maintain the values monitored by the sensors at an acceptable level for process operation considering factors like product quality, safety, and physical limitations of the equipment. In addition to describing the normal activity of the process, a control program also describes how the process will initialize at the start of each day and how the controller will respond to an emergency outside of the normal operating conditions

of the system. Unlike a linear computer program, logic programs are continuously monitoring and responding without a specific order. Before constructing logical control programs, it is important to understand the conditional statements, such as IF-THEN and WHIlE statements, that govern process logic.

The following logic controls found below are written in pseudocode. Pseudocode is a compact and informal way of writing computer program algorithms. It is intended for human reading instead of machine reading, and does not require stringent syntax for people to understand. Pseudocode is typically used for planning computer program development and to outline a program before actual coding occurs.

IF-THEN statements

IF-THEN statements compare a value from a sensor to a set value and describe what should happen if the relationship holds. The IF-THEN statement takes the form IF X, THEN Y where X and Y can be single variables or combinations of variables. For example, consider the following statements 1 and 2. In statement 1, both X and Y are single variables whereas in statement 2, X is a combination of two variables. The ability in conditional logic to combine different conditions makes it more flexible than incidence graphs, which can only describe monotonic relationships between two variables (See incidence graphs). A monotonic relationship is one where if X is increasing, Y is always decreasing or if X is increasing, Y is always increasing. For complex processes, it is important to be able to express non-monotonic relationships.

1. IF T>200 C, THEN open V1
2. IF T> 200 C and P> 200 psi, THEN open V1.

Where T is Temperature, P is pressure, and V represents a valve.

In statement 1, if the temperature happens to be above 200 C, valve 1 will be opened.
In statement 2, if the temperature is above 200 C and the pressure is above 200 psi, then the valve will be opened.

Otherwise, no action will be taken on valve 1

If the conditions in the IF statement are met, the THEN statement is executed, and depending on the command, the physical system is acted upon. Otherwise, no action is taken in response to the sensor input. In order to describe an alternate action if the IF condition does not hold true, ELSE statements are necessary.

ELSE statements

The simple form of an IF-THEN-ELSE statement is IF X, THEN Y, ELSE Z where again X, Y, and Z can be single variables or combinations of variables (as explained in the IF-THEN section above). The variable(s) in the ELSE statement are executed if the conditions in the IF statement are not true. This statement works similar to the IF-THEN statements, in that the statements are processed in order. The ELSE statement is referred to last, and is a condition that is often specified to keep the program running. An example is the following:

IF P>200 psi, THEN close V1
ELSE open V4

In this statement, if the pressure happens to be 200psi or less, the THEN statement will be skipped and the ELSE statement will be executed, opening valve 4.

Sometimes, if X, Y or Z represent many variables and several AND or OR statements are used, a WHILE statement may be employed.

CASE statements

CASE statement is an alternative syntax that can be cleaner than many IF..THEN and ELSE statements. The example shown in the table below shows its importance.

CASE	IF..THEN and ELSE

T>Tset+2: v2=v2+0.1 T>Tset+1: v2=v2+0.05 T v2=v2-0.05 T v2=v2-0.1	IF T>Tset+1: IF T>Tset+2: v2=v2+0.1 ELSE: v2=v2+0.05 ELSE IF T IF T v2=v2-0.1 ELSE: v2=v2-0.05

Thus CASE statements make the code easier to read for debugging.

WHILE statements

The WHILE condition is used to compare a variable to a range of values. The WHILE statement is used in place of a statement of the form (IF A>B AND IF A

Example:
A tank that is initially empty needs to be filled with 1000 gallons of water 500 seconds after the process has been started-up. The water flow rate is exactly 1 gallon/second if V1 is completely open and V1 controls the flow of water into the tank.

Using a IF...THEN statement the program could be written as follows:
IF t > 500 and t < 1501 THEN set V1 to open
ELSE set V1 to close.

The WHILE statement used to describe this relationship is as follows:

WHILE 500 < t < 1501 set V1 to open
ELSE set V1 to close.

It may not seem like much of a change between the two forms of the code. However, this is a very simple model. If modeling a process with multiple variables you could need many IF...THEN statements to write the code when a single WHILE condition could replace it.

Example:
V1 controls reactants entering a reactor that can only run safely if the temperature is under 500K.

The WHILE can be used to control the process as follows:
WHILE T < 500 set V1 to open
ELSE set V1 to close. ALARM.

This example shows how a WHILE condition can be used as a safety measure to prevent a process from becoming unstable or unsafe.

In addition to lists of IF-THEN-ELSE-WHILE statements, control logic can be alternately represented by truth tables and state transition diagrams. Truth tables show all the possible states of a model governed by conditional statements and state transition diagrams represent truth tables graphically. Oftentimes, these are used in conjunction with booleans, variables that can only have

two values, TRUE OR FALSE. A Boolean model or Boolean function follows the format of IF-THEN statements described here. A description of Boolean models, truth tables, and state transition diagrams is given here.

GO TO statements

The GO TO statement helps to break out of current run to go to a different configuration. It can be an important operator in logical programming because a lot of common functions are accessed using GO TO. However, many programmers feel that GO TO statements should not be used in programming since it adds an extra and often unnecessary level of complex that can make the code unreadable and hard to analyze. Even though the GO TO operator has its downsides, it is still an important operator since it can make help to simplify basic functions. It can simplify code by allowing for a function, such as a fail safe, be referenced multiple times with out having to rewrite the function every time it is called. The GO TO operator is also important because even advanced languages that do not have a GO TO function often have a different operator that functions in a similar manner but with limitations. For example, the C, C++ and java languages each have functions break and continue which are similar to the GO TO operator. Break is a function that allows the program to exit a loop before it reaches completion, while the continue function returns control to the loop without executing the code after the continue command. A function is a part of code within a larger program, which performs a specific task and is relatively independent of the remaining code. Some examples of functions are as follows:

FUNCTION INITIALIZE: It runs at the beginning of a process to make sure all the valves and motor are in correct position. The operation of this function could be to close all valves, reset counters and timers, turn of motors and turn off heaters.

FUNCTION PROGRAM: It is the main run of the process.

FUNCTION FAIL SAFE: It runs only when an emergency situation arises. The operation of this function could be to open or close valves to stop the system, quench reactions via cooling, dilution, mixing or other method.

FUNCTION SHUTDOWN: It is run at the end of the process in order to shutdown.

FUNCTION IDLE: It is run to power down process.

All the functions mentioned above except FUNCTION IDLE are used in all chemical processes.

ALARM statements

The ALARM statement is used to caution the operators in case a problem arises in the process. Alarms may not be sufficient danger to shut down the process, but requires outside attention. Some example when ALARM statements are used in the process are as follows:

- If the storage tank of a reactant is low, then ALARM.
- If the pressure of a reactor is low, then ALARM.
- If no flow is detected even when the valve is open, then ALARM.
- If the temperature of the reactor is low even after heating, then ALARM.
- If redundant sensors disagree, then ALARM.

In conclusion ALARM functions are very important in order to run a process safely.

Control Language in Industry

As stated before, these commands are all in pseudocode, and not specific to any programming language. Once the general structure of the controller is determined, the pseudocode can be coded into a specific programming language. Although there are many proprietary languages in industry, some popular ones are:

- Visual Basic
- C++
- Database programming (ex. Structured Query Language/SQL)
- Pascal
- Fortran

Pascal and Fortran are older languages that many newer languages are based on, but they are still used with some controllers, especially in older plants. Any experience with different computer languages is a definite plus in industry, and some chemical engineers make the transition into advanced controls designing, writing, and implementing code to make sure a plant keeps running smoothly.

Logical Functions in Microsoft Excel

Microsoft Excel has basic logical tools to help in constructing simple logical statements and if needed more complex logical systems. A list of the functions is shown below.

TRUE().....Returns the logical value, TRUE.

FALSE().....Returns the logical value, FALSE.

AND(logical_expression_A,B,C).....Returns TRUE if all the expressions are true.

OR(logical_expression_A,B,C).....Returns TRUE if one of the expressions are true.

NOT(logical_expression).....Returns the opposite of the expected logical value. If the expression is TRUE, it will return FALSE

IFERROR(value,value_if_error).....Returns the value unless there is an error in which it will return the value_if_error.

IF(logical_expression,value_if_true,value_if_false).....Checks the validity of the expression and returns TRUE or FALSE likewise.

The IF() function will be most useful in logical programing. It is essentially an IF THEN or IF ELSE function returning one of two values based on the logical expression it is testing. Excel also allows logical functions within functions. This allows for logical expressions involving more than one IF statement within itself for example. Use of these tools will be practical in quickly setting up control programs and other systems in Microsoft Excel.

Constructing a Logical Control Program

Understanding the conditional statements used in control logic is the first step in constructing a logical control program. The second step is developing a thorough understanding of the process to be controlled. Knowledge of the equipment, piping, and instrumentation (contained on a P&ID diagram), operating conditions, chemical compounds used, and safety concerns is necessary. Particularly it is important to know the measured and controlled variables. For example, the pressure limits of a tank must be known in order to develop a control plan to ensure safety; ignoring this constraint could lead to explosion and injury. Once the necessary controls are known, one can develop a plan using the logical statements described previously. The third step is constructing a logical control program is understanding that there is not always a right answer, meaning there are many different ways to ensure the same desired outcome.

Worked out Examples 1, 2, and 3 demonstrate the construction of simple logical control programs. The more complex the situation, the longer the logical control plan becomes yet the process is still the same. An example of a more complex logical control program is given here.This example is from the 2005 ChE 466 class at the University of Michigan and describes an entire chemical process from the delivery of raw materials to the output of the final product.

Determining Fail Safe Conditions

Fail safe is the practice of designing a system to default to safe conditions if anything or everything goes wrong. The goals of fail safe conditions are to:

- Protect plant personal
- Protect the local community around the plant
- Protect the environment
- Protect plant equipment

In order to establish safe conditions, fail safe programs must specify the desired positions of all valves and status of all motors and controlled equipment. For example, in an exothermic reactor, fail safe conditions would specify opening all cooling water

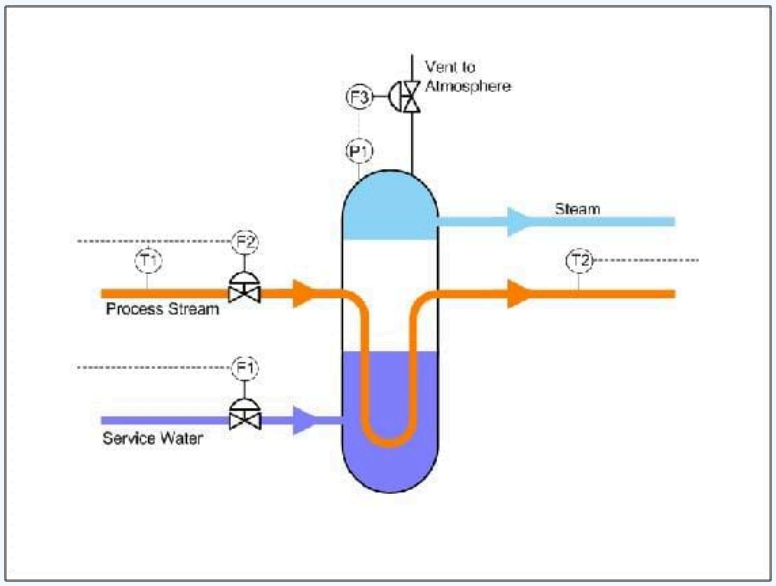
valves, closing all feed valves, shutting off feed pump motors, turning on agitator motor, and open all vent valves.

Control programs frequently define fail safe conditions at the beginning of the program. These conditions are then activated using a GO TO command when process conditions exceed the maximum or fall below the minimum allowable values.

All processes must be evaluated for conditions that could cause hazards and fail safe procedure must be designed counteract the effects.

Example 5.2.1: Reboiler

Reboilers are used in industry to cool down process streams by creating steam from water. This chemical process involves a phase change from liquid (water) to gas (steam) and it is important to monitor flowrates, temperatures, and pressures. Below is a diagram of a reboiler. The controlled variables are F1, F2, and F3; these are controlled by manipulating the corresponding valves. The measured variables are P1, T1, and T2. Considering the operating conditions and constraints given below, write a logical control program for the reboiler.



The operating constraints are as follows:

- T1 must not exceed 350 ˚C
- T2 must be between 100 and 200 ˚C
- P1 cannot exceed 150 psi

The normal operating conditions for the controlled variables are as follows:

- F1 is 20 gal/min
- F2 is 10 gal/min
- F3 is closed

Solution

The first step is understanding the system to be controlled. Looking at the diagram, one can see that F1 controls the flow rate of the service water, F2 controls the flow rate of the process stream, and F3 controls the vent to the atmosphere. T1 is the temperature of the process stream entering the reboiler and T2 is the temperature of the process stream exiting the reboiler. P1 is the pressure inside the reboiler.

Below is a possible control program to ensure the operating constraints are met; there may be other solutions to achieve the same objective:

- IF T2 > 200 ˚C, THEN F1 = 30 gal/min
- IF T2< 100 ˚C, THEN F1 = 10 gal/min
- IF P1 > 150 psi, THEN open F3
- IF T1 > 350 ˚C, THEN F2 = 2 gal/min

In order to control the exit temperature of the process stream, one can increase or decrease the service water from the normal flow rate. Opening the vent to the atmosphere reduces the pressure if it reaches an unsafe value. If the entering temperature is too high for the process stream, reducing the entering flow rate will ensure the exit temperature is not too high for the remainder of the process. Note: reducing the process stream may negatively impact the rest of the plant.

Example 5.2.2: Thermostat

Chemical engineers are increasingly involved in biological applications. In many biological processes, temperature control can be very critical to the process. Take, for example, a simulated cell on a computer chip. Let's say that it is imperative for the chip to remain at 97°F ± 1°F, very similar to the temperature of the human body. Some of the reactions and processes on the chip will not function at temperatures outside this range and can become irreversibly damaged unless the power supply to the chip is turned off. The power is turned on and off by a switch S1. There are flows of cool water with an automatic valve V1 and warm water with an automatic valve V2 for cooling and heating purposes. The chip is attached to a thermostat to sense and control the temperature, T. Write out a controlled logic scheme to maintain the chip temperature and to prevent damage.

Hint: Heat or cool the chip before reaching the limits (about half a degree)

Solution

To control the temperature, the chip must be heated or cooled depending on the environment

IF (T<96.5) THEN V2 is open

ELSE V2 is closed

IF (T>97.5) THEN V1 is open

ELSE V1 is closed

The control is executed before the temperature limit is reached to allow for lag time in the heating or cooling water flow.

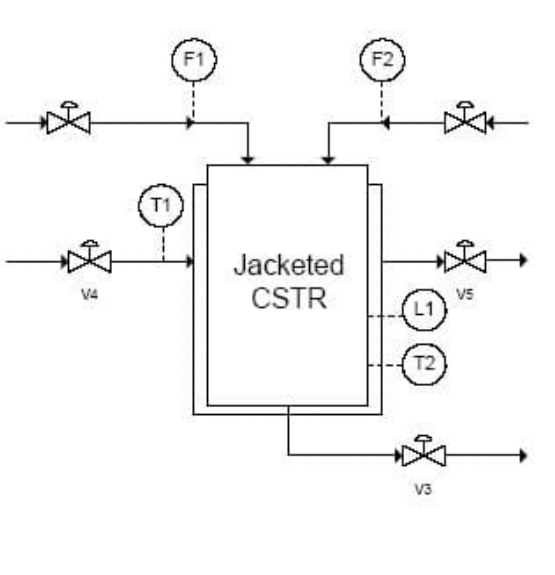
Also, to ensure that the chip is not damaged, the power needs to be shut off if the temperature goes above 98°F or below 96°F

WHILE (96

ELSE S1 is off

Example 5.2.3: Chemical Reactor

There is an exothermic chemical reaction occurring in a CSTR (Continuous Stirred Tank Reactor) that involves two reactants being fed at a 1:1 ratio. All valves are set to be 50% open normally. Write a control program that keeps the level in the CSTR vessel less than 8 meters (the tank is 10 meters tall) and the reactor temperature below 450 degrees Celsius.



Solution

While L1 > 8 set V3 to 100% open and close V1 and V2

Else set V1 and V2 and V3 to 50% open

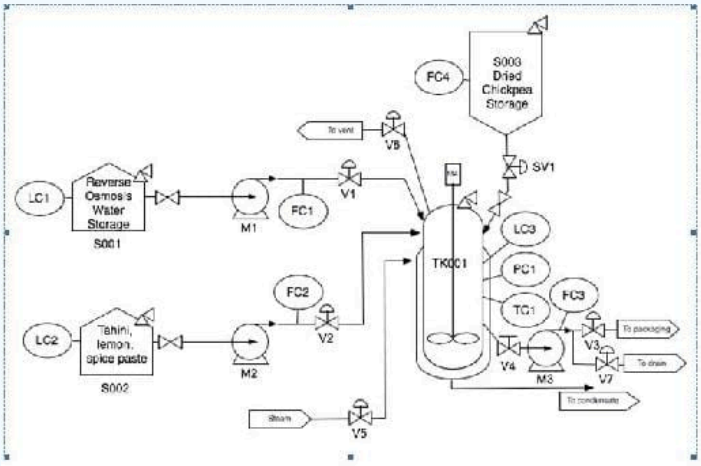
If T2 > 450 THEN set V5 and V4 to 100% open

Else set V5 and V4 to 50% open

This solution gives an example of using AND statements to control multiple valves with just one condition.

Example 5.2.4: Programming and Alarms

There is a process that is run by the P&ID shown below.



Based on this process and the steps listed below, write out a detailed control program for the process. Use comments (denoted with #) if necessary to explain the logic behind the steps.

1. Measure Qw units of water into a tank
2. Add Qc units of dried chickpeas
3. Let dried chickpeas soak for 20 hours without mixing
4. Drain off soaking water to waste (assume the filter in the tank will not allow whole chickpeas through the pump) 5) Add Qw units of fresh water to the tank.
5. Heat the tank to Tcook and maintain the pressure at 4 atm. Note that your tank is rated to withstand pressures between 0.5 and 6 atm, while outside of that range the tank may implode or explode.
6. Cook chickpeas for 20 minutes.
7. After cooking, turn off heat and allow the system to return to ambient temperature (Tamb) and ambient pressure. Beware of a strong vacuum forming in the tank as the water vapor condenses!
8. Drain cooking water to drain.
9. Pump in Qs units of the tahini spice mix
10. Blend the mixture for 10 minutes to produce a smooth hummus paste.
11. Pump out product to packaging.
12. Fill tank with clean water and agitate to clean the reactor.
13. Pump wash water to drain.

Solution

FUNCTION INITIALIZE

Turn off M1, M2, M3, M4
Close V1, V2, V3, V5, V6, V7, SV1
Set all timers to zero
Set all totalizers to zero

FUNCTION FAILSAFE

Turn off M1, M2, M3, M4
Close V1, V2, V3, V5, V7, SV1
Open V6

FUNCTION PROGRAM

#Step 1 – Measure Qw unites of water into a tank

Turn on M1
Open V1
WHILE FC1tot < Qw:

Adjust V1 to FC1set

IF LC1 < LC1min:

ALARM

Close V1

Turn off M1

#FC1tot is the total amount of fluid that has gone through the flow meter

#FC1set is the set point (amount the valve is open) for V1 that FC1 has already programmed into it

#LC1min is the minimum acceptable level of fluid in S001

#Step 2 – Add Qc units of dried chickpeas

Open SV1
WHILE FC4 < Qc:

Adjust SV1 to FC4set

IF LC4 < LC4min:

ALARM

Close SV1

#FC4set is the set point for SV1 that FC4 has already programmed into it

#LC4min is the minimum acceptable level of fluid in S003. LC4 is not on the P&ID, however, it makes sense to have one on it so the level on the tank can be properly monitored

#Step 3 – Let dried chickpeas soak for 20 hours without mixing

WAIT 20 hours

#Step 4 – Drain off soaking water to waste

Open V7
Turn on M3
WHILE FC3tot < Qw:

Adjust V7 to FC3set2

Turn off M3

Close V7

#FC3tot is the total amount of fluid that has gone through the flow meter

#FC3set2 is the set point for V7 that FC3 has already programmed into it

#Step 5 – Add Qw units of fresh water to the tank

Clear FC1tot
Turn on M1
Open V1
WHILE FC1tot < Qw:

Adjust V1 to FC1set

IF LC1 < LC1min:

ALARM

Close V1

Turn off M1

#Step 6 – Heat the tank to Tcook and maintain the pressure at 4 atm.

WHILE TC1 < Tcook:

Adjust v5 to Tcook

IF OR (PC1 < 0.5, PC1 > 6):

GO TO FAILSAFE

IF PC1 < PC1set:

Close V6

ELSE:

Adjust V6 to PC1set

IF LC3 > LC3max:

GO TO FAILSAFE

#PC1set is the setting that V6 must be set to for the tank to have 4 atm of pressure in it

#LC3max is the maximum level that the contents of the tank are allowed to get to. Anything higher indicates a problem with one of the flow meters.

#Step 7 – Cook chickpeas for 20 minutes#

WAIT 20 minutes

#Step 8 –After cooking, turn off heat and allow the system to return to ambient temperature and pressure

IF PC1 > PC1amb:

Adjust V6 to PC1set2

IF OR (PC1 < 0.5, PC1 > 6):

GO TO FAILSAFE

WHILE TC1 > Tamb:

Close V5

#PC1amb is the ambient pressure of 1 atm that the system needs to get to

#PC1set2 is the second setting on PC1 which effects how open V6 is

#Step 9 – Drain cooking water to drain

Clear FC3tot
Open V7
Turn on M3
WHILE FC3tot < Qw:

Adjust V7 to FC3set

Turn off M3

Close V7

#Step 10 – Pump in Qs units of the Tahini spice mix

```
Clear FC2tot
Turn on M2
Open V2
WHILE FC2tot < Qs:

Adjust V2 to FC2set

IF LC2 < LC2min:

ALARM

Close V2

Turn off M2

#FC2tot is the total amount of fluid that has gone through the flow meter

#FC2set is the set point for V2 that FC2 has already programmed into it

#Step 11 – Blend the mixture for 10 minutes to produce a smooth hummus paste

Turn on M4
WAIT 10 minutes
Turn off M4

#Step 12 – Pump out product to packaging

Open V3
Turn on M3
WHILE LC3 > 0:

Adjust V3 to FC3set

Turn off M3

Close V3

#FC3set is the set point for V3 that FC3 has already programmed into it

#Step 13 – Fill tank with clean water and agitate to clean the reactor

Turn on M1
Open V1
WHILE LC3 < LC3max:

Adjust V1 to FC1set

IF LC1 < LC1min:

ALARM

Close V1
Turn off M3
Turn on M4
WAIT 10 minutes

Turn off M4

#Step 14 – Pump was water to drain
```

Open V7
Turn on M3
WHILE LC3 > 0:

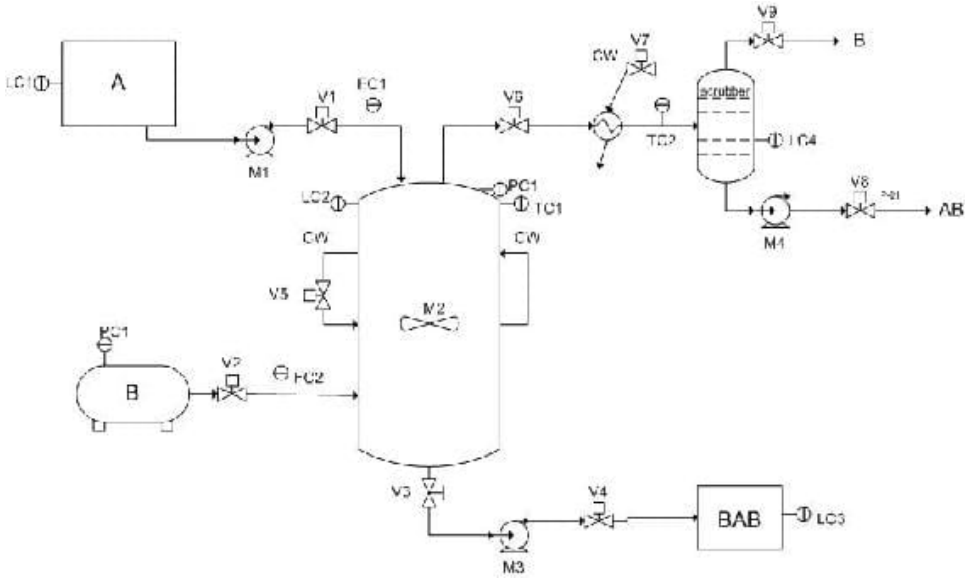
Adjust V7 to FC3set2

Turn off M3

Close V7

Worked out Example 4: Programming and Alarms

Worked Out Example 5: Another Chemical Reactor (taken from Prof. Barkel's 9/29/09 lecture)



WRITING THE CONTROL PROGRAM

Nomenclature:
x_A = input amount of A
y_B = input amount of B

Readings:
L_L = low level
L_H = high level
Lmax = maximum high level
Lmax = maximum low level

Things to note:
Be sure to use If, Then statements.
Lines in *italics* are comments used to organize the program.

CONTROL PROGRAM

Equal means within +/- 0.05 % of the value being held equal.

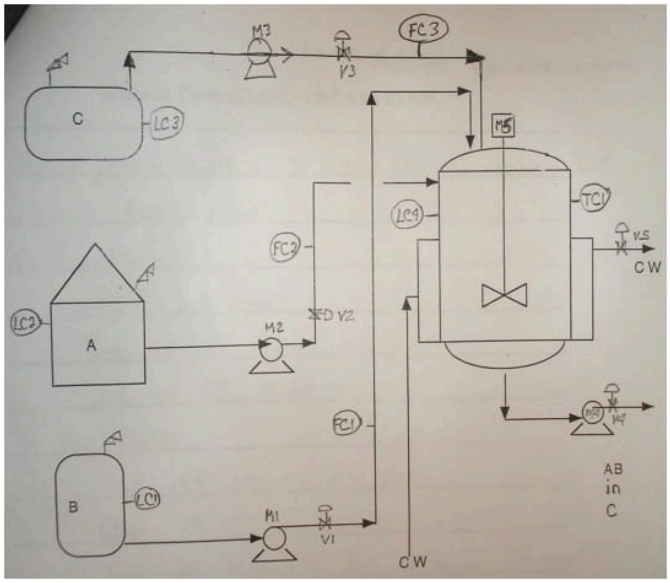
Initialize *This is used to set all control devices to the position you want them in.*
1. Turn off all valves
2. Shut off all motors

Fail Safe *This is used to shut down the process in case something goes wrong and must be terminated immediately.*
1. Turn off all valves except V5, V7, V9, V6
2. Shut off all motors except M2

Universal Statements
If T1>Tmax, then go Fail Safe.

1. Initialize
2. Open V6, V9, V1. *[so that systems doesn't build up pressure]*
3. If L1>L_L, then turn on M1. *[fill with A]*
4. If L2>=*Lag*, turn on M2. *[start stirrer]*
5. If F1>=x_A, turn off V1, shut off M1. *[stop flow of A]*
6. Open V5, V7. *[allow for cooling]*
7. If P1>=P_L, then open V2. *[fill with B]*
8. If L4>=L4min, open V8, turn on M4. *[release product AB]*
9. If L2<*Lag*, then shut off M2. *[so motor doesn't run dry]*
10. If L4<$L4_L$, close V8. *[stop outflow of product AB]*
11. If T1><T_H, then open V5. *[cool reactor temperature]*
12. If T2><T_H, then open V5. *[cool upper product stream]*
13. If T1<<T_L, then close V5. *[stop cooling of reactor]*
14. If F2>=<y_B, then turn off V2. *[stop flow of B]*
15. Close V8, M4, leave V6, V9 open. *[shut down process but allow it to vent]*
16. Pump out BAB.

Worked Out Example 6



$$\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{AB}$$

The reaction is exothermic
A, B, AB are all liquids
It goes to 100% completion

X_A = total amount of A used
Y_B = total amount of B used

B is added slowly into a full charge of A
The temperature is maintained at T_R
Upon completion of the reaction, the AB is cooled to T_P
Z_C, an amount of solvent C is added to the AB to make the final product

Using only the equipment and instrumentation shown, write the control logic for this BATCH reaction. Use **If**, **Then** logic.

Use of subscripts:
Ag for agitator
L for low-bottom of control range
H for high-top of control range
Min for minimum - lowest level allowed
Max for maximum - highest level allowed
tot for totalizer

Process Control Logic

- Equal means within +/- 0.05% of the value being held equal.

1) Initialize

Shut all valves
Turn off all motors
Set all FC_{tot} = 0

2) Fail safe

Shut off V1, V2, V3, V4
Open V5
Turn off M1, M2, M3
Turn off M4, Open V4
Turn on M5 if LC4>Lag, else turn off M5

3) If T1>T_{max}, then go Fail Safe
4) If L2>$L2_{min}$, then turn on M2, open V2
5) If L2<$L2_{min}$, then turn off M2, shut off V2
6) If $FC2_{tot}$ = X_A, then turn off M2, shut off V2
7) If L4>Lag, then turn on M5
8) If L1>$L1_{max}$, then go Fail Safe
9) If L1>$L1_{min}$, then turn on M1, open V1
10) If L1<$L1_{min}$,then turn off M1, shut off V1
11) If $FC1_{tot}$ = Y_B, then turn off M1, shut off V1
12) If T1>$T1_H$, then open V5
13) If T1<$T1_L$, then close V5
14) If L4>$L4_H$, then close V1, turn off M1
15) If T1>T_P, then open V5
16) If T1=T_P, then close V5
17) If L3>$L3_{min}$, then turn on M3, open V3
18) If L3<$L3_{min}$, then turn off M3, shut off V3
19) If $FC3_{tot}$ = Z_C, then turn off M3, shut off V3
20) If LC4 21) Open V4
22) Turn on M4
23) If L4<$L4_{min}$, then turn off M4, shut off V4

References

- Savitch, Walter. "Problem Solving with C++ The Object of Programing". Boston: Pearson Education, Inc. 2005.
- Stanford Encyclopedia of Philosophy
- Woolf, Peter. "Possible Useful Notes", Notes for Boolean Logic, September 22, 2005.
- Woolf, Peter. "Project 1 Sample Solution", October 6, 2005.
- Woolf, Peter. "Useful Definitions", Terms Review, October 11, 2005.

Contributors and Attributions

Authors: Stephanie Fraley, Michael Oom, Benjamin Terrien, John Zalewski
Stewards: Ross Bredeweg, Jessica Morga, Ryan Sekol, Ryan Wong

This page titled 5.2: Logical Control Programs - IF... THEN... WHILE... is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

6: MODELING CASE STUDIES

6.1: Surge Tank Model

6.2: Heated Surge Tank

6.3: Bacterial Chemostat

6.4: ODE and Excel CSTR model with heat exchange

6.5: ODE and Excel model of a Simple Distillation Column

6.6: ODE and Excel model of a Heat Exchanger

6.7: ODE and Excel Model of an Adiabatic PFR

6.8: Modeling and PID Controller Example - Cruise Control for an Electric Vehicle

6.9: Blood Glucose Control in Diabetic Patients

This page titled 6: Modeling Case Studies is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

6.1: Surge Tank Model

Introduction

Used to regulate fluid levels in systems, surge tanks act as standpipe or storage reservoirs that store and supply excess fluid. In a system that has experienced a surge of fluid, surge tanks can modify fluctuations in flow rate, composition, temperature, or pressure. Typically, these tanks (or “surge drums”) are located downstream from closed aqueducts or feeders for water wheels. Depending upon its placement, a surge tank can reduce the pressure and volume of liquid, thereby reducing velocity. Therefore, a surge tank acts as a level and pressure control within the entire system.

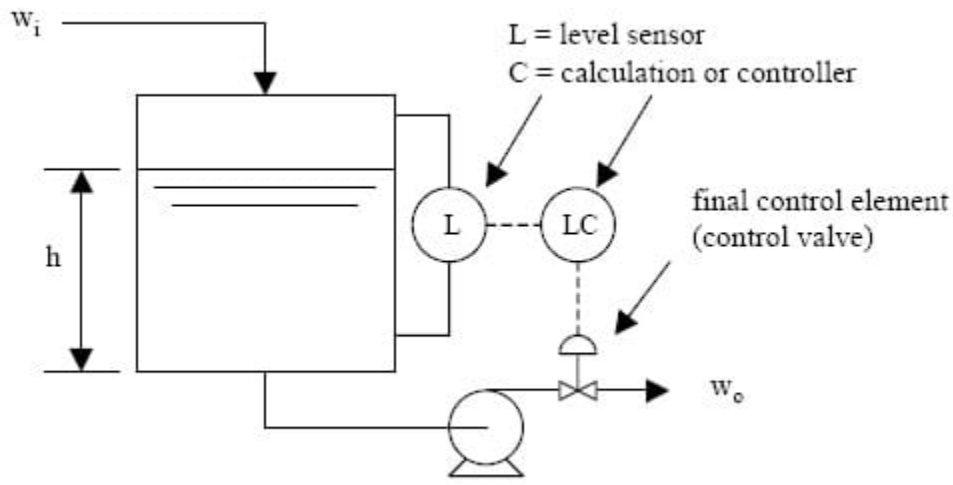
Since the flow to the surge tank is unregulated and the fluid that is output from the surge tank is pumped out, the system can be labeled as unsteady-state [MIT], but the approach to an approximate solution (below) utilizes techniques commonly adhered to when solving similar steady-state problems.

The technology behind surge tanks has been used for decades, but researchers have had difficulty fully finding a solution due to the non-linear nature of the governing equations. Early approximations involving surge tanks used graphical and arithmetical means to propose a solution, but with the evolution of computerized solving techniques, complete solutions can be obtained. [Wiley InterScience].

Derivation of Ordinary Differential Equation

To accurately model a surge tank, mass and energy balances need to be considered across the tank. From these balances, we will be able to develop relationships for various characteristics of the surge tank

Diagram of Surge Tank System



A surge tank relies on the level sensor to determine whether or not fluid stored in the tank should be removed. This regulated outflow is pumped out in response to calculations made by a controller that ultimately opens and closes the control valve that releases the fluid from the tank.

Governing Equations of Surge Tank Model

A surge tank's components must be divided up and evaluated individually at first, then as a whole. First, the expression for the inlet stream must be obtained. The simplistic sine function will be used as the basis for the expression of a stream because it

typically describes the tidal surge pattern of a low-viscous fluid, like water. The flowrate, w, will be given in units of kg h^{-1}.

$$w_i = a + b\sin(c\pi t)$$

where *a* and *b* are constants determined by the specific problem circumstance.

A mass balance must now be performed on the tank as a system. Using the concept that *mass in = mass out*, and assuming that the tank is a perfect cylinder devoid of diversions, an expression can be derived:

$$\text{rate mass in} - \text{rate mass out} = \text{accumulated fluid in tank}$$

Rewriting,

$$w_i - w_o = \rho\frac{dV}{dt}$$

$$\frac{dh}{dt} = \frac{w_i - w_o}{\rho A}$$

$$\int_0^h dh = \int_0^t \frac{w_i - w_o}{\rho A}dt$$

$$h(t) - h_o = \frac{1}{\rho A}\int_0^t w_i - w_o dt$$

Where, at time $t = 0$, the amount of fluid in the surge tank is constant; thus $h(0) = h_0$.

Substituting the original equation for the inlet stream, w_i, into the expression for the height of the tank, *h(t)*, the governing equation for the height of the tank is obtained:

$$h(t) = h_o + \frac{1}{\rho A}\int_0^t \left[(a + b\sin(c\pi t) - w_0\right]dt$$

Integrating by parts,

$$h(t) = h_o + \frac{x}{\rho A}(1 - \cos(c\pi t))$$

where *x* is formed from the constants *a* and *b* during integration.

Secondary uses of Surge Tanks

Surge tanks are most commonly used to protect systems from changes in fluid levels; they act as a reservoir that stores and supply excess fluid. In addition, the tanks shield the systems from dramatic changes in pressure, temperature, and concentration. They can also allow one unit to be shut down for maintenance without shutting down the entire plant.

Pressure

There can be moments of high pressure, called hammer shock, in a system when the liquid (incompressible) flow is stopped and started. The energy that liquids possess while traveling through pipes can be categorized as potential or kinetic. If the liquid is stopped quickly, the momentum that the liquid carries must be redirected elsewhere. As a result, the pipes vibrate from the reactive force and weight of the shock waves within the liquid. In extreme cases, pipes can burst, joints can develop leaks, and valves and meters can be damaged.

The extreme amounts of pressure are dampened when the fluid enters the surge tanks. The surge tank acts as a buffer to the system, dispersing the pressure across a greater area. These tanks make it possible for a system to more safely execute their tasks.

Temperature

The temperature of a fluid can either be controlled or changed through the use of a surge tank. The surge tank allows for a rapid change in fluid temperature. This is exemplified by the process of pasteurization; the milk needs to be at a high temperature for just a short period of time, so it is exposed to the high temperature and then moved to the surge tank where it can be stored and cooled (see heated surge tank).

A substance can enter the surge tank at room temperature, and it will instantaneously mix with the rest of the tank. Substances entering the tank will also subsequently rise to meet the high temperature and then exit the surge tank quickly thereafter.

Concentration

Concentration inside the surge tank is kept relatively constant, thus the fluid exiting the surge tank is the same as the fluid in the tank. This is favorable when there is a concentration gradient in the incoming fluid to the surge tank. The tank homogenizes the entering fluid, keeping the concentration of the reactants the same throughout the system, therefore eliminating any concentration gradient.

Example 6.1.1: Modeling Surge Tank

Suppose we are to design a surge tank that deals with flow swings of +/- 40% over a 10 minute period modeled by the following equation:

$$w_1 = 500 + 200 \sin \frac{\pi t}{10} \tag{6.1.1}$$

where flow is m^3h^{-1} and time in hours. The outlet flow rate, w_o, is desired to be 500 m^3h^{-1}.

The surge tank is 30m tall and has a cross sectional area of $5m^2$. The initial height of the fluid ($\rho = kg / m^3$) is 10m.

a. Model this hypothetical example.
b. The surge tank has upper and lower limit controls that open or close the valve if the level gets too high or low. What is the highest maximum set point that can be used without the surge tank overflowing?

Solution

a.

$$\rho A \frac{dh}{dt} = w_i - w_o \tag{6.1.2}$$

with $h(0) = h_o$.

$$h(t) = h_o + \frac{1}{\rho A} \int_0^t (w_i(t) - w_o(t))\, dt$$

Substituting Equation 6.1.1 into 6.1.2, we get

$$h(t) = h_o + \frac{2000}{\rho \pi A} \left(1 - \cos \frac{2\pi t}{20}\right) \tag{6.1.3}$$

To complete the design, we must have the cross-sectional area A of the surge tank, this would be given.

If we were to apply the following condition,

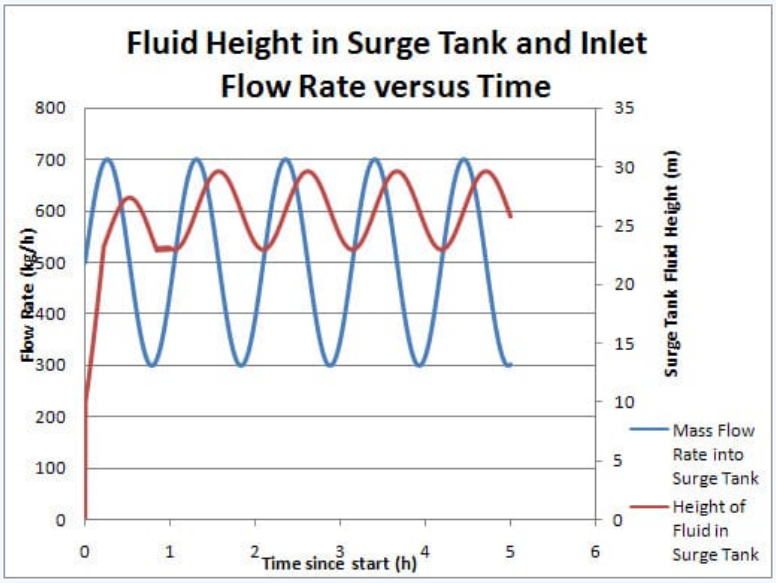
$$w_i - w_o = C \tag{6.1.4}$$

and were to substitute Equation 6.1.4 into Equation 6.1.2, we discover

$$h(t) = h_o + \frac{C}{\rho A}t \tag{6.1.5}$$

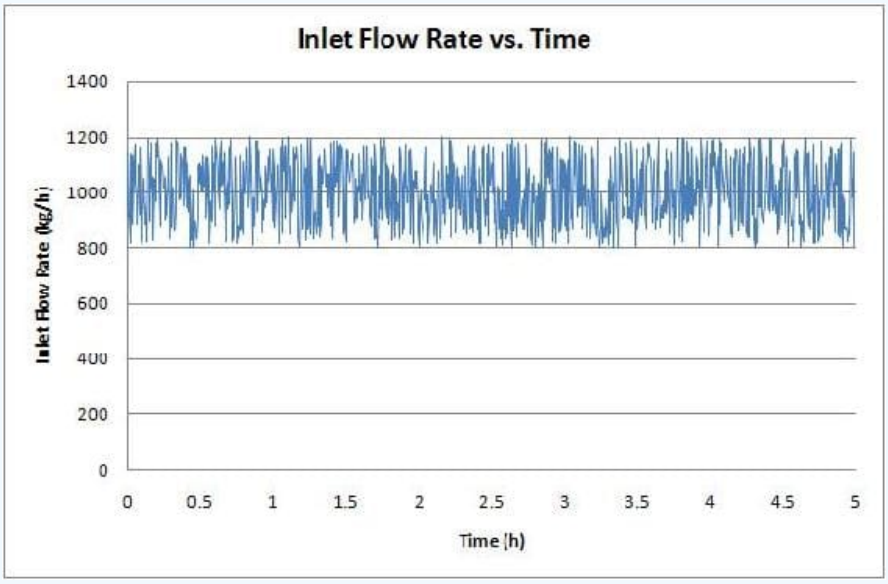
b.

Using the Microsoft Excel model provided, the highest possible maximum set point is 23m. Media:Model Surge Tank.xls



Example 6.1.2

In this example, we simulate random fluctuations that might be observed in a real world process stream. A random number generating function (in MS Excel) is used to generate pseudo-random fluctuations of $\pm$ 200 kg/h about an average value of 1000 kg/h for the inlet stream, $w_i(t)$ over a period of 5 hours. The result is as follows:



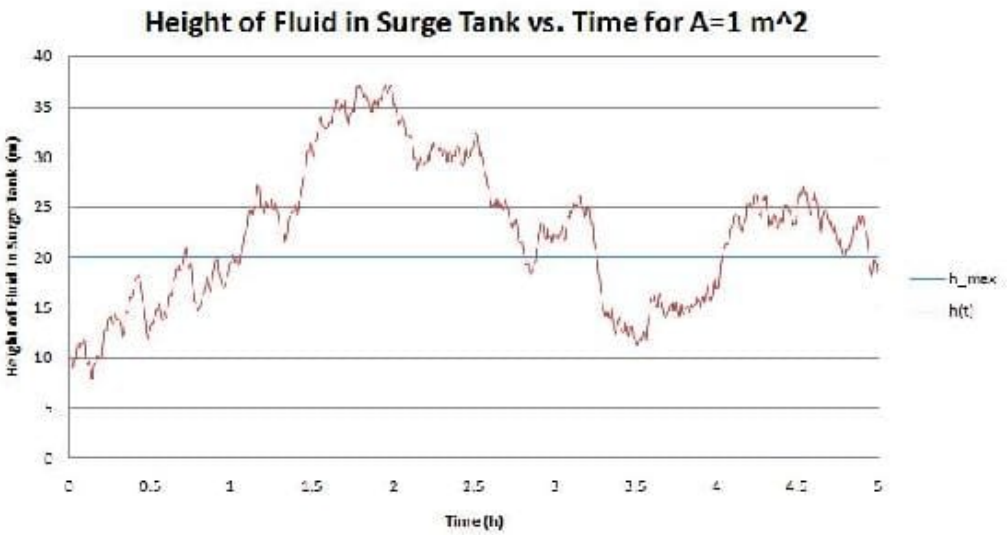
Examine the effect that these fluctuations have on the fluid level in surge tanks of various volumes. To vary the volume, assume tanks of constant height h_{max}=20m, and vary the cross sectional area from A=1m^2 to A=5m^2. Use the following parameters for the surge tank: Initial fluid level h_0=10m, ρ=1 kg / m^3, w_0=1000 kg / m^3, t_0=0 h and t_f=5h.

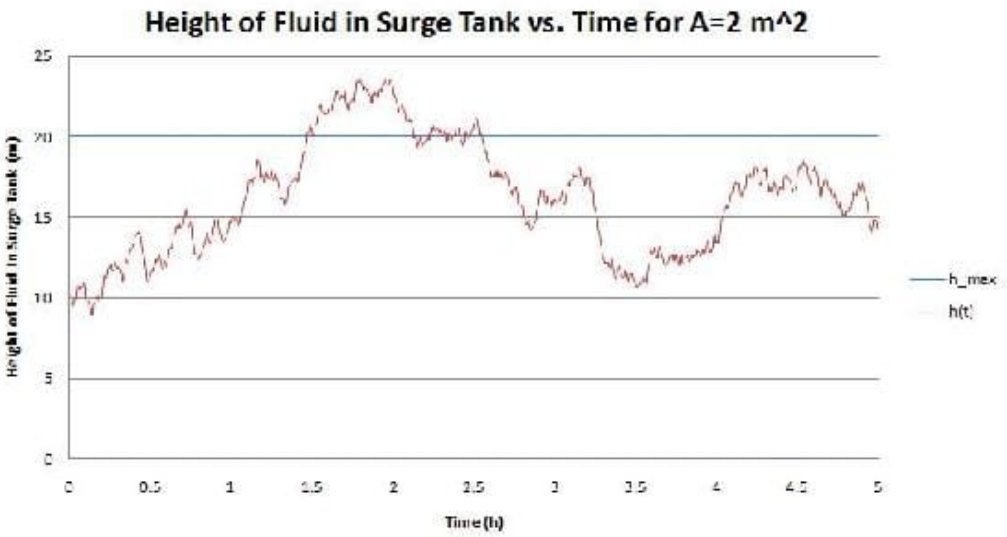
Solution

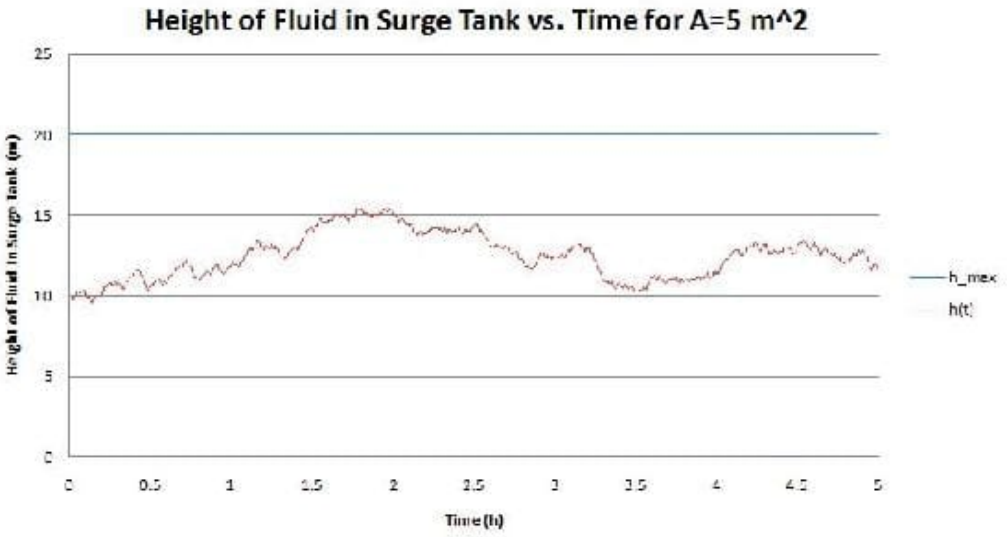
In the solution to the first example problem, we used a trigonometric function to simulate fluctuations in the inlet stream, and we obtained an analytical solution to the differential equation. Since we are using a pseudo-random fluctuating inlet stream in this example, we will solve this problem via numerical methods. Using the pseudo-random data for $w_i(t)$, we perform a numerical integration on the previously derived expression:

$$h(t) = h_o + \frac{1}{\rho A}\int_0^t (w_i(t) - w_o(t))\, dt \tag{6.1.6}$$

This integration was performed with the trapezoid rule (in MS Excel, using a slight modification of the posted Excel model), using the specified surge tank parameters, with A=1 m^2, 2 m^2, and 5 m^2 l. The following results were obtained:


Height of Fluid in Surge Tank vs. Time for A=1 m^2
Height of Fluid in Surge Tank (m)
Time (h)
h_max
h(t)


Height of Fluid in Surge Tank vs. Time for A=2 m^2
Height of Fluid in Surge Tank (m)
Time (h)
h_max
h(t)


Height of Fluid in Surge Tank vs. Time for A=5 m^2
Height of Fluid in Surge Tank (m)
Time (h)
h_max
h(t)

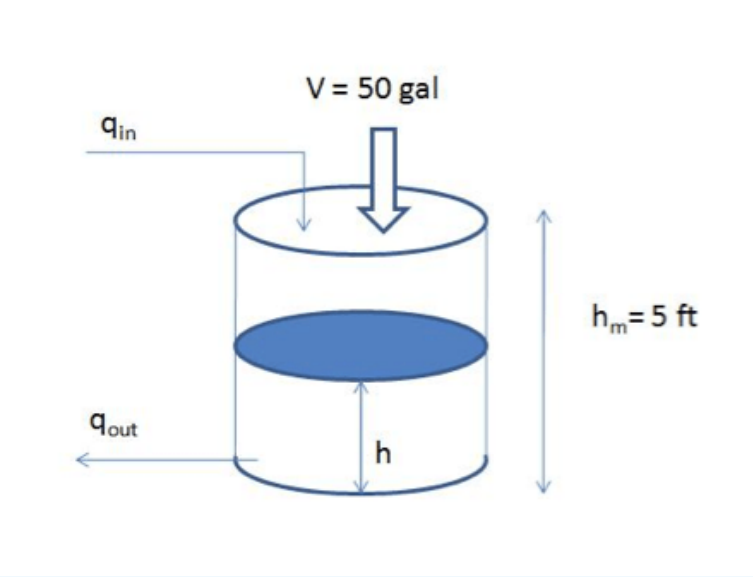
We see that increasing the volume of the surge tank by increasing the cross sectional area A reduces the magnitude of the fluctuations of the fluid level h in the surge tank. For A=1 m^2 and $2m^2$, the capacity of the surge tank is exceeded and it overflows.

Example 6.1.3

An operator quickly adds 50 gallons from a drum of water into a cylindrical surge tank with a diameter of 4 feet. The initial volume of liquid in the tank is 40 cubic feet and the total height of the tank is 5 feet. The inflow and exit flow rates are initally 6 cubic feet per minute. The resistance has a linear relationship with liquid height.

a. Derive a first order linear model for the tank height and determine the value of R for the outflow
b. Will the tank overflow when the drum is added?
c. Show the height h(t) after the drum of water is added; treat the change in h(t) as instantaneous
d. Does the tank height return to steady state? What is it?

Solution



(a)

Assume the operator adds the water at t = 0, so when t < 0, the tank is under steady state which can be described as:q_{IN} – q_{OUT} = 0 (1)

Assume: $q_{OUT} = \left(\frac{h}{R}\right)$ where R is the resistance and h is the height of the liquid. Initial volume of the tank is:

$$V_{SS} = \pi \left(\frac{d^2}{4}\right) h_{SS}$$

$$40 = \pi \left(\frac{4^2}{4}\right) h_{SS}$$

Thus: $h_{SS} = \left(\frac{10}{\pi}\right) = 3.183ft$

Substitute $q_{IN} = \left(\frac{6ft^3}{min}\right)$ and $q_{OUT} = \left(\frac{h_{SS}}{R}\right)$ into equation 1 above:

0 = 6 – h_{SS} / R

$$R = \left(\frac{h_{SS}}{6}\right) = \left(\frac{0.531min}{ft^2}\right)$$

At time t > 0, the 50 gallons (6.685 cubic feet) of water is added to the system which disrupts the steady state and the corresponding initial condition of the system is:

$V_0 = V_{SS} + V_{operator} = 40 + 6.685 = 46.685ft^3$

$$h_0 = \left(\frac{46.685}{4\pi}\right) = 3.7151ft$$

The corresponding dynamic linear model:

Label 'ex2:2' multiply defined

with initial condition: $h_0 = 3.7151\ \ ft$

(b)

After adding the drum, h_0 = 3.7151 which is lower than 5 feet. The system will not overflow at t_0. Because the liquid level is higher than the steady state level and $q_{out} = \left(\frac{h}{R}\right)$, the out flow rate will be larger than q_{IN} which pushes the liquid level drop. As a result, water will not over flow.

(c)

Define deviation variables. $h' = h - h_{SS}$; $q_{IN}' = q_{IN} - q_{INSS} = 0$ and substitute into equation (2)

Missing \left or extra \right

(3)

Solve the linear model above (Eq. 3): (Either by direct integration or laplace transform and inverse lapace transform)

$$h' = h'_0 exp(-\left(\frac{t}{AR}\right)) \quad (4)$$

Put $h' = h - h_{SS}$ into equation 4:

$$h - h_{SS} = (h_0 - h_{SS})exp(-\left(\frac{t}{AR}\right))$$

Then:

$$h = h_{SS} + (h_0 - h_{SS})exp(-\left(\frac{t}{AR}\right))$$

$$h = 3.183 + (0.5321)exp(-\left(\frac{t}{AR}\right))$$

(d)

Yes. Since

$$\lim_{t\to\infty} h = 3.183 + 0.5321\exp(-0.1499t) = 3.183\text{ft}$$

which was the original steady state

Sage's Corner

If you are a sage for this page, please link your narrated powerpoint presentation here.

References

- Kundur, Prabha. *Power System Stability and Control*, McGraw-Hill 1994.
- Cheng-Nan Lin and John S. Gladwell. *Non-dimensional Surge Tank Analysis,* State of Washington Water Research Center, Washington State University 1970.
- Slocum, Stephen Elmer. *Elements of Hydraulics*, McGraw-Hill 1915.
- MIT OpenCourseWare. www.ocw.cn/OcwWeb/Chemical-Engineering/10-450Process-Dynamics--Operations--and-ControlSpring2003/LectureNotes/index.htm. Design of a surge tank to smooth out fluctuations in flow. Definition of important process control terms
- Use the Harvard Referencing style for references in the document.
- For more information on when to reference see the following Wikipedia entry.

Contributors and Attributions

Authors: (September 6, 2007) Cara Canady, David Carpenter, Che Martinez, Jeremy Minty, Bradley Novak

This page titled 6.1: Surge Tank Model is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

6.2: Heated Surge Tank

Introduction

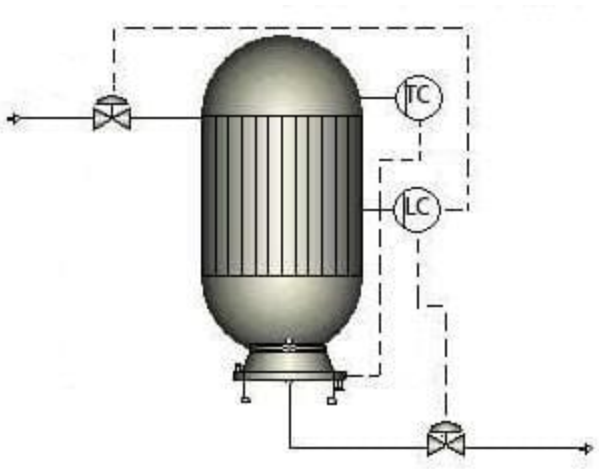
A surge tank is an additional safety or storage tank that provides additional product or material storage in case it becomes needed. Heat exchange can be added to surge tanks, which provides temperature control for the tank. Within a system these tanks can appear as distillation columns, reboilers, heated CSTR's, and heated storage. They can increase production rates by allowing a batch of product to finish reacting while the initial tank is reloaded, provide constant system parameters during start up and shut down, or create additional storage space for product overflow or backup material.

Uses for Heated Surge Tanks:

- Fuel surges caused by motion of a vehicle: If fuel cannot be drawn from the primary tank, the engine resorts to a surge tank. The heat maintains the fuel's temperature.
- Caramelization: During the formation of caramel, the mixture must be maintained at a specific temperature for a predetermined amount of time. Once the ingredients are thoroughly dissolved, the mixture is transferred to a heated surge tank and maintained until the caramel has thickened and is ready to be drawn out.
- Mixing of gases: Bulk gas lines can be connected to a heated surge tank with a pressure sensor. The pressure sensor would control the temperature. By heating the gas when it first enters the tank, there is no risk of explosion later due to expansion.
- Heated pools: Surge tanks are used to catch and store displaced water from a pool. If the pool is heated, a heated surge tank should be used to maintain the temperature of the water.
- De-aeration: Heated surge tanks are often used with de-aerators. They heat the component that will enter the de-aerator, because if the component is not preheated, the de-aerator must wait until the component reaches the correct temperature. This could waste a lot of time and energy.
- Chemical Baths: Often in industry, things need to be treated with a chemical bath. The chemicals usually need to be at a certain temperature so that it will adhere to the object. A heated surge tank is perfect for this application.
- Reboilers: Liquids coming off of a distillation column can be reheated to enter the column again at a higher temperature to drive the separation process. Many industries use this tool to obtain a more efficient separation and produce a higher net profit.
- Product or Material Backup: Heated surge tanks can also be used as simple storage in two ways. First, a surge tank can be used excess product not yet sold or otherwise moved out of the production system. Second, heated surge tanks can serve as backup for chemical or fuel supplies to a production plant, such as outdoor gasoline tanks for a backup generator in case of power failure.

Basic Design for Heated Surge Tanks

Above is a basic example of a heated surge tank. While surge tanks can have multiple inputs and outputs, for simplicity we have only included one of each here.



Connected to the tank is a temperature control, which controls the heater. Depending on the temperature of the fluid, this control will increase or decrease the heating to the tank. This will keep the fluid at the necessary temperature to meet the process requirements.

There is also a level control connected to the tank to indicate when the tank has neared maximum capacity. When this happens, the control will open the valve at the bottom of the tank, allowing the product to flow further down the process. The control can also slow or stop the flow coming through the input valve. These mechanisms will prevent the tank from overfilling. The position of the level control depends on the type of material in the process, the phase of the material, the type of level control, and the requirements of the system.

Useful Equations

The basic equations that govern heated surge tanks are shown below. First, a simple mass balance is done on the system. Second, the energy balance was simplified using the assumptions listed below. Most problems involving this type of tank can be described by these equations. Additional considerations may require additional variables and equations.

Assumptions:

1. The substance coming into the tank is uniform.
2. No reaction is taking place.
3. The tank is well mixed, which means the temperature profile is constant throughout the tank.

Mass Balance

Since there is no generation from reactions inside the heated surge tank, we obtain the rate of accumulation or level inside the tank by subtracting what is coming out from what is coming in.

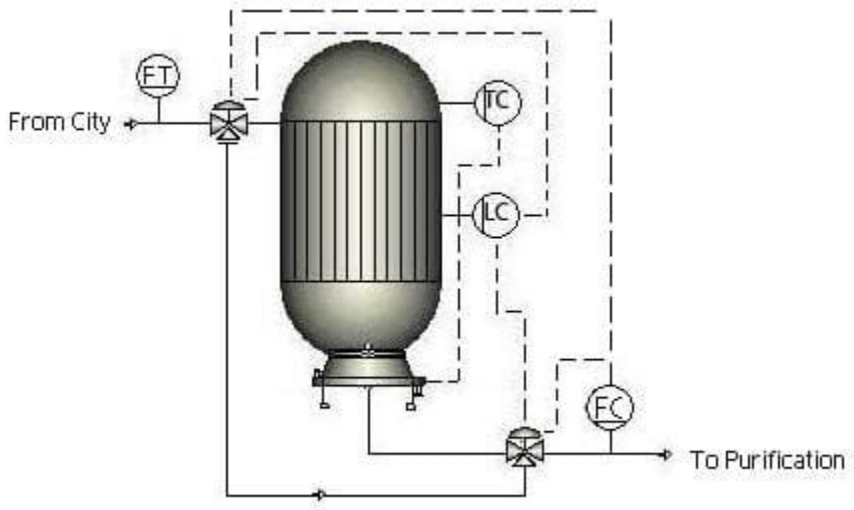
$$Rate of Accumulation = (\text{Flow In}) - (\text{Flow Out})$$

Energy Balance

The temperature change with respect to time is essential for the purpose of configuring a system to reach steady state. When turning a system on or off, there is a time period in which the system is in unsteady state. During this time, the system is difficult to model. In steady state, the system is easier to model because once steady state is reached the left hand term will become zero.

$$\frac{dT}{dt} = \frac{v\rho C_p\left(T_0 - T\right) + UA\left(T_C - T\right)}{V\rho C_p}$$

Case Study - Water Purification at IBM



At IBM's manufacturing facility outside Burlington, Vermont, a heated surge tank is used in the de-ionized water system. In order to wash semi-conductor wafers in manufacturing, the water has to be about 1,000,000 times cleaner than the incoming city water. All of this purification is done on site.

The water comes in from the municipal water source at a constant flow rate, but manufacturing demand is not constant. In order to compensate for this, when the demand in manufacturing is low, a surge tank is used to store extra water for high demand periods. Because the large tank is located outside and the winter in Vermont is very cold, the tank is heated to prevent the water inside from freezing.

During normal operation of the system, the surge tank is bypassed. When a flow controller downstream has low demand, the inlet valve opens, letting water into the surge tank. A level controller monitors the tank to make sure it doesn't overfill and can shut off the inlet valve and let water out. A temperature controller controls the heater jacket to maintain the water around 50°C. When the demand for water increases, the flow controller near the outlet can shut off the inlet valve to the tank, and/or further open the outlet valve to access the extra water supply in the tank.

Example 6.2.1: Determining Temperature of Heating Fluid for Tank

A heated surge tank is being designed to hold paraffin wax coming from a distillation column at an oil refinery. High pressure steam will be used as a heating fluid in the heat exchanger to heat up and maintain the paraffin at 51°C (to maintain high viscosity and prevent solidification). The physical parameters of the tank (volume of 5 m^3) and heat exchanger within it are given. The tank is originally filled with paraffin at room temperature. At what temperature must the high pressure steam be to sufficiently heat the paraffin; will a proposed temperature of 130°C be adequate?

The paraffin comes into the tank at 37°C at a volumetric flow rate of 0.0005 m^3/s. The heat exchange coefficient is equal to 50 W/m^2/K and the area of the heat exchanger is 2.0 ^{2}. The heat capacity of the paraffin is 2900 J/Kg/K and the density is 800 Kg/m3.

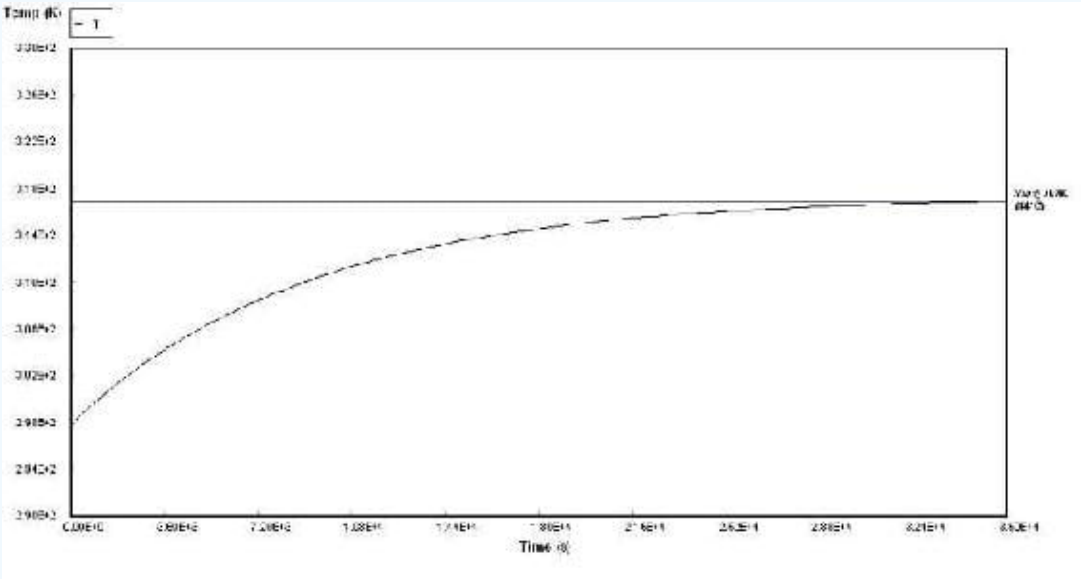
Solution

POLYMATH CODE:

$$\dfrac{d(T)}{d(t)} = \frac{-v_{in} * \rho * C_p*(T - T_{in} + U * A * (T_a - T)}{C_p * V * \rho}$$

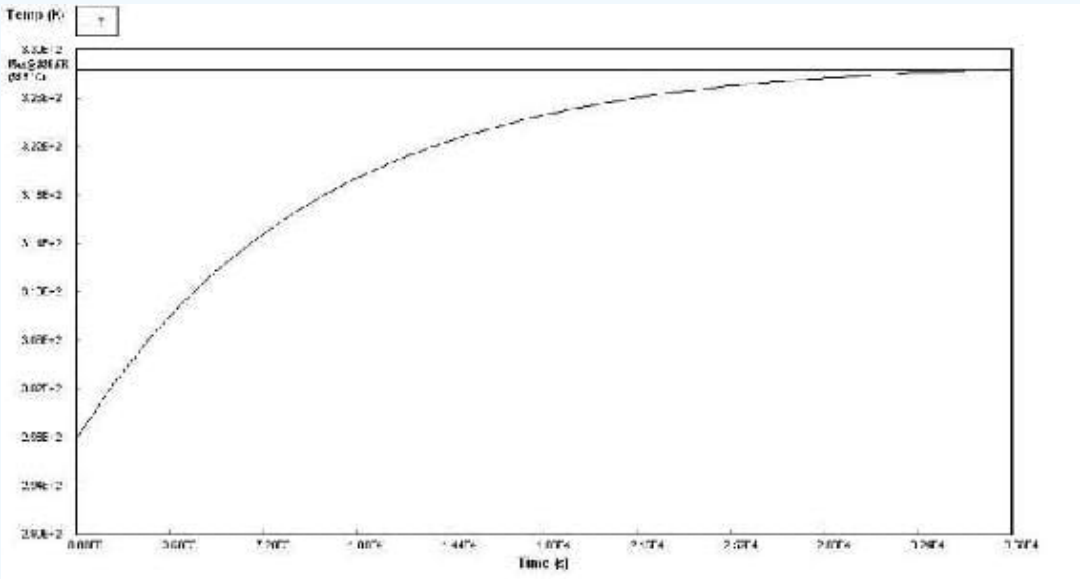
- vin= 0.0005 m^3/s

- U= 50 W/m^2/K
- A= 2 m^2
- Ta= 130°C= 403K
- Cp= 2900 J/Kg/K
- rho= 800 Kg/m3
- Tin= 37°C= 310K
- V= 5 m^3
- t(0)= 0 s
- T(0)= 298K
- t(f)= 36000 s



With a heating fluid at a temperature of 130°C, the fluid only reaches a temperature of 44°C (317K). A higher capacity heating fluid must be used.

Trying a heating fluid at 277°C, we generate the plot below.



The plot shows that the tank will reach a temperature of about 55.5°C (328.5K) with a heating fluid at 277°C. This will be sufficient to maintain the paraffin in the liquid phase.

Example 6.2.2: Time to Reach Steady State after Heating Failure

For the same surge tank from problem 1, if the heater fails for 2 hours after 10 hours of operation and is then restarted, how long will it take after it is restarted to reach steady state?

Solution

Approximately 20 hours

The first manipulation that must be done to your Polymath code is to create an "if-then-else" statement for the dT/dt line for the case before hour 10, the case between hour 10 and hour 12, and the case after hour 12. For the time ranging between 10 and 12 hours, because the heating element has failed, the differential equation over this period must reflect that. Thus, the section $(U * A * (T_a - T))$ from the differential equation is dropped and the equation appears as follows:

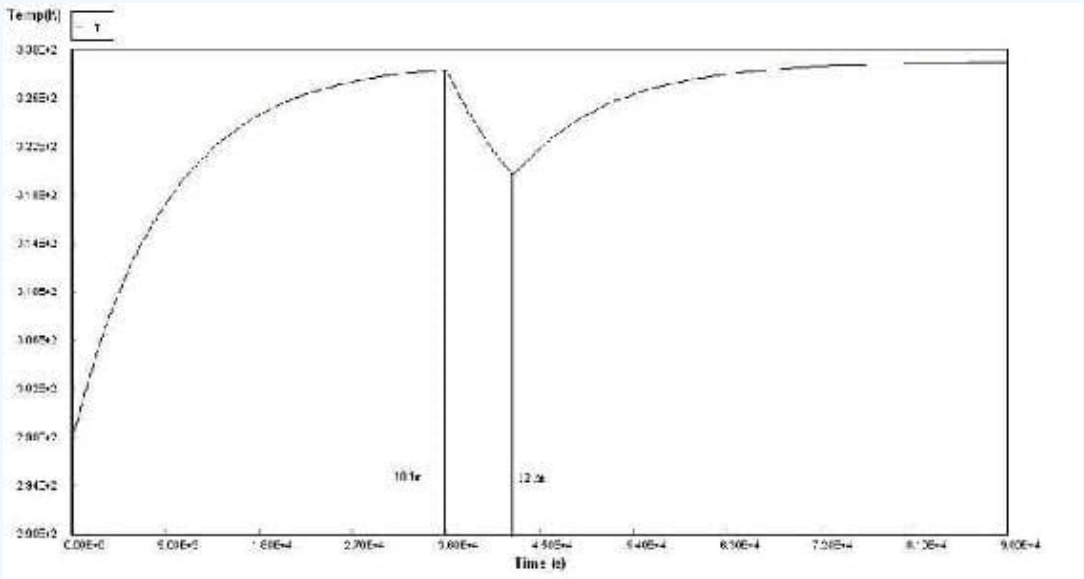
$$d(T)/d(t) = \frac{-v_{in} * \rho * C_p * (T - T_{in})}{C_p * V * \rho}.$$

For any other time in this simulation, the normal dT/dt equation is used. In order to determine where the surge tank reaches steady state again, the final time is increased until dT/dt approximately reaches zero. The graph generated in Polymath will look like the following, using the code below as a parameter:

\[\dfrac{d(T)}{d(t)= if(36000

then

$$(-v_{in} * \rho * Cp * (T - T_{in})/(C_p * V * \rho))else((-vin * rho * Cp * (T - T_{in}) + (U * A * (T_a - T)))/(Cp * V * rho))$$



Contributors and Attributions

- **Authors:** *Angela Antosiewicz, Christopher Kline, Peter Heisler, Paul Niezguski*
- **Stewards:** *Karen Staubach, Soo Kim, Kerry Braxton-Andrew, Joshua Katzenstein*

This page titled 6.2: Heated Surge Tank is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

6.3: Bacterial Chemostat

Introduction

Bioreactors are used to grow, harvest, and maintain desired cells in a controlled manner. These cells grow and replicate in the presence of a suitable environment with media supplying the essential nutrients for growth. Cells grown in these bioreactors are collected in order to enzymatically catalyze the synthesis of valuable products or alter the existing structure of a substrate rendering it useful. Other bioreactors are used to grow and maintain various types of tissue cultures. Process control systems must be used to optimize the product output while sustaining the delicate conditions required for life. These include, but are not limited to, temperature, oxygen levels (for aerobic processes), pH, substrate flowrate, and pressure. A bacterial chemostat is a specific type of bioreactor. One of the main benefits of a chemostat is that it is a continuous process (a CSTR), therefore the rate of bacterial growth can be maintained at steady state by controlling the volumetric feed rate. Bacterial chemostats have many applications, a few of which are listed below.

- Pharmaceuticals: Used to study a number of different bacteria, a specific example being analyzing how bacteria respond to different antibiotics. Bacteria are also used in the production of therapeutic proteins such as insulin for diabetics.
- Manufacturing: Used to produce ethanol, the fermentation of sugar by bacteria takes place in a series of chemostats. Also, many different antibiotics are produced in chemostats.
- Food Industry: Used in the production of fermented foods such as cheese.
- Research: Used to collect data to be used in the creation of a mathematical model of growth for specific cells or organisms.

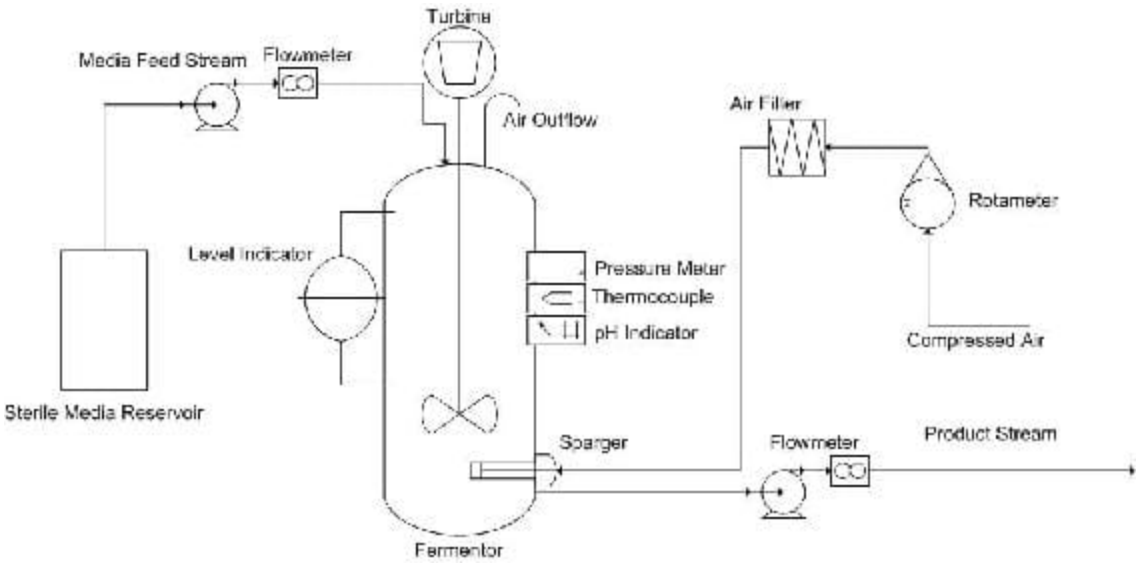
The following sections cover the information that is needed to evaluate bacterial chemostats.

Bacterial Chemostat Design

The bacterial chemostat is a continuous stirred-tank reactor (CSTR) used for the continuous production of microbial biomass.

Chemostat Setup

The chemostat setup consists of a sterile fresh nutrient reservoir connected to a growth chamber or reactor. Fresh medium containing nutrients essential for cell growth is pumped continuously to the chamber from the medium reservoir. The medium contains a specific concentration of growth-limiting nutrient (C_S), which allows for a maximum concentration of cells within the growth chamber. Varying the concentration of this growth-limiting nutrient will, in turn, change the steady state concentration of cells (C_C). Another means of controlling the steady state cell concentration is manipulating the rate at which the medium flows into the growth chamber. The medium drips into culture through the air break to prevent bacteria from traveling upstream and contaminating the sterile medium reservoir.



The well-mixed contents of the vessel, consisting of unused nutrients, metabolic wastes, and bacteria, are removed from the vessel and monitored by a level indicator, in order to maintain a constant volume of fluid in the chemostat. This effluent flow can be controlled by either a pump or a port in the side of the reactor that allows for removal of the excess reaction liquid. In either case, the effluent stream needs to be capable of removing excess liquid faster than the feed stream can supply new medium in order to prevent the reactor from overflowing.

Temperature and pressure must also be controlled within the chemostat in order to maintain optimum conditions for cell growth. Using a jacketed CSTR for the growth chamber allows for easy temperature control. Some processes such as biological fermentation are quite exothermic, so cooling water is used to keep the temperature at its optimum level. As for the reactor pressure, it is controlled by an exit air stream that allows for the removal of excess gas.

For aerobic cultures, purified air is bubbled throughout the vessel's contents by a sparger. This ensures enough oxygen can dissolve into the reaction medium. For anaerobic processes, there generally is not a need for an air inlet, but there must be a gas outlet in order to prevent a build up in pressure within the reactor.

In order to prevent the reaction mixture from becoming too acidic (cell respiration causes the medium to become acidic) or too basic, which could hinder cell growth, a pH controller is needed in order to bring pH balance to the system.

The stirrer ensures that the contents of the vessel are well mixed. If the stirring speed is too high, it could damage the cells in culture, but if it is too low, gradients could build up in the system. Significant gradients of any kind (temperature, pH, concentration, etc.) can be a detriment to cell production, and can prevent the reactor from reaching steady state operation.

Another concern in reactor design is fouling. Fouling is generally defined as the deposition and accumulation of unwanted materials on the submerged surfaces or surfaces in contact with fluid flow. When the deposited material is biological in nature, it is called biofouling. The fouling or biofouling in a system like this can cause a decrease in the efficiency of heat exchangers or decreased cross-sectional area in pipes. Fouling on heat exchanger surfaces leads to the system not performing optimally, being outside the target range of temperature, or spending excess energy to maintain optimum temperature. Fouling in pipes leads to an increase in pressure drop, which can cause complications down the line. To minimize these effects, industrial chemostat reactors are commonly cylindrical, containing volumes of up to 1300 cubic meters, and are often constructed from stainless steel. The cylindrical shape and smooth stainless steel surface allow for easy cleaning.

Design Equations

The design equations for contiuous stirred-tank reactors (CSTRs) are applicable to chemostats. Balances have to be made on both the cells in culture and the medium (substrate).

Mass Balance

The mass balance on the microorganisms in a CSTR of constant volume is:

[Rate of accumulation of cells, g/s] = [Rate of cells entering, g/s] – [Rate of cells leaving, g/s] + [Net rate of generation of live cells, g/s]

The mass balance on the substrate in a CSTR of constant volume is:

[Rate of accumulation of substrate, g/s] = [Rate of substrate entering, g/s] – [Rate of substrate leaving, g/s] + [Net rate of consumption of substrate, g/s]

Assuming no cells are entering the reactor from the feed stream, the cell mass balance can be reworked in the following manner:

$$(RateAccumulationCells) = V\frac{dC_C}{dt} \tag{6.3.1}$$

$$(\text{Flow Entering}) \ -(\text{Flow Leaving}) = 0 - \nu_0 C_C \tag{6.3.2}$$

$$\text{Rate Cell Generation} = V\left(r_g - r_d\right) \tag{6.3.3}$$

Similarly, the substrate mass balance may be reworked in the following manner:

$$\text{Rate Accumulation Substrate} = V\frac{dC_S}{dt} \tag{6.3.4}$$

$$(\text{Flow Entering}) \ -(\text{Flow Leaving}) = \nu_0 C_{S0} - \nu_0 C_S \tag{6.3.5}$$

$$\text{Rate Substrate Consumption} = Vr_S \tag{6.3.6}$$

Putting Equations 6.3.1-6.3.3 together gives the design equation for cells in a chemostat:

$$V\frac{dC_C}{dt} = 0 - \nu_0 C_C + V\left(r_g - r_d\right) \tag{6.3.7}$$

Similarly, Equations 6.3.4-6.3.6 together gives the design equation for substrate in a chemostat:

$$V\frac{dC_S}{dt} = \nu_0 C_{S0} - \nu_0 C_S + V\left(r_g - r_d\right) \tag{6.3.8}$$

Assumptions made about the CSTR include perfect mixing, constant density of the contents of the reactor, isothermal conditions, and a single, irreversible reaction.

Rate Laws

Many laws exist for the rate of new cell growth.

Monod Equation

The Monod equation is the most commonly used model for the growth rate response curve of bacteria.

$$r_g = \mu C_c \tag{6.3.9}$$

where

- r_g = cell growth rate
- C_c = cell cencentration
- μ = specific growth rate

The specific cell growth rate, μ, can be expressed as

$$\mu = \mu_{\max}\frac{C_s}{K_s + C_s} \tag{6.3.10}$$

where

- μ_{max} = a maximum specific growth reaction rate

- K_S = the Monod constant
- C_S = substrate concentration

Tessier Equation and Moser Equation

Two additional equations are commonly used to describe cell growth rate. They are the **Tessier and Moser Equations**. These growth laws would be used when they are found to better fit experimental data, specifically at the beginning or end of fermentation.

Tessier Equation:

$$r_g = \mu_{\max}\left[1 - \exp\left(-\frac{C_s}{k}\right)\right] C_c \tag{6.3.11}$$

Moser Equation:

$$r_g = \frac{\mu_{\max} C_s}{1 + kC_s^{-\lambda}} \tag{6.3.12}$$

where λ and k are empirical constants determined by measured data.

Death Rate

The death rate of cells, r_d, takes into account natural death, k_d, and death from toxic by-product, k_t, where C_t is the concentration of toxic by-product.

$$r_d = (k_d + k_t C_t)\, C_c \tag{6.3.13}$$

Death Phase The death phase of bacteria cell growth is where a decrease in live cell concentration occurs. This decline could be a result of a toxic by-product, harsh environments, or depletion of nutrients.

Stoichiometry

In order to model the amount of substrate and product being consumed/produced in following equations, yield coefficients are utilized. Y_{sc} and Y_{pc} are the yield coefficients for substrate-to-cells and product-to-cells, respectively. Yield coefficients have the units of g variable/g cells. Equation 6.3.14 represents the depletion rate of substrate:

$$-r_s = Y_{sc} r_g + mC_c \tag{6.3.14}$$

Equation 6.3.15 represents the rate of product formation:

$$r_p = Y_{pc} r_g \tag{6.3.15}$$

Control Factors

The growth and survival of bacteria depend on the close monitoring and control of many conditions within the chemostat such as the pH level, temperature, dissolved oxygen level, dilution rate, and agitation speed. As expected with CSTRs, the pumps delivering the fresh medium and removing the effluent are controlled such that the fluid volume in the vessel remains constant.

pH level

Different cells favor different pH environments. The operators need to determine an optimal pH and maintain the CSTR at it for efficient operation. Controlling the pH at a desired value during the process is extremely important because there is a tendency towards a lower pH associated with cell growth due to cell respiration (carbon dioxide is produced when cells respire and it forms carbonic acid which in turn causes a lower pH). Under extreme pH conditions, cells cannot grow properly, therefore appropriate action needs to be taken to restore the original pH (i.e. adding acid or base).

Temperature

Controlling the temperature is also crucial because cell growth can be significantly affected by environmental conditions.

Choosing the appropriate temperature can maximize the cell growth rate as many of the enzymatic activates function the best at its optimal temperature due to the protein nature of enzymes.

Dilution rate

One of the important features of the chemostat is that it allows the operator to control the cell growth rate. The most common way is controlling the dilution rate, although other methods such as controlling temperature, pH or oxygen transfer rate can be used. Dilution rate is simply defined as the volumetric flow rate of nutrient supplied to the reactor divided by the volume of the culture (unit: time-1). While using a chemostat, it is useful to keep in mind that the specific growth rate of bacteria equals the dilution rate at steady state. At this steady state, the temperature, pH, flow rate, and feed substrate concentration will all remain stable. Similarly, the number of cells in the reactor, as well as the concentration of reactant and product in the effluent stream will remain constant.

Negative consequences can occur if the dilution rate exceeds the specific growth rate. As can be seen in Equation 6.3.16 below, when the dilution rate is greater than the specific growth rate (D > μ), the dC_C/dt term becomes negative.

$$\frac{dC_C}{dt} = (\mu - D)C_C \tag{6.3.16}$$

This shows that the concentration of cells in the reactor will decrease and eventually become zero. This is called wash-out, where cells can no longer maintain themselves in the reactor. Equation 6.3.17 represents the dilution rate at which wash-out will occur.

$$D_{max} = \frac{\mu_{max} C_{s0}}{K_s + C_{s0}} \tag{6.3.17}$$

In general, increasing the dilution rate will increase the growth of cells. However, the dilution rate still needs to be controlled relative to the specific growth rate to prevent wash-out. The dilution rate should be regulated so as to maximize the cell production rate. Figure 1 below shows how the dilution rate affects cell production rate(DC_C), cell concentration (C_C), and substrate concentration (C_S).

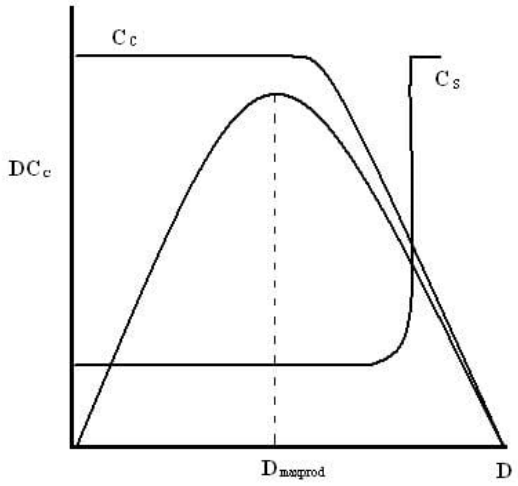


Figure 1: Cell concentration, cell production, and substrate concentration as a function of dilution rate

Initially, the rate of cell production increases as dilution rate increases. When $D_{maxprod}$ is reached, the rate of cell production is at a maximum. This is the point where cells will not grow any faster. D = μ (dilution rate = specific growth rate) is also established at this point, where the steady-state equilibrium is reached. The concentration of cells (C_C) starts to decrease once the dilution rate exceeds the $D_{maxprod}$. The cell concentration will continue to decrease until it reaches a point where all cells are washed out. At this stage, there will be a steep increase in substrate concentration because fewer and fewer cells are present to consume the substrate.

Oxygen transfer rate

Since oxygen is an essential nutrient for all aerobic growth, maintaining an adequate supply of oxygen during aerobic processes is crucial. Therefore, in order to maximize the cell growth, optimization of oxygen transfer between the air bubbles and the cells becomes extremely important. The oxygen transfer rate (OTR) tells us how much oxygen is consumed per unit time when given concentrations of cells are cultured in the bioreactor. This relationship is expressed in Equation 6.3.18 below.

$$\text{Oxygen Transfer Rate (OTR)} = Q_{O_2} C_C \tag{6.3.18}$$

Where C_C is simply the concentration of cell in the reactor and Q_{O2} is the microbial respiration rate or specific oxygen uptake rate. The chemostat is a very convenient tool to study the growth of specific cells because it allows the operators to control the amount of oxygen supplied to the reactor. Therefore it is essential that the oxygen level be maintained at an appropriate level because the cell growth can be seriously limited if inadequate oxygen is supplied.

Agitation speed

A stirrer, usually automated and powered with a motor, mixes the contents of the chemostat to provide a homogeneous suspension. This enables individual cells in the culture to come into contact with the growth-limiting nutrient and to achieve optimal distribution of oxygen when aerobic cultures are present. Faster, more rigorous stirring expedites cell growth. Stirring may also be required to break agglutinations of bacterial cells that may form.

Q&A

Q1: Why is a chemostat called a chemostat?

A1: Because the chemical environment is static, or at steady state. The fluid volume, concentration of nutrients, pH, cell density, and other parameters all are assumed to remain constant throughout the operation of the vessel.

Q2: What are some concerns regarding chemostats?

A2: a) Foaming results in overflow so the liquid volume will not be constant. b) Changing pumping rate by turning the pump on/off over short time periods may not work. Cells respond to these changes by altering rates. A very short interval is needed for it to respond correctly. c) Fragile and vulnerable cells can be damaged/ruptured when they are caught between magnetic stirring bar and vessel glass. d)Bacteria contamination occurs because bacteria travel upstream easily and contaminate the sterile medium. This can be solved by interrupting the liquid path with an air break.

Q3: The Monod equation uses a Michaelis-Menten relationship which is based on a quasi-state assumption. (T/F)

A3: T

Q4: An important feature of chemostat is the dilution rate. Define dilution rate.

A4: Dilution Rate = volume of nutrient medium supplied per hour divided by the volume of the culture.

Q5: What are the advantages/disadvantages over choosing a chemostat instead of a batch reactor for bioreactions?

A5: Advantages: 1. A chemostat has better productivity than a batch reactor. There is a higher rate of product per time per volume. A batch process wastes time. 2. A chemostat is operated at steady state, therefore it has better control maintaining the same conditions for all product produced.

Disadvantages: 1. A chemostat is less flexible than a batch reactor. A batch reactor can be used to make more than one product. 2. It is harder to maintain a sterile system in a chemostat. A batch reactor is easier to clean.

Q6: What is the physical meaning of the Monod constant?

A6: The Monod constant is a substrate concentration at which the growth rate of the biomass of microbial cells participating in the reaction is half the maximum growth rate.

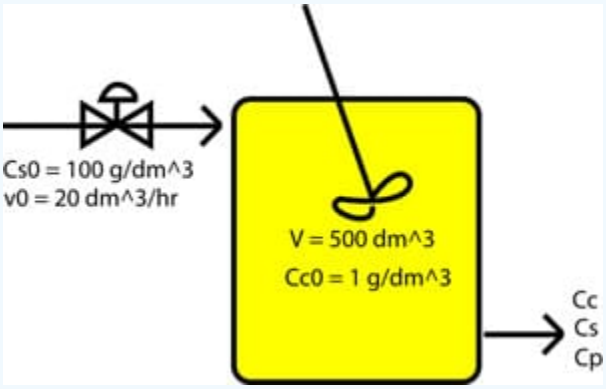
Example 6.3.1

Researchers at the University of Michigan are using a bacterial chemostat to model the intestinal tract of a pig in order to study the metabolism of E. Coli bacteria in that particular environment. The growth chamber of the chemostat has a volume of 500 dm^3.

The initial concentration of E. Coli bacteria inoculated in the chemostat growth chamber is 1 g/dm^3. A 100g/dm^3 substrate feed is fed to the chemostat at a volumetric flow rate of 20 dm^3/hr. How much time is required for this biochemical process to reach steady rate from the point of startup? Assume the growth rate is the Monod equation for bacteria bacterial cell growth, shown above.

Additional data pertaining to the problem is given: μ_{max} = 0.8; K_s = 1.7 g/dm^3; $Y_{s/c}$ = 8; $Y_{p/c}$ = 5; m = 0; r_d = 0;

Schematic Drawing



Answer = 3.7 hours

Solution

The Chemostat was modeled in Excel using the design equations above and Euler's Method. A graph of Cell Concentration (g/dm^3) vs Time(hr) was then plotted. When the Cell Concentration become stable, steady state has been reached and the time can be read off the graph. Below is a screen shot of the model and the graph created.

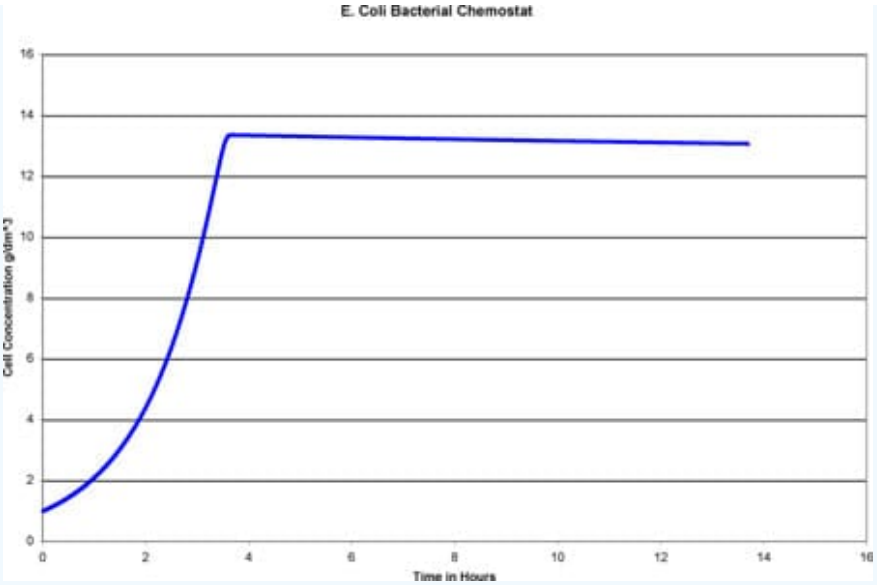
Excel Model Screen Shot

Given Information

Symbols	Values	Units	Description
Cs0	100	g/dm^3	Initial Feed Concentration of Substrate
m.max	0.8	h-1	Max specific growth rate
Ks	1.7	g/dm^3	Monod Constant
V	500	dm^3	Volume of Chemostat
v0	20	dm^3/hr	Initial Volumetric Flowrate
Ys\c	8	g Substrate/g Cells	Yield Coefficient Substrate/Cells
Yp\c	5	g Product/g Cells	Yield Coefficient Product/Cells
Cc0	1	g/dm^3	Initial Concentration of Cells
m	0	g Substrate/(g Cells*hr	Maintenance Constant for Cells
kvalve	0	dm^3/hr^2	Increase Fowrate Constant
deltat	0.001	hr	Step size

Calculation

t (s)	Cs	dCs/dt	Cc	dCc/dt	Cp	dCp/dt	v	dv/dt	rg
0	100		1		0		20	0	0.786627
0.001	100	-6.293018682	1	0.746621	0	3.933137	20	0	0.786627
0.002	99.9937	-6.293018682	1.000746627	0.746621	0.003933	3.933137	20	0	0.787214
0.003	99.9874	-6.297458876	1.001493255	0.747184	0.007866	3.936069	20	0	0.7878
0.004	99.9811	-6.301899059	1.002240439	0.747741	0.011802	3.939002	20	0	0.788387
0.005	99.9748	-6.306342552	1.002988179	0.748298	0.015741	3.941936	20	0	0.788975
0.006	99.9685	-6.310789355	1.003736477	0.748855	0.019683	3.944873	20	0	0.789562
0.007	99.9622	-6.31523947	1.004485332	0.749413	0.023628	3.947812	20	0	0.790151
0.008	99.9559	-6.319692829	1.005234745	0.749971	0.027576	3.950753	20	0	0.790739



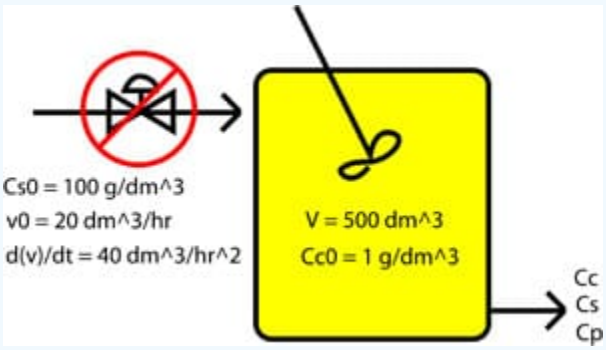
Excel Graph

This graph clearly shows that steady state is reached 3.7 hours after start up.

Example 6.3.2

After calculating the time required to reach steady state, the researchers decide to start up the chemostat. While do so, the control valve for the inlet substrate feed malfunctions. The flow rate of substrate into the chemostat is accelerating at 40 dm^3/hr^2. Determine how long they have to correct the problem before wash-out occurs and all of the bacteria in the chemostat is lost.

Schematic Drawing



Modeling the Malfunction

$$fracdv_0dt = kvalve = 40$$

Answer = 20 hours

Solution

The Chemostat was modeled in Excel using the design equations above and Euler's Method. A graph of Cell Concentration (g/dm^3) vs Time(hr) was then plotted. When the Cell Concentration becomes zero wash-out of the bacteria took place. Below is a screen shot of the model and the graph created.

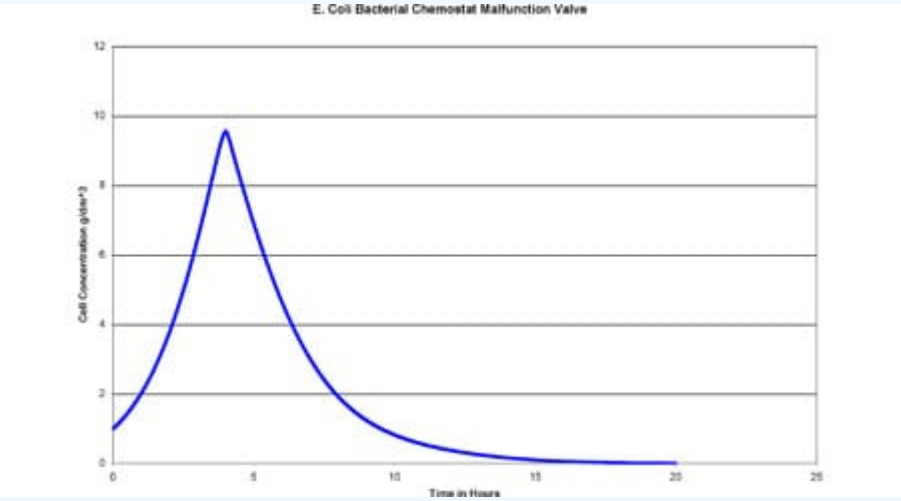
Excel Model Screen Shot

Given Information

Symbols	Values	Units	Description
Cs0	100	g/dm^3	Initial Feed Concentration of Substrate
mumax	0.8	h-1	Max specific growth rate
Ks	1.7	g/dm^3	Monod Constant
V	500	dm^3	Volume of Chemostat
vo	20	dm^3/hr	Initial Volumetric Flowrate
Ys/c	0	g Substrate/g Cells	Yield Coefficient Substrate/Cells
Yp/c	5	g Productg Cells	Yield Coefficient Product/Cells
Cc0	1	g/dm^3	Initial Concentration of Cells
m	0	g Substrate/(g Cells*hr	Maintenance Constant for Cells
kvalve	40	dm^3/hr^2	Increase Flowrate Constant
deltat	0.001	hr	Step size

Calculation

t (s)	Cs	dCs/dt	Cc	cCc/dt	Cp	dCp/dt	v	dv/dt	rg
0	100		1		0		20	40	0.786627
0.001	100	-6.293018682	1	0.746627	0	3.933137	20.04	40	0.786627
0.002	99.9937	-6.293018682	1.000746127	0.746647	0.003933	3.933137	20.08	40	0.787214
0.003	99.9874	-6.297458876	1.001493175	0.747024	0.007866	3.936059	20.12	40	0.7878
0.004	99.9811	-6.301090556	1.002240199	0.7475	0.011802	3.939001	20.16	40	0.788387
0.005	99.9748	-6.306341041	1.002987199	0.747977	0.015741	3.941935	20.2	40	0.788974
0.006	99.9685	-6.310786331	1.003735175	0.748454	0.019683	3.944871	20.24	40	0.789562
0.007	99.9622	-6.315234427	1.004484129	0.748931	0.023628	3.947809	20.28	40	0.79015
0.008	99.9559	-6.31968533	1.005233059	0.749408	0.027576	3.950748	20.32	40	0.790738



Excel Graph

This graph clearly shows wash-out occurs 20 hours after start up. We can see in example that process controls are extremely important for Bacterial Chemostats.

The template model used for both Worked Out Example 1 and 2 can be downloaded here Media: Bacterial Chemostat Template.xls

References

- "Chemostat." *McGraw-Hill Dictionary of Scientific and Technical Terms*. McGraw-Hill Companies, Inc., 2003. Accessed September 16, 2007. Available http://www.answers.com/topic/chemostat
- Fogler, H. Scott (2006). *Elements of Chemical Reaction Engineering*. New Jersey: Prentice Hall PTR. ISBN 0-13-047394-4
- Kelly, C. "Ch 9 - Reactors" 17 February, 2004. Retrieved 2007-09-24. Available www.ecs.syr.edu/faculty/kelly/handouts/biochemical%20engineering/chapter%209%20reactors.ppt
- Smith, H.L. (2006). "Bacterial Growth". Retrieved on 2007-09-15.
- Strandberg, Per Erik (2004). "Mathematical models of bacteria population growth in bioreactors: formulation, phase space pictures, optimisation and control". Retrieved on 2007-09-16.
- Strandberg, Per Erik (2003). "The Chemostat". Retrieved on 2007-09-15.
- "What is a Bioreactor?" Transgalactic Ltd., 25 May, 2005. Retrieved 2007-09-24. Available http://www.bionewsonline.com/o/what_is_bioreactor.htm

Contributors and Attributions

Authors: Shoko Asei, Brian Byers, Alexander Eng, Nicholas James, Jeffrey Leto
Stewards: Jeffrey Falta, Taylor Lebeis, Shawn Mayfield, Marc Stewart, Thomas Welch
Stewards: Sarah Hebert, Valerie Lee, Matthew Morabito, Jamie Polan

This page titled 6.3: Bacterial Chemostat is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

6.4: ODE and Excel CSTR model with heat exchange

Introduction

A **Continuous Stirred-Tank Reactor (CSTR)** is a chemical reaction vessel in which an impeller continuously stirs the contents ensuring proper mixing of the reagents to achieve a specific output (Figure 6.4.1). Useful in most all chemical processes, it is a cornerstone to the Chemical Engineering toolkit. Proper knowledge of how to manipulate the equations for control of the CSTR are tantamount to the successful operation and production of desired products. The purpose of this section is to model *dynamic conditions* within a CSTR for different process conditions. Simplicity within the model is used as the focus is to understand the dynamic control process.

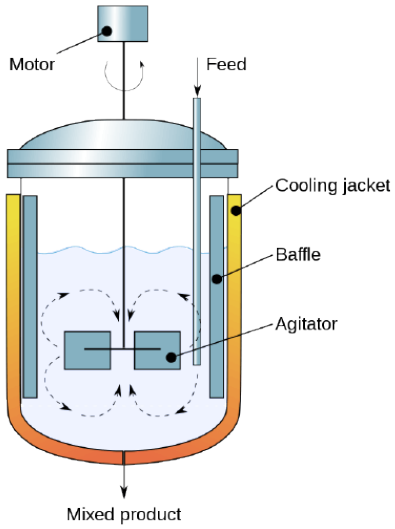


Figure 6.4.1: Cross-sectional diagram of a CSTR. (CC BY-SA 3.0; Daniele Pugliesi via Wikipedia)

Assumptions

For the purposes of this section, we have made the following assumptions to explain CSTR with heat exchange modeling.

- **Perfect mixing:** The agitator within the CSTR will create an environment of perfect mixing within the vessel. Exit stream will have same concentration and temperature as reactor fluid.
- **Single, 1st order reaction:** To avoid confusion, complex kinetics are not considered in the following modeling. \[\-r_{a}=k C_{A}]
- **Parameters specified:** We assume that the necessary parameters to solve the problem have been specified.
- **Volume specified:** In a control environment, the size of the vessel is usually already specified.
- **Constant Properties:** For this model, we have made the assumption that the properties of the species we are trying to model will remain constant over the temperature range at which we will be looking at. It is important to make this assumption, as otherwise we will be dealing with a much more complex problem that has a varying heat capacity, heat of reaction, etc.

ODE Modeling in Excel

To setup the model, the mass and energy balances need to be considered across the reactor. From these energy balances, we will be able to develop relationships for the temperature of the reactor and the concentration of the limiting reactant inside of it.

Variable Definitions

The following table gives a summary of all of the variables that we used in our mathematical formulas.

Symbol Explanations

Symbol	Meaning	Symbol	Meaning
V	Volume of Reactor	ρ	Density of Stream
m	Mass Flow Rate	C_{A0}	Original Concentration
C_A	Current Concentration	T_0	Original Temperature
T	Current Temperature	T_C	Coolant Temperature
ΔH_{rxn}	Heat of Reaction	ΔC_p	Overall change in Heat Capacity
k_0	Rate Law Constant	E	Activation Energy
R	Ideal Gas Constant	UA	Overall Heat Transfer Coefficient

Mass Balance

From our energy and material balances coursework, we know that the general equation for a mass balance in any system is as follows:

\mathrm is only supported in math mode

$$\text{Rate Accumulation} = V\frac{dC_A}{dt}$$

$$\text{Flow In} - \text{Flow Out} = \frac{m}{\rho}\left(C_{A0} - C_A\right)$$

$$\text{Rate Generation} = V\left(-r_a\right) = VC_Ak_0e^{-E/RT}$$

In the case of a CSTR, we know that the rate of accumulation will be equal to $V\frac{dC_A}{dt}$. This comes from the fact that the overall number of moles in the CSTR is VC_A, so the accumulation of moles will just be the differential of this. Since V is a constant, this can be pulled out of the differential, and we are left with our earlier result. We also know that the flow of moles in versus the flow of moles out is equal to $\frac{m}{\rho}\left(C_{A0} - C_A\right)$, which is the mass flow rate, divided by the density of the flow, and then multiplied by the difference in the concentration of moles in the feed stream and the product stream. Finally, we can determine what the rate of generation of moles in the system is by using the Arrhenius Equation. This will give us the rate of generation equal to

.

Combining all of these equations and then solving for $\frac{dC_A}{dt}$, we get that:

$$\frac{dC_A}{dt} = \frac{m}{\rho V}\left(C_{A0} - C_A\right) - k_0C_Ae^{-E/RT}$$

Energy Balance

From our thermodynamics coursework, we know that the general equation for an energy balance in any system is as follows:

$$\{Rate\ Energy Accumulation\} - (Heat\ In) \quad \{Heat\ Out\} \mid \{Rate\ Heat\ Generation\} \mid (Heat\ Transfer)$$

$$(Rate\ Energy Accumulation) = V\rho C_p\frac{dT}{dt}$$

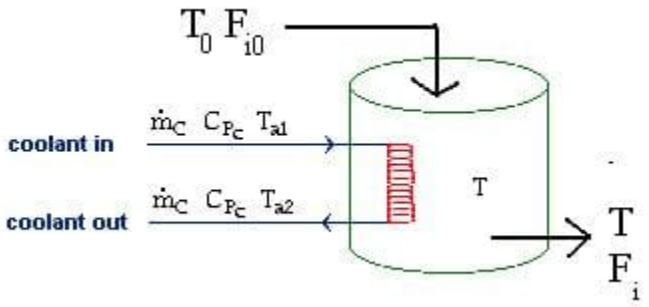
$$(Heat\ In) - (Heat\ Out) = m\Delta C_p(T_0 - T)$$

$$(Rate\ Heat\ Generation) = -V\Delta H_{rxn}(-r_a) = -V\Delta H_{rxn}k_0C_Ae^{-E/RT}$$

$$(Heat\ Transfer) = UA(T_C - T)$$

In the case of a CSTR, we know that the rate of energy accumulation within the reactor will be equal to $V\rho C_p\frac{dT}{dt}$. This equation is basically the total number of moles (mass actually) in the reactor multiplied by the heat capacity and the change in temperature. We also know that the heat generated by this reaction is , which is the rate of mass generation (– Vr_a) times the specific heat of reaction (ΔH_{rxn}). The overall rate of heat transfer into and out of the system is given by $m\Delta C_p(T_0 - T)$. This equation is the flow rate multiplied by the heat capacity and the temperature difference, which gives us the total amount of heat flow for the system. Finally, the amount of heat transferred into the system is given by $UA(T_C - T)$. Combining all of these equations and solving the energy balance for $\frac{dT}{dt}$, we get that:

In a realistic situation in which many chemical processed deal with multiple reactions and heat effects slight changes to the modeled equation must be done. The diagram below evaluates the heat exchanger under heat effects in which there is an inlet and outlet temperature that is accounted for in the enthalpy term in the newly modeled equation.



To model a heat exchanger that accounts for multiple reactions simply take the δH_{rxn} and δC_p term and add the Greek letter sigma for summation in front of the terms. When considering a case with multiple reactions and heat effects, the enthalpy and heat capacity of each reaction must be implemented in the energy balance, hence i and j represents the individual reaction species. The equation now looks something like this:

Euler's Method

In order to model an ODE in Excel, we will be using Euler's Method. The reason why we chose Euler's Method over other methods (such as Runge-Kutta) was due to the simplicity associated with Euler's. Trying to model with methods such as Runge-Kutta are much more difficult in Excel, while Euler's method is quite simple. We also checked the accuracy of our results using Euler's Method by comparing our answers with that of Polymath, a highly accurate ODE modeling program. We found that our answers were the same as that of Polymath, so we are quite comfortable using Euler's Method for our model. Here are the full details on Euler's Method. For this model, we will primarily be interested in the change in concentration of a reactant and the temperature of the reactor. These are the two differential equations that we were able to obtain from the mass and energy balances in the previous section. The following is an application of Euler's method for the CSTR on the change in concentration:

$$C_{Ai} = C_{A(i-1)} + \frac{dC_A}{dt}\Delta t$$

This same application can be made to the change in temperature with respect to time:

$$T_i = T_{i-1} + \frac{dT}{dt}\Delta t$$

Assuming that all values in the ODE's remain constant except for T and C_A , the new value is then found by taking the pervious value and adding the differential change multiplied by the time step.

List of Equations

The following are a summary list of all of the equations to be used when modeling CSTR with a heat exchange.

$$\frac{dC_A}{dt} = \frac{m}{\rho V}(C_{A0} - C_A) - k_0 C_A e^{-E/RT}$$

$$C_{Ai} = C_{A(i-1)} + \frac{dC_A}{dt}\Delta t$$

$$T_i = T_{i-1} + \frac{dT}{dt}\Delta t$$

Combining Using Excel

To model the CSTR using Excel you use Euler's Method, Energy Balance and Mass Balance together to solve for the concentration at time *t*.

	A	B	C	D	E	F	G	H
10	m	4	kg/s	Mass feec rate and product rate				
11	k_0	1.97E+20	s^{-1}	Rate constant				
12	UA	7000	J/s	Heat Transfer Coefficiert and Area				
13	T_0	400	K	Feed temperature				
14	$T_{0,2}$	400	K	Change Initial Feed Temperature (leave the same if this				
15	t for $T_{c,2}$	0	s	Time when Initial Feed Temperature Changes				
16	T_c	350	K	Coolant temperature				
17	$T_{c,2}$	350	K	Change Initial Coolant Concentration (leave the same if				
18	t for $T_{c,2}$	0	s	Time when Initial Coolant Temperature Changes				
19	V	3000	L	Reactor volume				
20	ρ	1	kg/L	Density of reactor feed and product (assumed constant)				
21	ΔH_{Rx}	-200	J/mol	Heat of reaction				
22								
23								
24	t (s)	T_0	T (K)	T_c	C_{A0}	C_A	dC_A/dt	dT/dt
25	0	400	400	350	9	9	0.00E+00	0
26	1	400	400	350	9	9	-3.72E-01	-0.04227
27	2	400	399.9577	350	9	8.627994	-3.54E-01	-0.04556
28	3	400	399.9122	350	9	8.273736	-3.37E-01	-0.04869
29	4	400	399.8635	350	9	7.936446	-3.21E-01	-0.05167
30	5	400	399.8118	350	9	7.615362	-3.06E-01	-0.05448

Screenshot of Excel Model for CSTR with Heat Exchange

How to Use Our Model

In order to help facilitate understanding of this process, we have developed an Excel spreadsheet specifically for looking at the changes in concentration and temperature given some change in the input to the CSTR system. An example of a change to the system could be that the temperature of the feed stream has dropped by a given number of degrees, or that the rate at which the feed stream is being delivered has changed by some amount. By using our spreadsheet, you will be able to easily plug in your given parameters, and look at the trend of the concentration and temperature over a wide time interval.

The way in which this spreadsheet works is quite simple. Boxes are provided for you to input all of the given information for your CSTR problem. Various columns containing values for the temperature, concentration, etc., with respect to time have also been provided. There are then more columns that contain the values for the various differential equations from above. With the time derivative in hand, we are then able to predict the value of the temperature or concentration at the next given time interval.

Our easy-to-use Excel model is given here: CSTR Modeling Template. In our model, you will find a column of unknowns that must be specified in order to solve for the optimal conversion temperature and optimal concentration of A. There are then two cells that will display the optimal temperature and concentration. Graphs are also provided to look at the change in temperature and concentration over time. Most of the variables in the model are self-explanatory. One important feature of our model is the option to have a change in the temperature of the feed stream or the concentration of A after a given time t. You do not need to input a value for these cells if there is no change in the feed; it just provides a convenient way to look at the change of temperature and concentration of A. You are also provided with a cell for the time step, Δt . Depending on what size time step you choose, you may need to choose a larger value if your graphs do not reach steady state. If this is the case, the output cells will tell you to increase the time step.

Applications to Process Control

The model created can account for many different situations. However, in process control, it is very important for the following to be considered.

1. **Exothermic reaction** - In industry, the exothermic reaction is typically of more importance to controls as there is a safety factor involved. (i.e. explosive reactor conditions)
2. **Volume constant** - For CSTRs, liquids are mostly used.

The model developed can assist in controlling an actual process in many ways including the following, but not limited to:

- Maintaining product concentration despite changes operating conditions
- Predicting and preventing explosive reactor conditions
- Flow rate optimization
- Input concentration optimization
- Operating temperature optimization
- Coolant temperature optimization

This model is very useful for many simple reactor situations and will aid in the understanding of a process and how dynamic situations will affect operation.

In real-life applications, the reaction may not be first order or irreversible or involve multiple reactions. In all these cases, we modify the Energy and Mass Balance with additional rate law considerations.

Example 6.4.1

You are contracted by WOW Chemical to control the operation of their 3000 L jacketed CSTR. They desire to create chemical B from chemical A at an optimal conversion. What is the temperature at which the optimal conversion is achieved?

A → B is a first order, irreversible reaction.

Some information about the process:

Feed stream temperature = 400 K

Coolant temperature = 350 K

Heat of reaction = -200 J/mol

Inlet concentration of A = 9 mol/L

Inlet flow rate = 4 kg/s

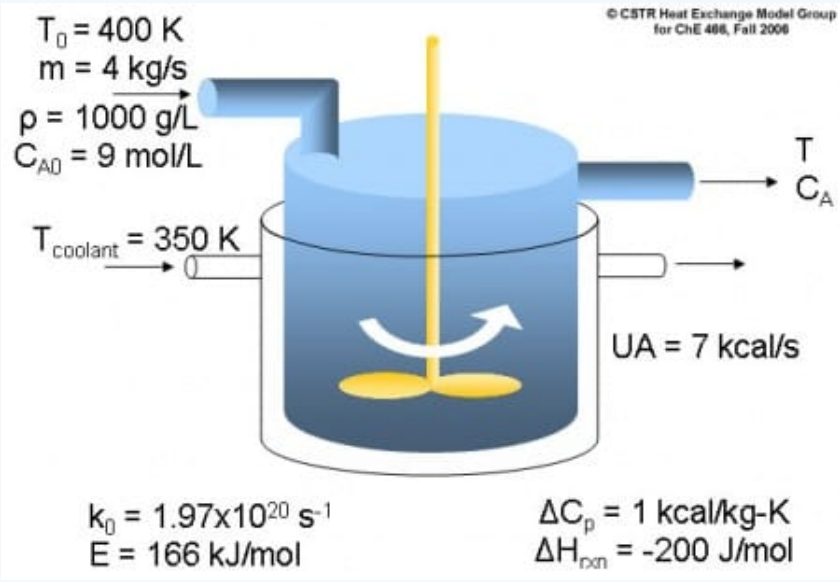
Density of A = 1000 g/L

UA of the heat exchanger = 7 kcal/s

Rate constant = 1.97×10^{20} s^{-1}

Activation energy = 166 kJ/mol

Overall change in Heat capacity = 1 kcal/kg-K.



Worked out Example 1

4.6.1 Answer to Example Problem #1

From using our handy-dandy worksheet, we see that the optimal conversion occurs at 368.4 K. This comes from inputting all of the given information from the problem statement into the Excel sheet, and then reading off the value for the optimal temperature. Also, make sure you are converting to proper units, such as from kJ to J, for use of our model (just to see if you were paying attention to units).

Example 1 Solution

	A	B	C	D
1	Enter Given Information HERE			Description
2	C_{A0}	9	mol/L	Initial Feed Concentration
3	$C_{A0,2}$	9	mol/L	Change Initial Feed Concentration (leave the same if this does not occur)
4	t for $C_{A0,2}$		s	Time when Initial Feed Conc changes
5	Δt	1	s	Time Step
6	ΔC_p	1000	J/kg-K	heat capacity of the reactor feed and product
7	E	166000	J/mol	Activation energy
8	m	4	kg/s	Mass feed rate and product rate
9	k_0	1.97E+20	s^{-1}	Rate constant
10	UA	7000	J/s	Heat Transfer Coefficient anc Area
11	T_0	400	K	Feed temperature
12	$T_{0,2}$	400	K	Change Initial Feed Temperature (leave the same if this does not occur)
13	t for $T_{0,2}$		s	Time when Initial Feed Temperature Changes
14	T_c	350	K	Coolant temperature
15	$T_{c,2}$	350	K	Change Initial Coolant Concentration (leave the same if this does not occur)
16	t for $T_{c,2}$		s	Time when Initial Coolant Temperature Changes
17	V	3000	L	Reactor volume
18	ρ	1	kg/L	Density of reactor feed and product (assumed constant)
19	ΔH_{Rx}	-200	J/mol	Heat of reaction
20				
21	Optimal Conversion Values			
22	T	368.377		K
23	C_A	6.301		mol/L

Worked out Example 2

1. In the previous example, the plant was operating in the daytime. At night, there is a drop of 5 K in feed stream temperature. In order to maintain the same concentration do we need to...

A. Decrease the temperature of the cooling water in the heat exchanger.

B. Increase the temperature of the cooling water in the heat exchanger.

C. Maintain the same temperature of the cooling water in the heat exchanger.

2. At night, the inlet flow rate of the system also decreases due to the plant not operating at full capacity. What happens to the temperature of the reaction if concentration of A in the feed stream stays the same?

A. Decrease

B. Increase

C. Stay the same

3. Finally, if the rate constant doubled, what effect will this have on the overall concentration if temperature is held constant?

A. Increase

B. Decrease

C. Stay the same

Answers to Example Problem #2

1: B. Since the temperature of the feed stream has decreased, we will need to increase the temperature of the cooling stream in order to maintain the same temperature at which the CSTR was operating at before. Using the Excel spreadsheet from Example 1 and lowering the feed stream temperature by 5 K, we can first see that the temperature of the CSTR has fallen to 366.5 K.

	A	B	C	D
1	Enter Given Information HERE			Description
2	C_{A0}	9	mol/L	Initial Feed Concentration
3	$C_{A0,2}$	9	mol/L	Change Initial Feed Concentration (leave the same if this does not occur)
4	t for $C_{A0,2}$		s	Time when Initial Feed Conc changes
5	Δt	1	s	Time Step
6	ΔC_p	1000	J/kg-K	heat capacity of the reactor feed and product
7	E	166000	J/mol	Activation energy
8	m	4	kg/s	Mass feed rate and product rate
9	k_0	1.97E+20	s^{-1}	Rate constant
10	UA	7000	J/s	Heat Transfer Coefficient and Area
11	T_0	395	K	Feed temperature
12	$T_{0,2}$	395	K	Change Initial Feed Temperature (leave the same if this does not occur)
13	t for $T_{0,2}$		s	Time when Initial Feed Temperature Changes
14	T_c	350	K	Coolant temperature
15	$T_{c,2}$	350	K	Change Initial Coolant Concentration (leave the same if this does not occur)
16	t for $T_{c,2}$		s	Time when Initial Coolant Temperature Changes
17	V	3000	L	Reactor volume
18	ρ	1	kg/L	Density of reactor feed and product (assumed constant)
19	ΔH_{Rx}	-200	J/mol	Heat of reaction
20				
21	Optimal Conversion Values			
22	T	366.524	K	
23	C_A	6.785	mol/L	

We can then adjust the temperature of the feed stream and see that it must be raised by 3 K in order to return to the original operating temperature of 368.4 K.

	A	B	C	D
1	Enter Given Information HERE			Description
2	C_{A0}	9	mol/L	Initial Feed Concentration
3	$C_{A0,2}$	9	mol/L	Change Initial Feed Concentration (leave the same if this does not occur)
4	t for $C_{A0,2}$		s	Time when Initial Feed Conc changes
5	Δt	1	s	Time Step
6	ΔC_p	1000	J/kg-K	heat capacity of the reactor feed and product
7	E	166000	J/mol	Activation energy
8	m	4	kg/s	Mass feed rate and product rate
9	k_0	1.97E+20	s^{-1}	Rate constant
10	UA	7000	J/s	Heat Transfer Coefficient and Area
11	T_0	395	K	Feed temperature
12	$T_{0,2}$	395	K	Change Initial Feed Temperature (leave the same if this does not occur)
13	t for $T_{0,2}$		s	Time when Initial Feed Temperature Changes
14	T_c	353	K	Coolant temperature
15	$T_{c,2}$	353	K	Change Initial Coolant Concentration (leave the same if this does not occur)
16	t for $T_{c,2}$		s	Time when Initial Coolant Temperature Changes
17	V	3000	L	Reactor volume
18	ρ	1	kg/L	Density of reactor feed and product (assumed constant)
19	ΔH_{Rx}	-200	J/mol	Heat of reaction
20				
21	Optimal Conversion Values			
22	T	363.470	K	
23	C_A	6.276	mol/L	

2: A. Since there is less feed coming into the CSTR, the optimal temperature of the reactor will decrease as well. This can again be seen by using our Excel model from Example 1. We know that originally the temperature of the reactor was at 368.4 K. By decreasing the flow rate to 3 kg/s, we can see that the temperature of the reactor decreases by over 3 K down to 365.1 K.

	A	D	C	D
1	Enter Given Information HERE			Description
2	C_{A0}	9	mol/l	Initial Feed Concentration
3	$C_{A0,2}$	9	mol/L	Change Initial Feed Concentration (leave the same if this does not occur)
4	t for $C_{A0,2}$		s	Time when Initial Feed Conc changes
5	Δt	1	s	Time Step
6	ΔC_p	1000	J/kg-K	heat capacity of the reactor feed and product
7	E	166000	J/mol	Activation energy
8	m	3	kg/s	Mass feed rate and product rate
9	k_0	1.97E+20	s^{-1}	Rate constant
10	UA	7000	J/s	Heat Transfer Coefficient and Area
11	T_0	400	K	Feed temperature
12	$T_{0,2}$	400	K	Change Initial Feed Temperature (leave the same if this does not occur)
13	t for $T_{0,2}$		s	Time when Initial Feed Temperature Changes
14	T_c	350	K	Coolant temperature
15	$T_{c,2}$	350	K	Change Initial Coolant Concentration (leave the same if this does not occur)
16	t for $T_{c,2}$		s	Time when Initial Coolant Temperature Changes
17	V	3000	L	Reactor volume
18	ρ	1	kg/L	Density of reactor feed and product (assumed constant)
19	ΔH_{Rx}	-200	J/mol	Heat of reaction
20				
21	Optimal Conversion Values			
22	T	365.139		K
23	C_A	6.610		mol/L

3: B. We see in the equations for the differential equations that the rate constant has a negative effect on the conctration of A. By doubling its value, we will then be decreasing C_A as well. This can also be seen in the Excel model from Example 1. We see that the original concentration of A is 6.3 mol/L. We can then double the value of the rate constant, and we see that the concentration of A drops down to 4.8 mol/L, which is what we had predicted would happen.

	A	B	C	D
1	Enter Given Information HERE			Description
2	C_{A0}	9	mol/L	Initial Feed Concentration
3	$C_{A0,2}$	9	mol/L	Change Initial Feed Concentration (leave the same if this does not occur)
4	t for $C_{A0,2}$		s	Time when Initial Feed Conc changes
5	Δt	1	s	Time Step
6	ΔC_p	1000	J/kg-K	heat capacity of the reactor feed and product
7	E	166000	J/mol	Activation energy
8	m	4	kg/s	Mass feed rate and product rate
9	k_0	3.94E+20	s^{-1}	Rate constant
10	UA	7000	J/s	Heat Transfer Coefficient and Area
11	T_0	400	K	Feed temperature
12	$T_{0,2}$	400	K	Change Initial Feed Temperature (leave the same if this does not occur)
13	t for $T_{0,2}$		s	Time when Initial Feed Temperature Changes
14	T_c	350	K	Coolant temperature
15	$T_{c,2}$	350	K	Change Initial Coolant Concentration (leave the same if this does not occur)
16	t for $T_{c,2}$		s	Time when Initial Coolant Temperature Changes
17	V	3000	L	Reactor volume
18	ρ	1	kg/L	Density of reactor feed and product (assumed constant)
19	ΔH_{Rx}	-200	J/mol	Heat of reaction
20				
21	Optimal Conversion Values			
22	T	363.486		K
23	C_A	4.818		mol/L

References

- Bequette, B. Wayne. *Process Dynamics Modeling, Analysis, and Simulation*, New Jersey: Prentice Hall.
- Fogler, H. Scott. *The Elements of Chemical Reaction Engineering*, Prentice Hall. ISBN 0-13-047394-4
- Kravaris, Costas. *Chemical Process Control: A Time-Domain Approach*, The University of Michigan.
- Riggs, James B.; Karim, M. Nazmul. *Chemical and Bio-Process Control*, Texas Tech University, Ferret Publishing.

Contributor

Authors: Jason Bourgeois, Michael Kravchenko, Nicholas Parsons, Andrew Wang

This page titled 6.4: ODE and Excel CSTR model with heat exchange is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

6.5: ODE and Excel model of a Simple Distillation Column

Introduction

Distillation is a commonly employed separation technique bases on difference in volatilities. The modern form of distillation as is is known today may be credited to early Arab alchemist, Jabir ibn Hayyan and the development of one of his inventions, the alembic. The distillation apparatus is commonly referred to as a still and consists of a minimum of a reboiler in which mixture to be separated is heated, a condenser in which the vapor components are cooled back to liquid form, and a receiver in which the concentrated liquid component fractions are collected. Ideally, distillation is governed by the principles of Raoult's Law and Dalton's Law. Dalton's Law states that for a mixture, the total vapor pressure is equal to the sum of the individual vapor pressures of the pure components which comprise this mixture. The relationship giving the vapor pressure of a volatile component in a mixture, P_A, is Raoult's Law and is governed by the following equation:

$$P_A = X_A P_A^o$$

Where X_A is the mole fraction of component A in the mixture and $P_A°$ is the vapor pressure of pure component A. This ideal model is based on a binary mixture of benzene and toluene but for other mixtures severe deviations from Raoult's Law may be observed due to molecular interactions. For these aforementioned mixtures where the components are not similar the only accurate alternative is obtaining vapor-liquid equilibrium by measurement.

In simple distillation, two liquids with differing boiling points are separated by immediately passing the vapors from the reboiler to a condensing column which condenses the vapor components. As a result the distillate is not usually pure but its composition may be determined by Raoult's Law at the temperature and pressure at which the vapors were boiled off. Consequently, simple distillation is usually used to separate binary mixtures where the boiling temperatures of the individual components are usually significantly different or to separate volatile liquids from non-volatile solids.

A reference of terms used in distillation reference are included at the end of this article.

Distillation Control

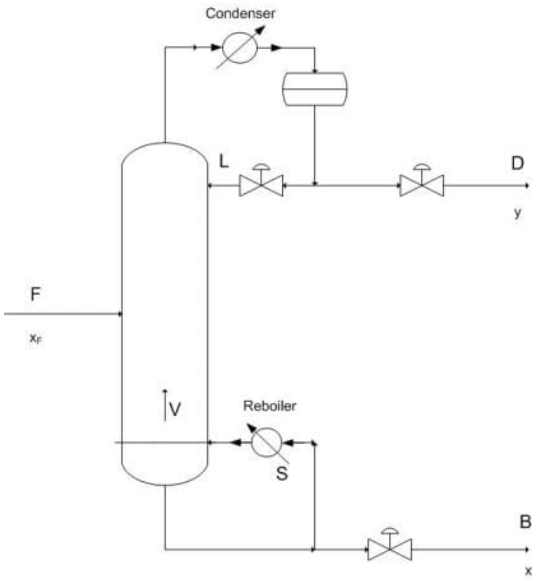
Distillation columns comprise an enormous amount of the separation processes of chemical industries. Because of their wide range of uses in these industries and because their proper operation contributes to product quality, production rates and other capital costs, it is clear that their optimization and control is of great importance to the chemical engineer. Distillation control becomes problematic because of the wide variety of thermodynamic factors stemming from the separation process. For example:

- Separations deviate from linearity of equations as product purity increases.
- Coupling of process variables occurs when compositions are controlled.
- Disturbances occur due to feed and flow agitation
- Efficiency changes in trays lead to non-steady state behavior.

To improve upon distillation control you must be able to characterize these potential problems and realize when they occur because they lead to dynamic behavior of the column. Of key importance to control is the maintenance of material and energy balances and their due effects on the column. Shown below is a schematic of a simple binary distillation column. Using the material balance formulas

$$\frac{D}{F} = \frac{z - x}{y - x}$$

where z, x, and y are the feed, bottoms and distillate concentrations respectively, you find that as D (Distillate) increases, its purity decreases. This leads to the idea that purity level varies indirectly with the flow rate of that product. Energy input is also key because it determines the vapor flow rate (V) up the column which has direct effects on the L/D ratio (reflux ratio) and hence relates to an increase in the amount of separation occurring. To summarize, energy input determines the amount of separation, while material flow relates the ratio of separation in the products.



Vapor liquid dynamics within the column also contributes to the theory behind process control because of a few important relations. Changing V (by changing the reboiler energy), causes an extremely rapid response in the overhead composition, while changing the reflux ratio requires a longer response to its effect on the reboiler.

In lower pressure columns, a phenomena known as entrainment or flooding occurs in which liquid is blown up into trays instead of dropping down into trays. This significantly decreases separation efficiency and therefore less product gain occurs. Using a packed column in these low pressure applications provides greater efficiency over tray columns and also allows for the faster accomplishment of a steady state profile. Controlling the occurrence of entrainment in either case is another crucial aspect which should be recognized when designing control systems for columns.

Regulatory Controls

For the distillation process it is imperative that regulatory controls such as level, flow, and pressure controllers are functioning properly to further ensure the effectiveness of the product composition controllers.

In terms of regulatory control, *level controls* are used to maintain specified levels in the reboiler, accumulator, and in the case of a distillation column with two columns due to high tray numbers for a single column, also maintain the level in an intermediate accumulator. Inept use of level controls may lead to problems elsewhere in the distillation process. For example, poor level control on the accumulator and reboiler may lead to problems with composition control for material balance control configurations. Also if the reboiler duty is maintained by one of these level controllers and the controller causes oscillation in the reboiler, consequently cycling may also occur in column pressure.

Flow controllers are used to manipulate and maintain desired flow rates of the reflux, distillate and bottoms products, and the heating medium employed in the reboiler. The setpoints of this specific type of controller is determined by the various composition and level controllers in the process.

Pressure controls are located to the top of the distillation column in the distillate vicinity. Here the column overhead pressure, caused by the accumulation of components in the vapor phase, acts as an integrator causing a change in the level of the accumulator. This pressure may be controlled using a variety of methods.

Pressure Control through Condenser Operation

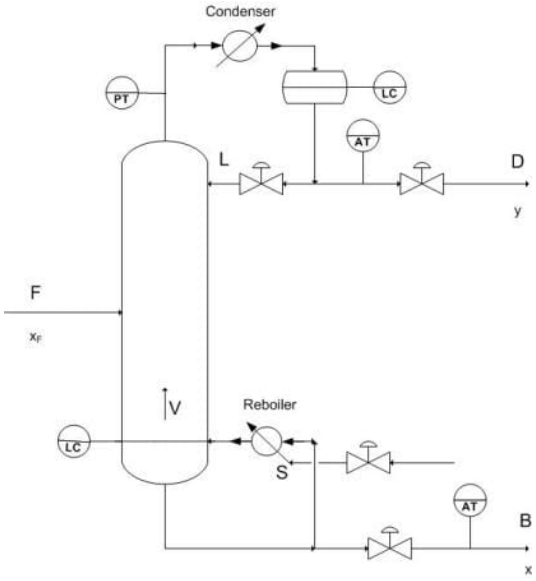
1. Maximize cooling water flow rate to the condenser (operate at minimum column pressure)
2. Adjust the rate of condensation of the overhead (for example by adjusting the flow rate of refrigerant to the condenser)
3. Adjust level of liquid in the condenser (changes heat transfer area)

Pressure Control through Accumulator Operation

1. Purge vapor from the overhead accumulator
2. Directly changing amount of material phase (for example by pumping in inerts)

Composition and Constraint Control

Because distillation requires a desired product concentration or flowrate, constraint control is used to ensure the desired operation conditions by having setpoints designated for the requirements of the system. The constraint is usually a concentration and the control of this concentration can vary depending on the application. In almost all industry applications only one product concentration is controlled, while the other is allowed to vary. This is known as Single Composition control and is much easier to achieve and maintain than the Dual composition control which specifies both product concentrations. The advantage to Dual Control however being increased energy efficiency because of increased separation. P&ID of common control placement used for composition control are shown below. Control lines are not included in this diagram because of the various number of control possibilites for a system.



Setting constraints on a column allows for proper control of the product as well as points to issues of safety and maintenance. The most common constraint controls are in the maximum reboiler and condenser duty which results from a number of variables including fouling , improper valve sizing, and excessive increases in feed. Other common constraints are flooding and weeping points which indicate incorrect L/V ratios and can be compensated for by adjusting the pressure drop across the column.

The most common adjustments for constraint control involve changing the reboiler duty to satisfy the constraints and almost always follow with subsequent fall back mechanisms to ensure product quality.

1. Ensuring single composition control
2. Reducing feed rate
3. Increasing product purity setpoints

Effective Distillation Control

Before doing an in depth analysis of distillation control, it is of paramount importance that the following basics are attended to.

- Firstly, ensure that regulatory controls are indeed functioning congruously.
- For changes in reflux temperature, employ the use of reflux controls.
- Make sure to check and evaluate analyzer deadtime, accuracy, and reliability. This is required to account for the lagtime from the product composition analyzer when using feedback control to control feed flow, reflux ratio and reboiler power. Refer to for deadtime. To select proper analyzer refer to
- Ensure that any thermistors or RTD's employed to measure tray temperatures for composition inference are fully operational and correctly situated. Care should be taken here to ensure that pressure corrected temperatures are used. Refer to the following for temperature sensors

- When streams such as D, B, L, and V are used as manipulated variables for composition control, they should be changed with respect to the measured feed rate when column feed rate changes are a common disturbance.

Summary of Distillation Control Methods

Control Parameter	Example of Control Method
Distillate flow rate (D)	Flow controller (Setpoint controlled by accumulator level)
Bottoms flow rate (W)	Flow controller (Setpoint controlled by level in bottom of column)
Reflux flow rate (L_D)	Flow controller (Setpoint controlled by top tray temperature)
Reboiler Steam flow rate	Flow controller (Setpoint controlled by bottoms composition analyzer)
Feed flow rate	Flow controller (Setpoint manually entered based on unit operations)
Distillate purity (x_D)	Reflux flow controller (Setpoint controlled by top tray temperature)
Bottoms purity (x_W)	Steam flow controller (Setpoint controlled by online analyzer)
Column pressure	Purge flow controller (Setpoint controlled by column pressure)

Control Problems and Disturbances

Feed Composition & Feed Flow Upsets

In order to properly determine the ultimate product purity and flow you must consider the impact of disturbances in the column system. The most significant, however most easily remedial of these disturbances is a feed composition upset, in which there is a change in the feed composition resulting in a major disturbance in the product composition. Hence, configuration of a distillation column must consider the control regulations of such a feed upset. Feed flow upsets are regulated using ratio control of L/F, D/F, V/F and B/F through level and reflux ratio sensors.

Feed and Reflux Enthalpy Disturbances

Feed Enthalpy upsets become an issue for columns operating at low reflux ratios and cause a large deviation from expected product concentrations due to the changes in vapor and liquid flowrates in the column. The usual compensation for this is using a feed heat exchanger to control the proper enthalpy to the column. In most cases, the feed is preheated before entering the column. By adjusting the duty of the preheater (i.e. decreasing or increasing the heating medium flow rate), a constant vapor/liquid ratio in the feed can be maintained. Rapid changes to the external conditions of a column, (especially large temperature deviations in rainstorms) can cause a subcooled reflux, changing the composition of the products. Reflux control can amend this properly.

Steam and Column Pressure Upsets

The most severe disturbance occurring in a distillation column occurs when there is a loss of steam pressure in the reboiler. A sharp drop in steam pressure results in a drop in reboiler effectiveness and therefore a huge increase in the impurity of the product. This can be avoided using an override control loop for this particular occurrence. Due to the effect of pressure on the relative volatility of the components in the system, a disturbance in column pressure leads to altered product quality. This can be effectively maintained by a composition controller to compensate for these pressure differences.

ODE Modeling of a Distillation Column

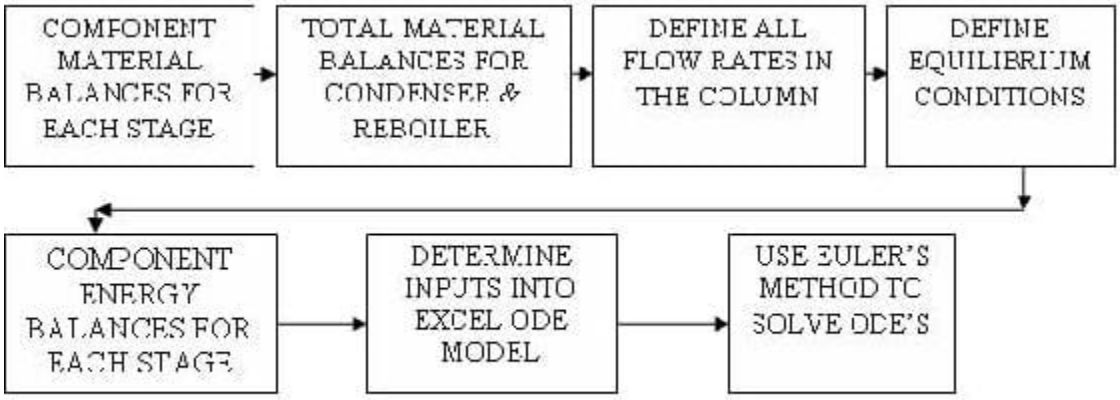
There are two methods used for modeling a distillation column: dynamic modeling and steady state modeling. The key difference between the models is that dynamic modeling is used to monitor changes in the distillation column as a function of time, while the steady state model looks at a given set of conditions at one particular time (i.e. when the column is at steady state).

Dynamic Model

Dynamic modeling of a distillation column may be used for a variety of different reasons: monitoring variations in the column as a result of feed changes, to predict the effects of tray fouling, and to predict when flooding with occur. The dynamic model

allows the user to improve distillation control by being able to deal with disturbances that cause upsets to the column's normal operation.

The algorithm for developing a dynamic distillation column model is as follows:



Step 1) Write component material balance for each stage in the column.

$$accumulation \quad liquid\ entering\ stage\ i - vapor\ entering\ stage\ i - liquid\ leaving\ stage\ i - vapor\ leaving\ stage\ i$$

Component material balance for all stages, except the feed tray, overhead condenser, and reboiler:

$$\frac{dM_i x_i}{dt} = L_{i-1}x_{i-1} + V_{i+1}y_{i+1} - L_i x_i - V_i y_i$$

Assumption: For simplicity, accumulation in the each stage is constant;

$$\frac{dM_i}{dt} = 0$$

Simplified component material balance for each stage (only composition changes with time):

$$M_i \frac{dx_i}{dt} = L_{i-1}x_{i-1} + V_{i+1}y_{i+1} - L_i x_i - V_i y_i$$

The following are examples of equations used in the Excel Interactive ODE Distillation Column Model, which are provided to help the user understand how the model works.

ODE used to solve for the liquid composition leaving tray 2 (rectifying section):

$$\frac{dx_2}{dt} = \frac{1}{M_2}\left[L_1x_1 + V_3y_3 - L_2x_2 - V_2y_2\right]$$

ODE used to solve for the liquid composition leaving tray 5 (stripping section):

$$\frac{dx_5}{dt} = \frac{1}{M_5}\left[L_4x_4 + V_6y_6 - L_5x_5 - V_5y_5\right]$$

Overhead condenser component balance:

$$\frac{dx_D}{dt} = \frac{1}{M_D}\left[V_1\left(y_1 - x_D\right)\right]$$

Feed tray component balance:

$$\frac{dx_3}{dt} = \frac{1}{M_3}\left[L_2x_2 + V_4y_4 - L_3x_3 - V_3y_3\right]$$

Reboiler component balance:

$$\frac{dx_W}{dt} = \frac{1}{M_W}\left[L_6x_6 - Wx_W - V_7y_7\right]$$

Step 2) Write total material balances around condenser and reboiler

Condenser material balance:

- Assumption 1: Total condenser (all vapor from the top of the column is condensed into a liquid).
- Assumption 2: Overhead accumulator liquid level remains constant.

$$D = [V_1 + L_D]$$

Reboiler material balance:

$$W = [F - D]$$

For these equations to work, the user must specify:

- reflux flow rate (mol/min)
- bottoms flow rate (mol/min).

Step 3) Define all flow rates

Vapor Leaving Feed Stage:

$$V_3 = V_4 + F\left(1 - q_F\right)$$

Liquid Leaving Feed Stage:

$$L_3 = L_2 + F\left(q_F\right)$$

Vapor flow rates in stripping section:

Assumption: Equimolal overflow for vapor in stripping section

$$V_4 = V_5 = V_6 = (V_7)$$

Vapor flow rates in rectifying section:

Assumption: Equimolal overflow for vapor in rectifying section

$$V_1 = V_2 = (V_3)$$

Liquid flow rates in rectifying section:

Assumption: Equimolal overflow for liquid in rectifying section

$$L_2 = L_1 = (L_D)$$

Liquid flow rates in stripping section:

Assumption: Equimolal overflow for liquid in stripping section

$$L_6 = L_5 = L_4 = (L_3)$$

Step 4) Define equilibrium conditions

The binary system used in the Excel ODE model is a benzene-toluene system. The equilibrium data for this system was put in the model and the relative volatilities were calculated for various equilibrium compositions.

Relative Volatility (from equilibrium data):

$$\alpha = \frac{y_{\text{benzene}}\, x_{\text{toluene}}}{x_{\text{benzene}}\, y_{\text{toluene}}}$$

where α is defined as the relative volatility of the two components in the system.

These relative volatilities were plotted against temperature and linear regression was used to fit the data.

Relative volativity as a function of temperature:

$$\alpha = [-0.009T + 3.3157]$$

This equation models how the separation changes on each tray as a function of tray temperature, which decreases up the column.

Equilibrium Vapor Composition for each stage:

Assumption: Trays in the Column are 100% efficient (vapor and liquid leaving a tray are in equilibrium)

$$y_i = \frac{\alpha x_i}{1 + (\alpha - 1)x_i}$$

Replacing alpha with the temperature dependent equation shows how tray temperature affects the amount of benzene in the vapor leaving each tray.

Step 5) Write component energy balances for each stage.

The ODE energy balances are essential for the dynamic model to run properly. Mass transfer occurs within the column because the temperature varies from the top of the column to the bottom thereby allowing separation of the components in the system.

Since the energy input into the column is added in the reboiler, the reboiler ODE is the first equation entered in the model. In our model, this is given as:

$$\frac{dT_7}{dt} = \frac{1}{M_W}\left[L_6 x_6 - W x_W\right]\left[T_6 - T_7\right] + \frac{q_r}{M_W c_p}$$

The next step is adding energy balances for each subsequent stage in the distillation column. The only stage in the column which has a slightly different energy ODE is the feed stage, given by:

$$\frac{dT_3}{dt} = \frac{1}{M_3}\left[\left[L_2 x_2\right]\left[T_2 \quad T_3\right] + \left[V_4 y_4\right]\left[T_4 \quad T_3\right] - \left[L_3 x_3\right]\left[T_2 \quad T_3\right] + \left[V_3 y_3\right]\left[T_4 \quad T_3\right] + \left[F x_{feed}\right]\left[T_{feed} \quad T_3\right]\right]$$

The last energy balance is around the condenser.

Assumption: Reflux return temperature is constant (overhead condenser duty varies to compensate for this).

Step 6) Determine inputs into ODE model

Once all of the equations have been put into the model, all remaining unknown variables must be placed in a section so that the user can specify these input values when running the model. For the Excel ODE distillation model, the users inputs include:

1. Feed flow rate
2. Mole fraction of light key in the feed
3. Reflux flow rate

4. Condenser, reboiler, and tray levels
5. Phase of the feed (q-value)
6. Feed temperature
7. Integration step size

To model effects of disturbances, the user may also change these input values:

- Feed flow after 200 time steps
- Feed composition after 600 time steps

Step 7) Use Euler's Method to solve the ODE's

This step involves using Euler's method. to integrate each ODE over each timestep in the interval to solve for the parameter value at the next time step. Creating a plot of these values versus time allows the user to see how changes in the input values effect parameters such as distillate and bottoms composition or flowrates.

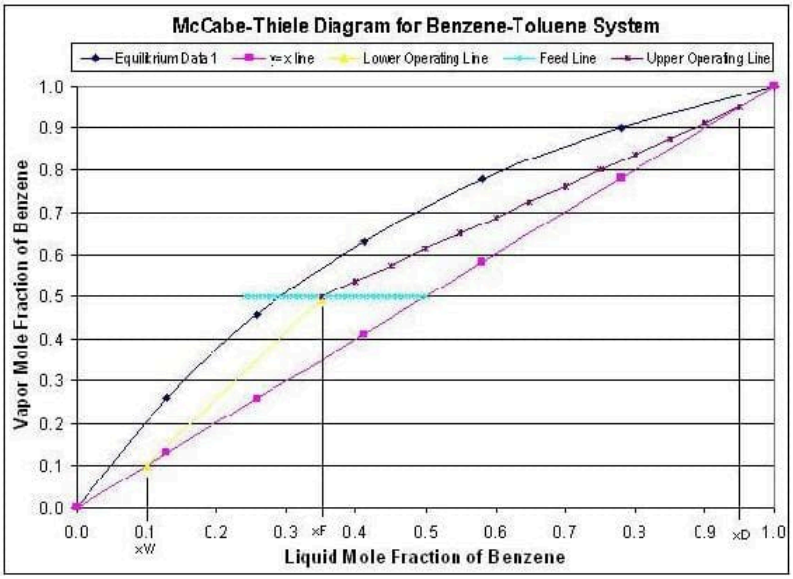
Additional Considerations for Dynamic Distillation Modeling

The Interactive Excel ODE Distillation Column Model does not take into account heat effects within or surrounding the column. It may be beneficial in modeling an actual column to determine the optimal temperatures that feed and reflux should enter at to achieve the greatest possible separation. The energy input into the reboiler is another consideration that may need to be modeled for economic purposes.

Also, to tune the dynamic distillation model with greater precision, additional parameters and equations may be added. One example of this would be to add a K_D or K_U input value to control the levels in the bottom of the column or overhead condenser.

Steady State Model

The steady state model which is described by using a McCabe-Thiele diagram shows the theoretical stages in a binary component distillation column. An example of a McCabe-Thiele diagram is shown below.



The upper operating line is a graphical representation of the vapor/liquid dynamics in each stage in the rectifying section of the column (above the feed stage), while the lower operating line represents the vapor/liquid dynamics in the stripping section of the column (below the feed stage). The starting point of the upper operating line represents the distillate composition and the bottom point of the lower operating line represents the bottoms compostion. The feed line shows the entering feed composition as well as whether the feed is vapor, liquid, or a combination of the two. To develop this steady state model, one must know the components in the system, so that equilibrium data can be obtained. Also, this model requires that the following parameters must

be known:

- reflux ratio
- distillate composition
- bottoms composition
- feed composition
- feed phase

The general equations used in the steady state model are given in the table below:

Media:EquationsforODEModeling.doc

The McCabe-Thiele diagrams are excellent for modeling steady state operation, but they do not describe how disturbances affect the column operation. For each change in a particular parameter, a separate McCabe-Thiele diagram must be made. The dynamic model, although more complex than the steady state model, shows how a column operates during start-ups, when disturbances occurs, and where steady state conditions occur. Therefore, the dynamic ODE model of a distillation column can allow the user to how product purities and flow rates change with time.

Glossary of Terms

- M_i = Molar holdup on tray i
- L_{i-1} = Liquid molar flowrate into tray i
- L_i = Liquid molar flowrate leaving tray i
- V_{i+1} = Vapor molar flowrate entering tray i
- V_i = Vapor molar flowrate leaving tray i
- x_i = mole fraction of light component in the Liquid phase of Tray i
- y_i = mole fraction of light component in the Gas phase of Tray i
- B = Bottoms flowrate
- D = Distillate flowrate
- f = Feed flow rate
- α = Relative volatility of Benzene-Toluene system.
- q = Vapor Liquid compostion value

Our Excel Distillation Model

Below you will find a link for our distillation model in Excel. It is an exceptionally large download so it may take some time if you are not using a high speed internet connection. Because it takes some time for the iterations to run to completion, calculations are not continuous within the spreadsheet. Be patient as it may take a few seconds for the calculations to run when operational parameters change. Also, because of the nature of the iterative process Excel uses to determine the cell values, large deviations may cause the model to crash resulting in #Error# for many cells. If this occurs, close the model and re-open it from the website to try different parameters.

In order to demonstrate dynamic changes occuring during operation, step changes are incorporated into the Excel model. To show the effects of a feed change, input the change into cell C-18. This change will occur at t=2min for a length 2 minutes and then the feed flowrate will revert back to the original value. To demonstrate a change in the feed compostion at t= 6min, input the change into cell C-19 on the Excel model. This compostion change will be constant over the remainder of the time and will not revert to the previous value. Notice the changes occuring in the graphs as these step changes occur. The Excel model can be found here => Media:ODEDistillationModel-Final.zip

This model has a specified control range because typical operation of distillation columns requires operation within a specific range of values where input parameters marginally vary. It is recommended to use the following variables for the initial model.

- Feed Flow to Column = 50.00
- Feed Composition (xF) = 0.5
- Reflux Flow = 15
- Bottoms Flow @ T= 0 = 35.00
- Mtray = 10

- Mcondenser = 50
- Mreboiler = 50
- qF = 0.4
- Feed Temperature = 80
- Reflux Temp = 80
- Initial Column Temp = 80
- Qreboiler = 100

Avoid extreme changes in the initial values, otherwise a significant error may occur causing the model to crash. The principal cause of this is due to the heat effects which are taken into account in the model and the fact that the number of trays on the column is set at 6.

Terms Commonly Used in Distillation

- **Active tray area:** The region of the tray where the upward moving vapor comes into contact with the downward flowing liquid (where the mass transfer occurs)
- **Downcomer:** Area on side of trays where liquid flows down through the distillation column
- **Disturbance (wrt distillation):** Any minor change in the distillation column caused by an external or internal source that causes product variability
- **Flooding:** Liquid from the active tray area is carried up into the vapor stream (occurs at low L/V ratios)
- **Ratio Control:** Controlled ratio of two manipulated variables
- **Tray Fouling:** Active area of tray is deminished, thus reducing separation efficiency within the column
- **Upset:** Any major change in the distillation caused by an external source that produces erratic column operation that requires manual override to gain control
- **Weeping Points:** Liquid from the active tray area seeps downward through the tray instead of flowing through the downcomer (occurs at high L/V ratios)

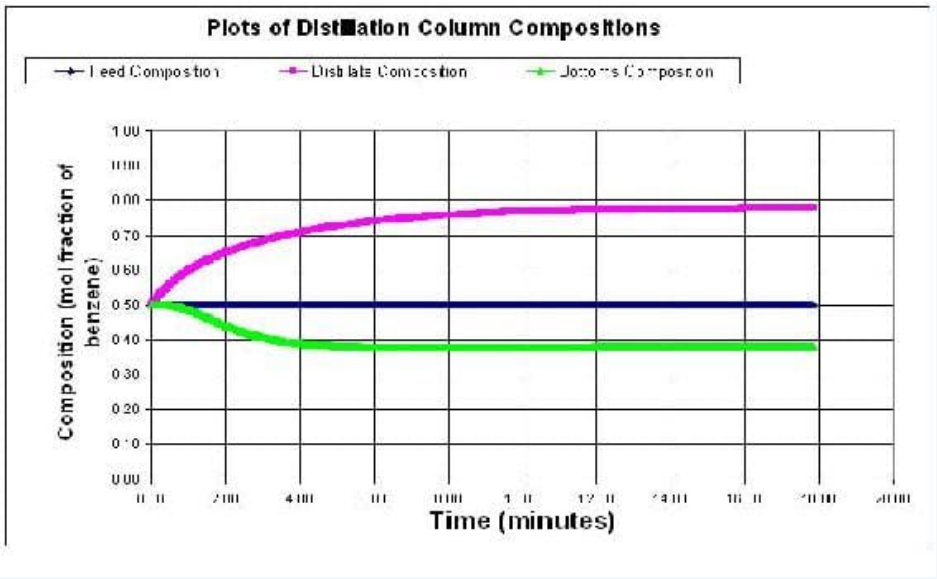
Example 6.5.1

Using the Excel Model of a Distillation Column posted above, see how a change in the feed concentration upsets the steadystate equilibration of the inital flow and how the column thereby adjusts to a change in feed composition at the time step indicated. Then determine the effect of changing the feed flowrate and reflux flow at the indicated times during the time the column is running. Also, determine the effect of making a larger step size for use in Eulers method.

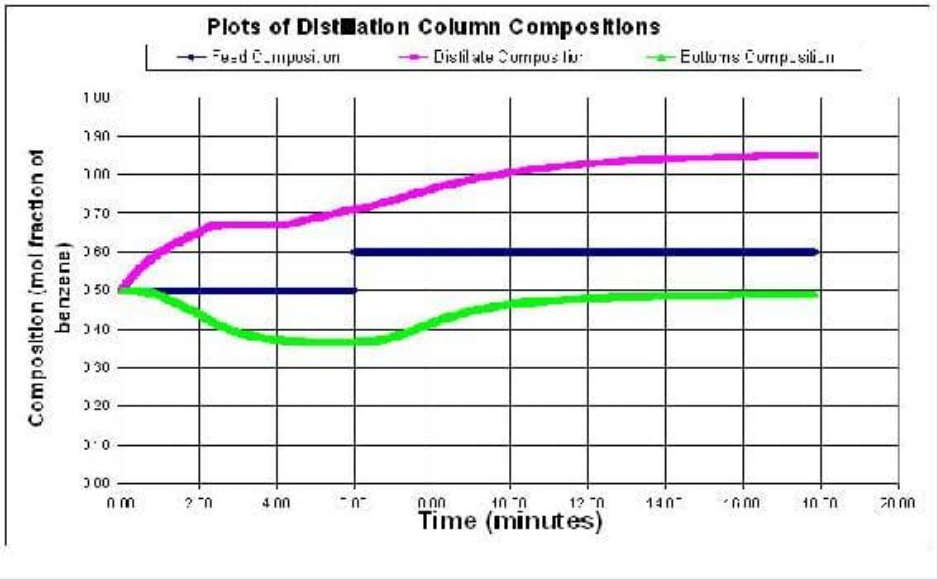
Remember: an excessive change may cause the distillation column to blow up, so take care in determining reasonable changes. Input the *changes* in the flowrate or concentration into C17-C20 of the spreadsheet. Input the new step size in cell D21.

Solutions

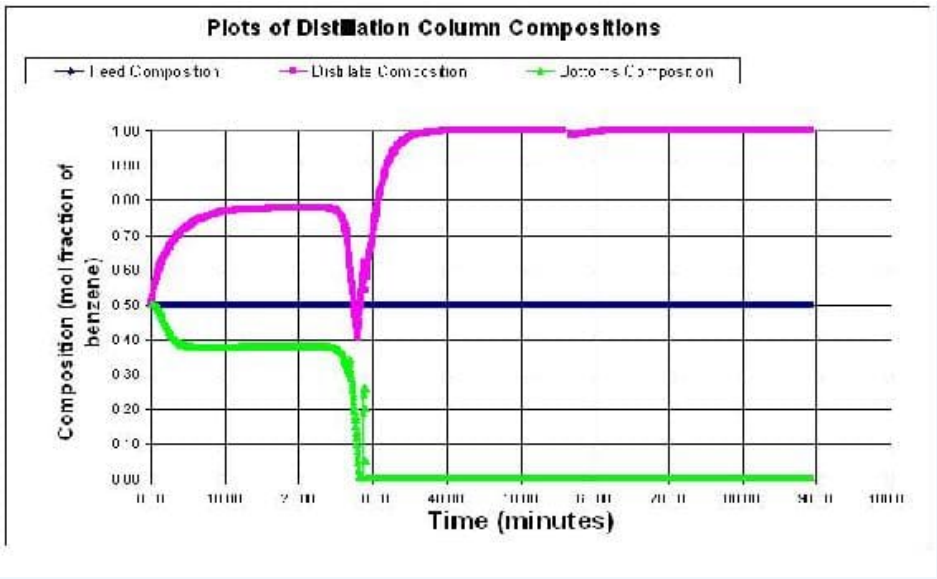
The normal concentration output for the distillation column is shown below. Feed flow rate is 50 mol/min, reflux flowrate is 15 mol/min, feed composition is 0.5 mol fraction of the light component. Step size is 0.01 minutes.



Students should see that an impulse change in flowrates does not really change the steady state profile of the column. However, a change in feed concentration will cause the column to jump and then equilibrate for the new concentration values. Below is a graph with a change in the feed flowrate to 100 mol/min, and a concentration change to 0.6 of the light component.



Larger step sizes will cause the model to crash because the error associated with the derivatives becomes larger with each derivative taken. Below the step size is 0.05 minutes. At some point the derivatives become uselsess data and therefore the distillation column model blows up with outrageous values as shown below.



Example 6.5.2

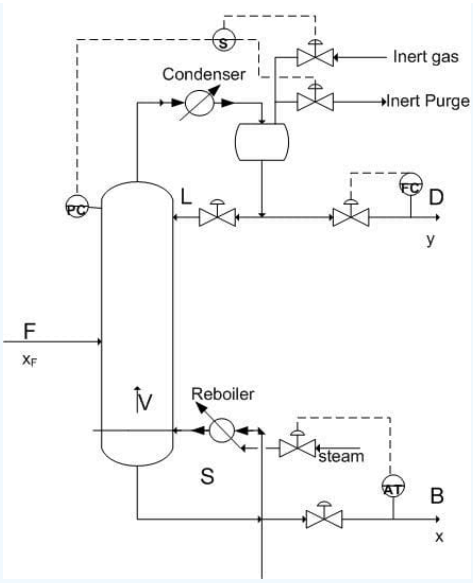
As a new hire at a specialty chemical company you are placed on project where you are asked to design and specify the controls on a state of the art distillation column used to separate a binary system. Your supervisor gives you a few clues to the control mechanism based on the chemistry occurring.

- The salable component is the heavy component and the mixture has a low relative volatility.
- The light component is not usable at all, however it is extremely toxic and must be controlled precisely.
- This is a small operation which requires a low pressure separation.
- The facility has had issues with the onsite steam that is typically used to heat the reboiler.

Draw the P&ID model of how you would model this column, mention all the items you must consider for the process and what you would do to control them.

Answer

Since the heavy component control is desired you will want a feedback loop based on the bottoms composition controlling reboiler duty. Also note how precise control of the vapor flowrate up the column can be best accomplished through reboiler control. Because it is low pressure you must begin to worry about entrainment and the possibility of flooding the column. Therefore there should be pressure sensors controlling the pressure in the column to maintain it at optimal levels using a packed column instead of a tray column to produce a better separation of the components. Because onsite steam has been shown to be a problem you would probably want to have a redundancy built into the reboiler steam or perhaps have its own autonomic steam delivery because this could potentially jeopardize the quality of your product.



References

- Bequette, B. Wayne. Process Dynamics Modeling, Analysis, and Simulation, New Jersey: Prentice Hall
- Perry, Robert H. Perry's Chemical Engineers' Handbook, 7th Ed. McGraw Hill. ISBN 0-07-049841-5
- Riggs, James B.; Karim, M. Nazmul. Chemical and Bio-Process Control, 3rd Ed., Texas Tech University, Ferret Publishing

Contributors and Attributions

- Authors: Jennifersue Bowker, George Cater, Kibwe John

This page titled 6.5: ODE and Excel model of a Simple Distillation Column is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

6.6: ODE and Excel model of a Heat Exchanger

Introduction

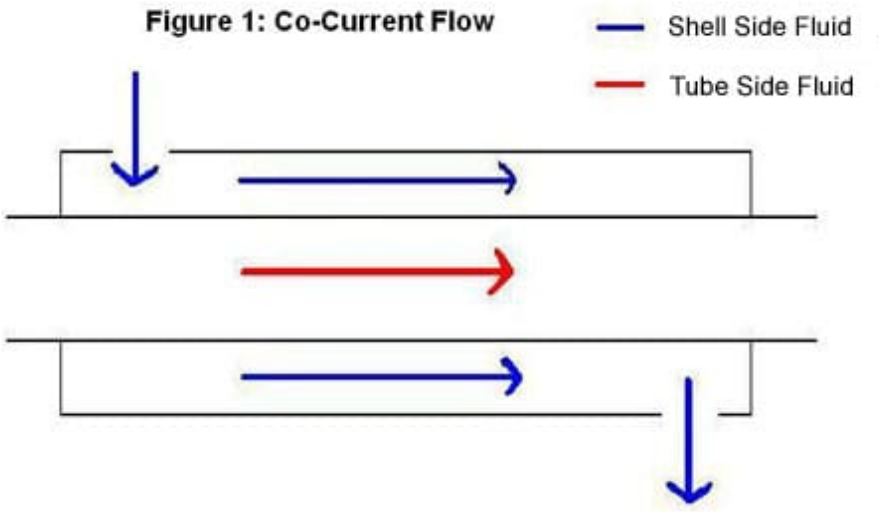
In process industries, heat exchangers are designed to transfer heat from one fluid to another. Heat exchangers have many different applications, especially in chemical processes, air conditioning, and refrigeration. They are classified by their design and type of flow. In some types of heat exchangers, the two fluids are separated by a wall or membrane and the heat transfer takes place by both convection and conduction. In another, less common type of exchanger, the two fluids physically come into contact with each other as the heat transfer occurs.

Since heat exchangers have a wide variety of applications and are commonly used in industry, control of the system is essential. A dynamic model may be created to allow the chemical engineer to optimize and control the heat exchanger. By utilizing this model, predictions can be made about how altering the independent variables of the system will change the outputs. There are many independent variables and considerations to account for in the model. If done so correctly, accurate predictions can be made about the system.

Types of Heat Exchangers

1. Double - Pipe Heat Exchanger

A double-pipe heat exchanger is the simplest type of heat exchanger and can operate with co-current (Figure 1) or counter-current (Figure 2) flow. The design consists of a single small pipe (tube-side) inside of a larger one (shell-side). A co-current heat exchanger is most commonly used when you want the exiting streams to leave the exchanger at the same temperature. A counter-current heat exchanger is used more often than co-current because they allow for a more efficient transfer of energy.



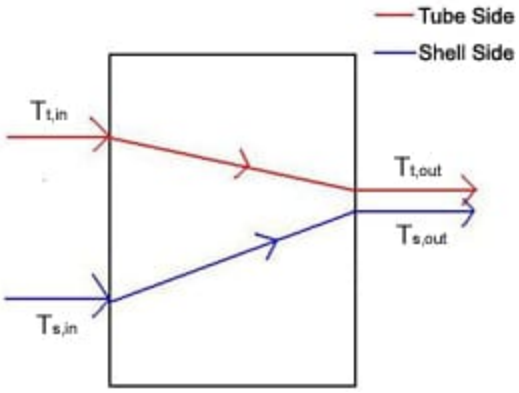
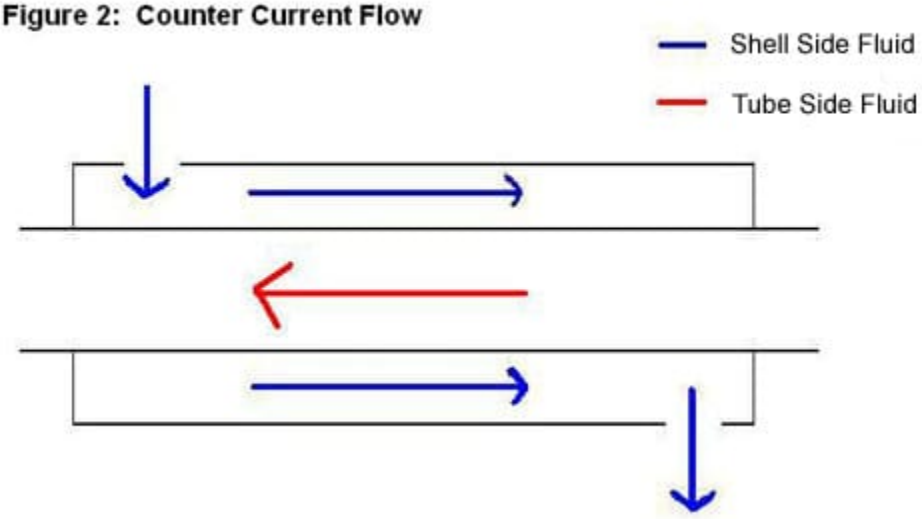
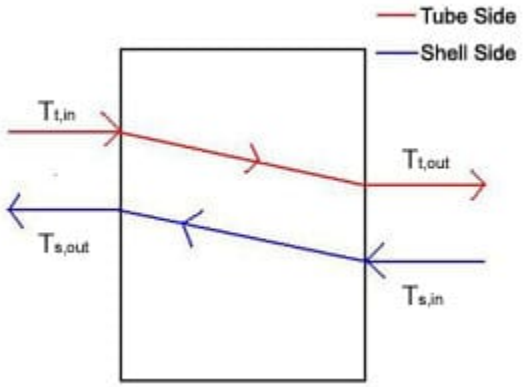


Figure 2: Counter Current Flow





2. Shell-and-Tube Exchanger

A shell-and-tube exchanger is used for larger flows, which are very common in chemical process industries. The design of this exchanger is a shell with a bundle of tubes inside. The tubes are in parallel and a fluid flows around them in the shell. There are many different arrangements such as straight (Figure3) or u-tube (Figure 4). Each arrangement allows for a different type of flow such as co-current, counter-current and cross flow. The tube-side can have one or more passes to increase the energy exchange from the tube-side fluid. The shell-side may contain baffles, or walls, that channel the fluid flow and induce turbulence, and thus, increase energy exchange. Correlations can be developed to predict the increase in energy exchange. These correlations are out of the scope of this article.

Figure 3: Straight Tube Heat Exchanger

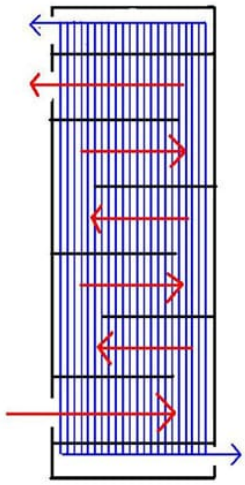
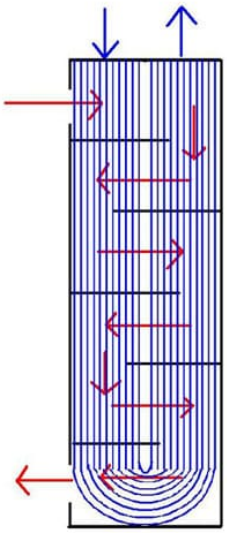


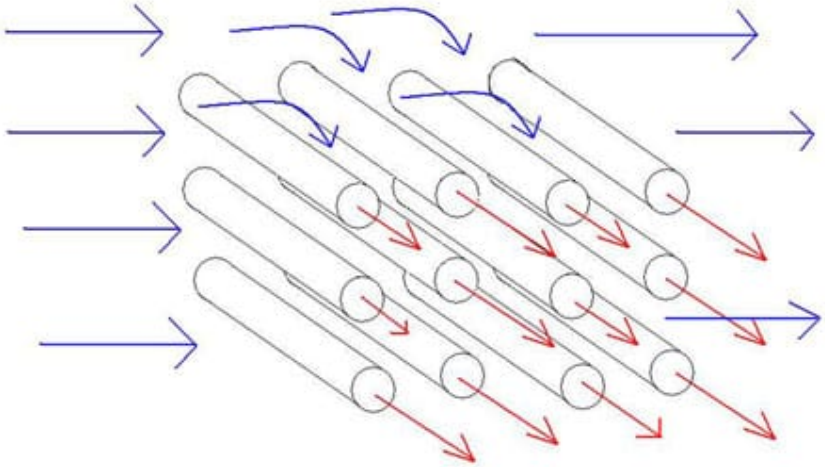
Figure 4: U-Tube Heat Exchanger



3. Cross-flow Exchanger

The most common application for a cross-flow heat exchanger is when a gas is being heated or cooled. This device consists of multiple tubes in parallel, usually containing a liquid, and the exterior gas flows across the tubes. In some cases the air is confined in separate flow chambers, such as fins, and in others it is open to flow freely (Figure 5).

Figure 5: Cross-Flow Heat Exchanger (free flow)



Dynamic Modeling of Heat Exchangers

Since heat exchangers are so widely used in industry, it necessary for a chemical engineer to be able to optimize and control the system and know how independent variables will affect the outputs from teh system. To do this, a dynamic model is developed and utilized.

A dynamic model of a heat exchanger may be used, for example, to predict how a change in the fluid flowrates or the addition of an insulating jacket will affect the outlet temperature of the product stream. The model uses ordinary differential equations (ODEs) to describe the process and, using a program like Microsoft EXCEL, gives plots of the variables vs. time for the entire process. There are many independent variables in a heat exchanger, which can cause modeling to be very complex since multiple ODEs are required to define all of the process variables.

Some of the independent variables in a heat exchanger system include:

Shell-side:

- fluid
- flowrate
- temperature
- number of baffles

Tube-side:

- fluid
- flowrate
- temperature
- number of passes

Flow configuration

- Co-current
- Counter-current
- Cross flow

Insulating jacket

The primary dependent variable of concern is the outlet temperature of the product stream (usually the tube-side fluid.) The outlet temperature, in a controlled system, is monitored by a sensor. The sensor then transmits a signal to an actuating device of one or more of the independent variables (usually shell-side flow controller) to perform some desired response.

Modeling with ODEs

The following section outlines the method for developing a dynamic model of a heat exchanger. The model is for a double-pipe heat exchanger that has the ability to flow in co-current or counter-current configurations and the option of an external insulating jacket. It assumes constant fluid properties and perfect heat transfer through the metal of the tubing. The outlet temperature of the tube-side fluid is monitored by a temperature sensor, and the flow rate of the shell-side fluid is controlled by a actuated flow-controlling device.

Energy Balance

An energy balance is first performed on the tube-side fluid.

$$\text{Rate of accumulation of thermal energy in tube-side fluid} = \text{rate of energy in} - \text{Rate of energy out} - \text{Heat transferred from sh}$$

The left-most term in the energy balance in Equation 6.6.1 is the amount of thermal energy that accumulates in the tube-side fluid and causes a change in its outlet temperature. The terms on the right side of the above energy balance describe the thermal energy of the fluid flowing in and the fluid flowing out and the amount of heat transfer from the shell side fluid. In the term for heat transfer from the shell-side, the temperatures are the temperatures of the outlet streams. The outlet temperatures will change according to whether you are running co-currently or counter-currently. The energy balance is written as:

$$mc_{p,t}\frac{dT_{t,out}}{dt} = \rho c_{p,t}F_{t,in}T_{t,in} - \rho c_{p,t}F_{t,out}T_{t,out} - \frac{kA_i}{\Delta z}\left(T_{t,out} - T_{s,out}\right) \tag{6.6.2}$$

where,

- m =mass of the fluid = $V\rho = \rho A_{crosssectional}\Delta z$
- c_p = constant pressure heat capacity of the fluid
- T =Temperature
- t =time

- k = conductive heat transfer coefficient
- A = surface area of tube that fluid contacts
- Δz = length of tube
- ρ = density of the fluid
- F = volumetric flowrate of the fluid

and subscripts denote

- t - tube-side fluid
- out - outlet
- in - inlet
- i - inside

An similar energy balance is next performed on the shell-side fluid.

$$\text{Rate of accumulation of thermal energy in tube-side fluid} = \text{rate of energy in} - \text{Rate of energy out} - \text{Heat transferred to tube}$$

The left-most term in the above energy balance is the amount of thermal energy that accumulates in the shell-side fluid and causes a change in its outlet temperature. The terms on the right side of the above energy balance describe the thermal energy of the fluid flowing in and the fluid flowing out, heat transfer to the tube-side fluid, and also the heat lost by convection to the surroundings. The energy is written as;

$$mc_{r,s}\frac{dT_{s,out}}{dt} - \rho c_{p,s}F_{s,sn}T_{s,in} - \rho c_{p,s}F_{s,cut}T_{s,out} - \frac{k \cdot A_o}{\Delta z}\left(T_{s,cut} - T_{t,out}\right) - h.A_s\left(T_s - T_\infty\right) \tag{6.6.4}$$

where,

- h = coefficient of convective heat transfer for air
- k - conductive heat transfer coefficient.

and subscripts denote

- s - shell-side fluid
- ∞ - air
- o - outside of tube
- o' - outside of shell

Considerations

There are considerations and simplifications you can make in order to solve the differential energy balances. The validity of these assumptions depends on how accurate of a model you require.

1. The heat capacity of the fluid may be temperature dependent. If this is the case, a polynomial equation could be written for the C_p values of each of the fluids. It would take the form-

$$C_p = a + bT + cT^2 + dT^3 \tag{6.6.5}$$

Values of a, b, c, and d are properties of the fluid and can be found in *Perry's Chemical Engineers' Handbook*.

It should also be noted that if the fluids in the process are gases, their fluid properties will be affected by changes in pressure and the use of a constant C_p value would be inappropriate.

2. The density of the fluid may be temperature dependent. This would be likely if the fluid were a vapor, as in the case of using steam as the shell-side fluid to heat the tube-side process fluid. If this is the case, a differential equation could be written for the ρ value of the fluid and would take the form-

$$\frac{d\rho}{dT} \propto C \tag{6.6.6}$$

where C is the coefficient of cubic expansion (relates kinetic energy to temperature).

3. Heat loss to the surroundings may be neglected. This would be the case if the heat exchanger is well insulated or if the shell-side fluid is about the same as ambient temperature. In this case-

$$hA\left(T_s - T_\infty\right) = 0 \tag{6.6.7}$$

4. The temperature sensor may have an inherent time delay. This means the temperature output from the sensor lags the actual temperature at the time of the reading. The temperature lag may be accounted for with the differential equation-

$$\frac{dT_{\text{outlet,sensor}}}{dt} = \frac{1}{\tau_{Ts}}\left(T_{\text{outlet,actual}} - T_{\text{outlet,sensor}}\right) \tag{6.6.8}$$

where τ_{Ts} = time constant for the temperature sensor. τ_{Ts} is a process parameter and usually ranges from 6 to 20 seconds depending on the properties of the sensor. This value would either be given in literature from the manufacturer or would have to be determined through experimentation.

5. The actuator system for the control valve may have a slow dynamic response. When the actuator system is based on a control valve, the response to the pressure change is slower than the flow through the valve. The flow rate through a control valve can be modeled with the differential equation-

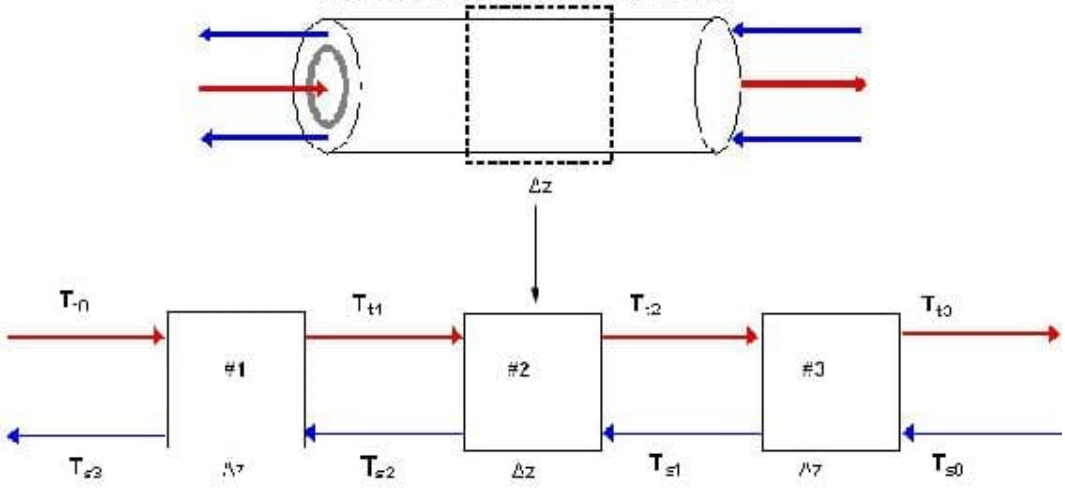
$$\frac{dQ_{\text{actual}}}{dt} = \frac{1}{\tau_v}\left(Q_{\text{setpoint}} - Q_{\text{actual}}\right) \tag{6.6.9}$$

where τ_v = time constant for the flow control valve. τ_v is a process parameter and usually ranges from 0.5 to 2 seconds depending on the properties of the sensor. This value would have to be determined through experimentation.

Using EXCEL to Solve ODEs

Since temperature of the heat exchanger varies across the metal as well as along the length of the pipe, a partial derivative for how temperature varies with length must be described. In order to solve this problem using excel, an approximation of length was used. Instead of taking the partial derivative across the length with respect to temperature, the pipe was divided into differential segments, Δz. Ideally Δz is an infinitesimally small cross section of the length of the heat exchanger. We will assume that through this differential segment, the temperature of the liquid that leaves the segment is the same as the temperature of the liquid within the segment. Since we are assuming the same temperature for the exiting streams as the inside of the segment, the choice of length for these Δz's helps dictate the accuracy of the solution. The combination of these differential units allows us to model a heat exchanger without the use of partial derivatives. The figure below shows an example of the simplification where the heat exchanger is split into three segments (Figure 6).

Figure 6: Heat Exchanger Simplification



To begin modeling, start by solving the appropriate energy balance (either Equation 6.6.2 or 6.6.4) for $\frac{dT}{dt}$ and make any simplifications necessary. Then use a method to solve ODEs in Excel, such as Euler's or Heun's. Go here: ExcelModelingODE for further explanation on how to set up an Excel spreadsheet using these methods. For each unit, Δz, and each fluid in that unit of the heat exhanger, you should have an approximation equation. Because ideally Δz represents an infinitesimal section, it follows from the first assumption that we can take the temperature in Δz as being the exit temperature of the hot and cold streams for the Δz's respectively.

The equation for Euler's Method takes the form:

$$T_{h1}\left(t_{i+h}\right) = T_{h1}\left(t_i\right) + \Delta t\left[\frac{dT}{dt}\left(t_i\right)\right] \tag{6.6.10}$$

The equation for Heun's Method takes the form:

$$T_{h_1}\left(t_{i+1h}\right) = T_{h_1}\left(t_i\right) + \Delta\left[\frac{1}{2}\frac{dT}{dt}\left(t_i T_i\right) + \frac{1}{2}\frac{dT}{dt}\left(t_{i+h} T_{i+1h_i}\right)\right] \tag{6.6.11}$$

The temperatures can then be plotted versus time in order to model how the system inputs affect heat exchange. The following examples walk you through a model of a simple heat exchanger in Excel.

Example 6.6.1: Euler's Method

As a process engineer in a chemical processing company, you were assigned to scale up a process for your company's latest and greatest chemical. Between one of the steps of the reaction, it is necessary to quickly and efficiently cool down the solution from 330 K before allowing it to enter the next step. From your previous knowledge, you know that a heat exchanger is necessary and the counter current heat exchanger is the most efficient to achieve a desired temperature. You decide to use cold water entering in at 250 K as your coolant since it is relatively inexpensive. In order to allow for the plant engineers in the future to have more control, create a model to determine when this process will go to steady state and graph the temperature profiles of the outlet streams. You can assume the heat exchanger is perfectly insulated and that there is perfect conduction through the metal from shell-side fluid to the tube-side fluid, and this conduction is described by the heat transfer coefficient, k, between the two fluids.

General Relevant Information:

- T_{air}=296.15 K (Temp. of the Air)
- r_i=0.1 m (Inner Diameter of Inner Tube)

- r_o=0.12 m (Outer Diameter of Inner Tube)
- r_o'=0.15 m (Outer Diameter of Outer Tube)
- Δz=1 m (Increment of length)
- k=450000 W/m2*K (heat transfer coefficient between shell and tube-side fluids)

Shell Side Fluid Information (Cold Water)

- C_{ps}=4185 J/kg*K (shell side fluid heat capacity)
- T_{0s}=250 K (input temp. of shell side fluid)
- A_s=0.02543 m (cross sectional area where shell side fluid is present)
- ρ_s=1000 kg/m^3 (density of shell side fluid)
- F_s=0.1 m3/s (volumetric flow rate of shell side fluid)

Tube Side Fluid Information (Solution)

- C_{pt}=1200 J/kg*K (tube side fluid heat capacity)
- T_{0t}=330 K (input temp. of tube side fluid)
- A_t=0.0314 m (cross sectional area where tube side fluid is present)
- ρ_t=1030 kg/m^3 (density of tube side fluid)
- F_t=0.2 m3/s (volumetric flow rate of tube side fluid)

Solution

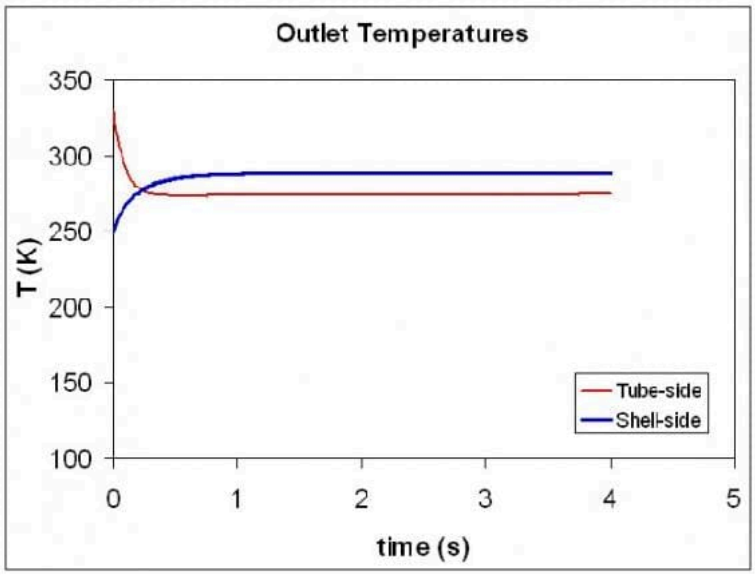
The general Algorithm can be followed to develop the Excel model.

- 1. Make a spreadsheet that has all relevant process information (ie: T inputs, ρ values, etc).
- 2. Perform an overall energy balance for the tube-side and shell-side fluids. Make simplifications and state assumptions. Solve for $\frac{dT}{dt}$
- 3. Divide the heat exchanger into smaller segments of length Δz
- 4. In your spreadsheet, set up a column for time. In the first row, start at t=0 and then increase time by Δh down the column.
- 5. Set up a column for the incoming tube-side temperature, T_{t0}. The value of incoming temperature is inputted by the user. Do the same for the shell-side temperature, T_{s0}.
- 6. Set up a column for the exiting tube-side temperature from each Δz unit. Do the same for the shell-side temperatures. For each of these temperature columns:

- a. In the first row, make an initial guess of the temperature. This is usually done by setting the temperature equal to the temperature of the stream coming into the heat exchanger.
- b. In subsequent rows, input the Euler's equation (Eq 10) by referencing the appropriate cells for the process information and take all temperatures from time, t_i, which are the temperatures in the row above.

- 7. Make a graph of variables of interest (Temperatures) vs. time.

The graph below illustrates how the exiting hot and cold temperature changes over a short period of time.



For a more detailed view of this process please see the attached spreadsheet:

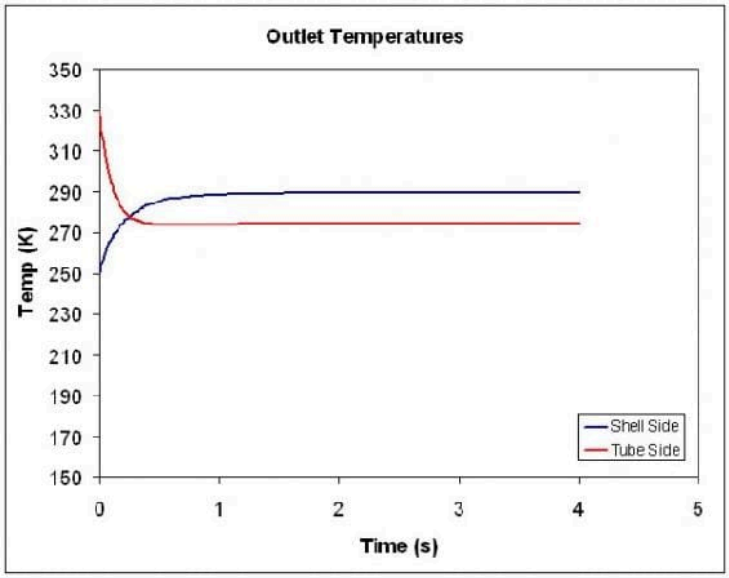
Example 6.6.2: Heun's Method

Again, you are working on the same process, but your supervisor requires you to have a more accurate description of the dynamics of the system. Being as bright as you are, you decide to utilize the 2nd order Runga Kutta method, known as the Heun's method.

Solution

Follow the Algorithm in Example 1, except in step 6b, use the equation for Heun's method, Eq 11.

The graph below illustrates how the exiting hot and cold temperature changes over a short period of time.



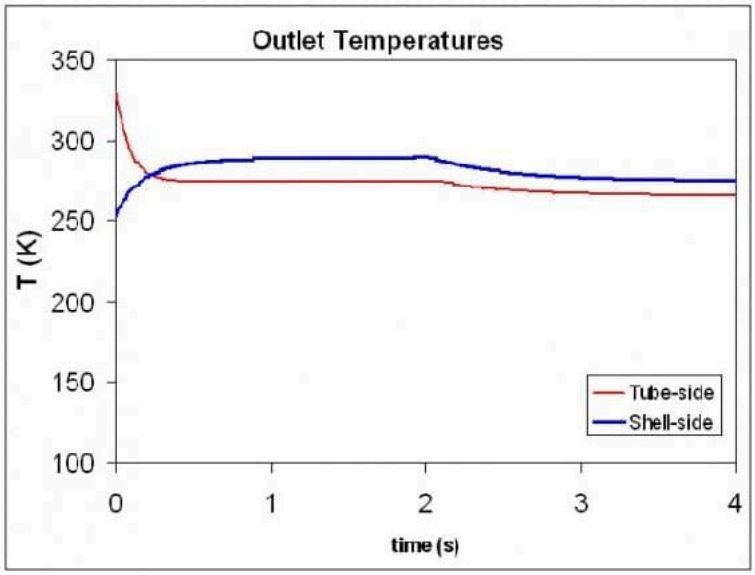
For a more detailed view of this process please see the attached spreadsheet:

Example 6.6.3

A month after you developed a model of the heat exchanger used in the process for your company's latest and greatest chemical, one of the plant engineers calls you and says that at t=2s into the process, the temperature of the incoming fluid changes abruptly from 330K to 300K. How do the temperatures of the outlet streams react to this change. Use the information and model from Example 1.

Solution

The Excel spreadsheet can be modified so that at time t=2s, the incoming temperature of the tube-side fluid is 300K. The graph below illustrates how the exiting hot and cold temperature changes over a short period of time.



For a more detailed view of this process please see the attached spreadsheet:

Tube-side Temp Change

References

- Riggs, James B.; Karim, M. Nazmul. Chemical and Bio-Process Control, Texas Tech University, Ferret Publishing.

Contributors and Attributions

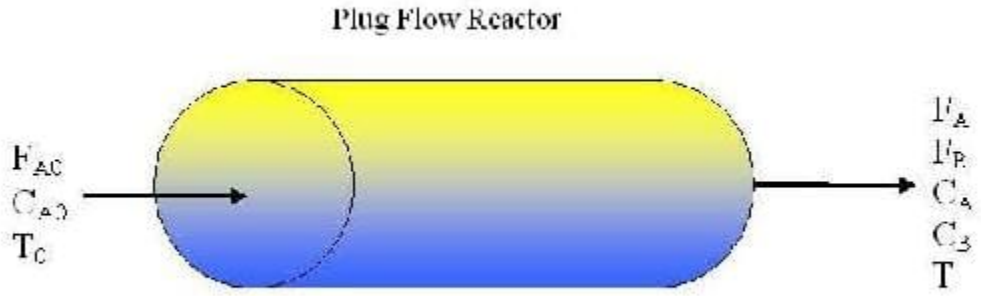
- Authors: Teresa Misiti, Carly Ehrenberger, Danesh Deonarain, Georgina Mang

This page titled 6.6: ODE and Excel model of a Heat Exchanger is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

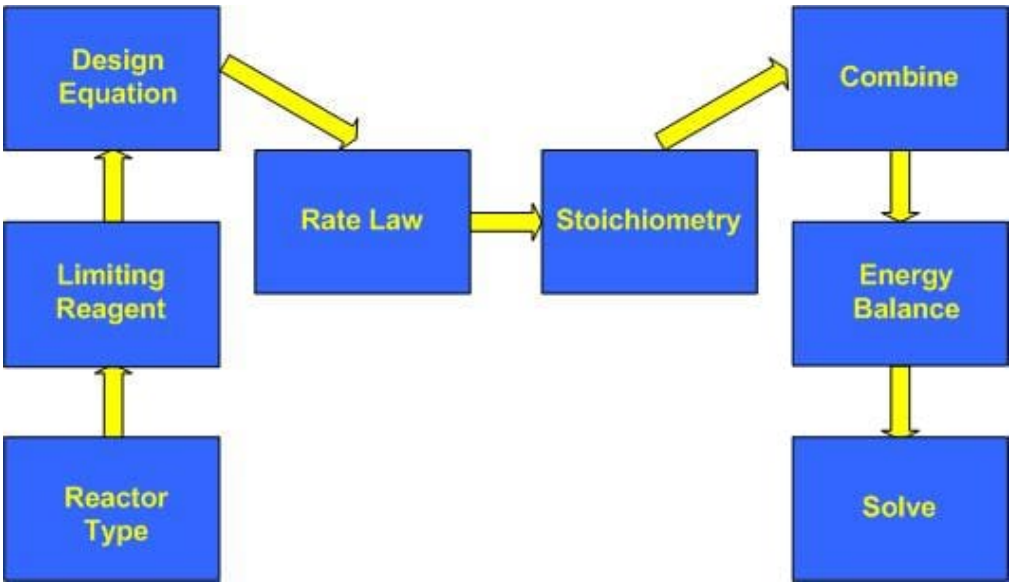
6.7: ODE and Excel Model of an Adiabatic PFR

Introduction

A **plug flow reactor (PFR)** is a tubular reactor used in chemical reactions. Reactants enter one end of the PFR, while products exit from the other end of the tube. A PFR is useful because of high volumetric conversion and good heat transfer. PFRs carry out power law reactions faster than CSTRs and generally require less volume. Excel modeling for an adiabatic plug flow reactor is useful for estimating conversion as a function of volume. By inputting the values of the constants and iteratively changing the volume, the conversion at a specific volume can be determined. This removes the need for complicated hand calculations and helps the user visualize the reaction and notice trends through the PFR. Alternatively, excel modeling can also determine the temperature inside the reactor as a function of volume. Euler's integration is used to estimate the conversion or temperature through the PFR. By ensuring that the volume intervals are small, the conversion or temperature values generated should be fairly accurate. Refer to the Euler's method section for more information on its implementation.



Basic Algorithm to Model Adiabatic PFR



The model presented will include the following assumptions:

1. Negligible pressure drop
2. Single elementary reaction
3. Power law kinetics
4. Constant fluid properties (i.e. heat capacity)

To describe the basic algorithm for an adiabatic PFR, a simple case will be considered where pure A enters the reactor:

$$a\,\mathrm{A} \rightleftharpoons a\,\mathrm{B}$$

1. Reactor Type

PFR

2. Limiting Reagent

The design and rate equations should be expressed in terms of the limiting reagent. Here, A is the limiting reagent.

3. Design Equation

The design equation for a PFR can be expressed in terms of several variables, including conversion, moles, and concentration. It is important to note that the design equation in terms of conversion may only be used if one reaction is occurring. Otherwise, it must be expressed in terms of moles or concentration. The equation to describe conversion as a function of volume, as derived from a simple mole balance, is shown below in Equation 6.7.1.

$$\frac{dX}{dV} = \frac{-r_A}{F_A O} \tag{6.7.1}$$

- X= conversion
- V= volume
- r_A= reaction rate of A
- F_{Ao}= initial moles of A

4. Rate Law

The rate law is independent of reactor type. It is expressed in terms of a rate constant and concentration. If the reaction is irreversible, the rate law is modeled according to Equation 6.7.2. If the reaction is reversible, the rate law models Equation 6.7.3 in which the concentration of B and the equilibrium constant must be accounted for.

$$-r_A = k c_A^a \tag{6.7.2}$$

$$-r_A = k\left(c_A^a - \frac{c_B^b}{K_c}\right) \tag{6.7.3}$$

where

- k= rate constant
- C_A^a = concentration of A with stoichiometric coefficient a
- C_B^b = concentration of B with stoichiometric coefficient b
- K_C= equilibrium constant

If the reaction is not isothermal, the rate constant and equilibrium constant should be written in terms of temperature as shown in Equation 6.7.4 and 6.7.5.

$$k = k_0 \exp\left(\frac{E}{R}\left(\frac{1}{T_0} - \frac{1}{T}\right)\right) \tag{6.7.4}$$

$$K_C = K_{C1} \exp\left(\frac{\Delta H_{RX}}{R}\left(\frac{1}{T_1} - \frac{1}{T}\right)\right) \tag{6.7.5}$$

where

- k_0= rate constant at T_0
- E= activation energy
- R= ideal gas constant
- T= temperature
- K_{C1} =equilbrium constant at T_1
- ΔH_{RX} = standard heat of reaction

5. Stoichiometry

Stoichiometry is used to express concentration of a species in terms of concentration of the limiting reactant. The stoichiometry equation depends on whether the reaction occurs in the liquid or gas phase. If it occurs in the gas phase, Equations 6.7.6 and 6.7.7 are used. For liquid phase, ε becomes zero and the correlation reduces to Equation 6.7.8. There is no pressure term included because it is assumed that pressure drop is negligible, and the initial pressure is equal to the final pressure.

$$C_A = C_{Ao}\frac{(1-X)}{(1+\epsilon X)}\frac{T_0}{T} \tag{6.7.6}$$

$$C_B = C_{Ao}\frac{\left(\theta_B + \frac{b}{a}X\right)}{(1+\epsilon X)}\frac{T_0}{T} \tag{6.7.7}$$

$$C_A = C_{AO}(1-X) \tag{6.7.8}$$

$$C_B = C_{Ao}\left(\theta_B + \frac{b}{a}X\right) \tag{6.7.9}$$

where

- $\varepsilon = y^*_{Ao}$ (stoichiometric coefficients of products - coefficients of reactants)
- y_{Ao} = initial mole fraction of A
- θ_i= ratio of initial moles (or concentration) of species i to initial moles (or concentration) of A

6. Combine

The design equation, rate law, and stoichiometric equations are combined. For example, assume the reaction is irreversible, liquid phase, and isothermal, with pure A entering the reactor and has the combined equation as shown below.

$$\frac{dX}{dV} = \frac{k * C^a_{A0}(1-X)^a}{F_{A0}} \tag{6.7.10}$$

7. Energy Balance

The energy balance can be used to relate temperature and conversion. It assumes heat flow and shaft work are not present.

$$T = \frac{X\left(-\Delta H_{RX}\right) + \Sigma\Theta_i C_{pi}T_o + X\Delta C_p T_r}{\Sigma\Theta_i C_{pi} + X\Delta C_p} \tag{6.7.11}$$

with

- C_{pi}= heat capacity of species i
- ΔCp= Cp products- Cp reactants

If \(ΔC_p= 0\), the energy balances reduces to the following:

$$T = \frac{X\left(-\Delta H_{RX}\right) + \Sigma\Theta_i C_{pi}T_o}{\Sigma\Theta_i C_{pi}} \tag{6.7.12}$$

8. Solve

Solve the system of equations in Excel. The ODE is approximated by Euler's method. See the attached spreadsheet for the ODE/

Excel modeling of the adiabatic PFR.

Example 6.7.1

A plant at the ETF (Equipment Testing Facility) has been acquired in a hostile takeover by Brown Industries, a company for which you are a process engineer. The Engineering Manager, Phil Savage, has assigned you to work on a PFR located within this new plant. The particular PFR in question is used to conduct the following reaction:

$$A \longrightarrow B$$

He wants you to provide him with two pieces of information. First, he wishes to know how far down the reactor must the reaction go to achieve 80% conversion. Secondly, he wants to know the temperature at this point.

You decide that first it would be best to ascertain all the thermodynamic and kinetic properties that are inherent to this reaction. Instead of looking them up in Perry's or your old ChemE books, however, you decide to ask the Lead Plant Engineer, Pablo LaValle. Pablo tells you the following:

"Listen wisely to what I say, young grasshopper. This reaction is irreversible, liquid-phase, and elementary. We will feed the reactants into the tube reactor of diameter 50cm at 273 K. We will begin with an initial concentration of A at 1.6 mol/L and an initial molar flow rate of A at 2 mol/min. The heat of reaction at 300 K is -20 kJ/mol A. The rate constant at 300K is 0.2 mol / L * min. The activation energy at 300K is 15.2 kJ/mol. The Cp values for the products and reactants are as follows.

- Cp A = 200 J / mol K
- Cp B = 100 J / mol K

"Go forth and use the Excel sheet that Group B has created and your job will be very easy."

Solution

The model for this example can be found here.

Pablo gave all the data needed to complete the model. After inputing the data, the final page of the spreadsheet was used to determine the reactor conditions Savage desired.

X = 0.8 at a reactor length of 33.4 meters. The temperature at this point of the reactor is 388.4K

Example 6.7.2

You have just completed your initiation in the ultra-super secret elite society known as the League of Chemical Engineers which is headquartered high in the Himalayans. They have an adiabatic PFR (diameter = 10cm) in their possession which they use to produce two gases, known as B and C, from a gas-phase reactant known as A. The reversible gas reaction is shown below:

$$A <= \longrightarrow B + C$$

Your sensei wishes to know what the maximum conversion possible is and also what length of reactor would be needed to reach a conversion of 15%. First, however, he orders you to climb K2 where the reaction data is stored by the society. After a perilous 5 day journey you discover the following information:

- Cp $_A$ = 40 kJ/mol K
- Cp $_B$ = 25 kJ/mol K
- Cp $_C$ = 15 kJ/mol K

- C_{AO} = 0.271 mol/L
- F_{AO} = 5 mol/min
- Initial Temp = 450K
- Heat of reaction at 300K = -20 kJ/mol A

- Activation energy = 31.4 kJ/mol
- K_{eq} at 300K = 0.01 mol/L
- k_forward at 300K = 0.133 mol /(L*min)

Solution

After climbing back down K2, and finding a computer, the data obtained can be input into the adiabatic model. The solution can be found here.

Be sure to set the values for both gas phase reaction and reversible reaction to 1.

Using the data table found on the last page of the spreadsheet, the desired information can be extracted.

The maximum conversion that can be achieved is X = 0.187

The reactor requires 64.8 meters to achieve a conversion of 15%.

References

Fogler, H. Scott. *Elements of Chemical Reaction Engineering*. Pearson Education, Inc., 2006.

Contributors and Attributions

- Brian McQuillan, Crystal Miranda, Brandon Quigley, and John Zhang

This page titled 6.7: ODE and Excel Model of an Adiabatic PFR is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

6.8: Modeling and PID Controller Example - Cruise Control for an Electric Vehicle

Introduction

Controls principles developed in this course can be applied to non-chemical engineering systems such as automobiles. Some companies, such as NAVTEQ, are developing adaptive cruise control products that use information about the upcoming terrain to shift gears in a more intelligent manner which improves speed regulation and fuel economy. This case study will examine the basics of develop a speed controller for an electric vehicle.

An electric vehicle was chosen for the following reasons:

- Electric vehicles are interesting from an engineering perspective and may become a reality for consumers in the future
- Torque produced by an electric motor is instantaneous (for all practical purposes). Thus actuator lag can be ignored, simplifying the development of said controller.
- Some electric vehicles feature motors directly integrated into the hub of the drive wheel(s). This eliminates the need for a transmission and simplifies vehicle dynamics models.

Forces

As shown in the free body diagram below, there are six forces acting on the vehicle:

1. Rolling Resistance
2. Aerodynamic Drag
3. Aerodynamic Lift
4. Gravity
5. Normal
6. Motor

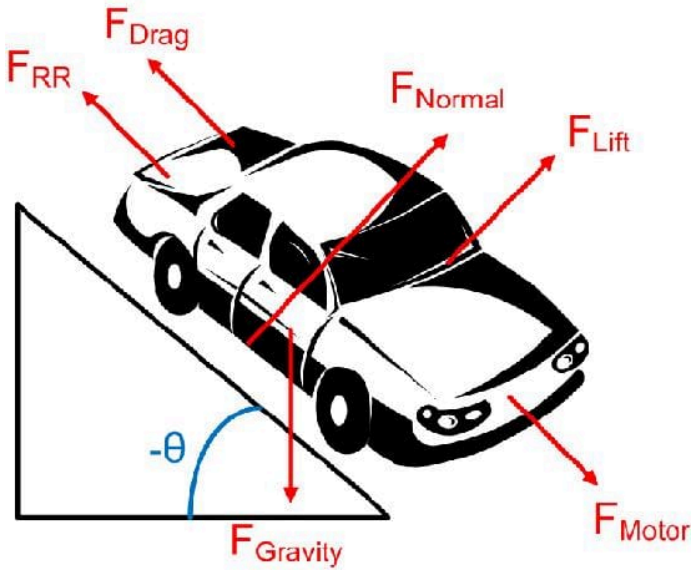


Figure 1: A free body diagram of the forces acting on the vehicle. $-\theta$ is used to denote the grade of the road such that a positive value of θ corresponds to the vehicle traveling uphill.

Rolling Resistance

Rolling resistance is due the tires deforming when contacting the surface of a road and varies depending on the surface being driven on. It can be model using the following equation:

$$F_{RR} = C_{rr1}v + C_{rr2}F_N$$

The two rolling resistance constants can be determined experimentally and may be provided by the tire manufacture. F_N is the normal force.

Aerodynamic Drag

Aerodynamic drag is caused by the momentum loss of air particles as they flow over the hood of the vehicle. The aerodynamic drag of a vehicle can be modeled using the following equation:

- ρ is the density of air. At 20°C and 101kPa, the density of air is 1.2041 kg/m^3.
- CdA is the coefficient of drag for the vehicle times the reference area. Typical values for automobiles are listed here.
- v is the velocity of the vehicle.

Aerodynamic Lift

Aerodynamic lift is caused by pressure difference between the roof and underside of the vehicle. Lift can be modeled using the following equation:

- ρ is the density of air. At 20°C and 101kPa, the density of air is 1.2041 kg/m^3.
- ClA is the coefficient of lift for the vehicle times the reference area.
- v is the velocity of the vehicle.

Gravity

In the diagram above, there is a component of gravity both in the dimension normal to the road and in the dimension the vehicle is traveling. Using simple trigonometry, the component in the dimension of travel can be calculated as follows:

$$F_{G,travel} = mg\sin(-\theta)$$

- m is the mass of the vehicle.
- g is the acceleration due to gravity.

Normal Force

The normal force is the force excerted by the road on the vehicle's tires. Because the vehicle is not moving up or down (relative to the road), the magnitude of the normal forces equals the magnitude of the force due to gravity in the direction normal to the road.

- m is the mass of the vehicle.
- g is the acceleration due to gravity.

Motor

The torque produced by an electric motor is roughly proportional to the current flowing through the stater of the motor. In this case study, the current applied to the motor will be controlled to regulate speed. Applying a negative current will cause the vehicle to regeneratively brake.

$$\tau = k_{\text{motor}} I$$

$$F_M = \frac{\tau}{r} = \frac{k_{\text{motor}}\, I}{r}$$

- τ is the torque produced by the motor.
- I is the current flowing through the motor.
- r is the radius of the tire.
- k_{motor} is a constant.

Newton's Second Law

Using Newton's Second Law, a differential for the vehicle's speed can be obtained.

$$\begin{aligned} ma &= \sum F \\ &= F_M - F_{drag} - F_{RR} + F_{G,\text{travel}} \end{aligned}$$

Substituting in the expressions for various forces detailed above yields the following:

Further substituting the expression for normal forces yields the following:

Substituting $\dfrac{dv}{dt}$ for a results in the following:

Grouping like turns results in the following:

In order to simply the remaining analysis, several constants are defined as follows:

$$\beta = rC_{rr1}$$

$$\gamma = rC_{rr2}mg\cos(-\theta) - rmg\sin(-\theta)$$

Substituting these into the differential equation results in the following expression:

$$\frac{dv}{dt} = \frac{\tau - \alpha v^2 - \beta v - \gamma(\theta)}{mr}$$

It is important that θ (and thus also γ) is a function of vehicle position.

PID Controller

One way to regulate vehicle speed is to control the torque generated by the electric motor. If the motor torque is greater than the resistive torque acting on the vehicle (a summation of aerodynamic drag, rolling resistance, etc.) the vehicle will accelerate. If the

motor torque is less than the resistive torque, the vehicle will slow down.

Expression for Phase Current

For an electric motor, the phase current flowing through the motor is proportional to the torque produced. Thus one strategy for controlling the vehicles speed is to controller the motor phase current.

Using a PID controller architecture, the expression for motor current is the following:

$$I = K_c\left(v_{set} - v\right) + \frac{1}{\tau_I}\int (v - v_{set})dt + \tau_D\frac{d\left(v_{set} - v\right)}{dt} + C_{offset}$$

Differential Equation for Velocity

Substituting this expression into the differential equation for vehicle position results in the following:

$$\frac{dv}{dt} = \frac{k_{motor}[K_c(v_{set} - v) + \frac{1}{\tau_I}\int (v_{set} - v)dt + \tau_D\frac{d(v_{set} - v)}{dt} + C_{offset}] - \alpha v^2 - \beta v - \gamma(\theta)}{mr}$$

Defining another variable x_1 allows for the removal of the integral from the expression.

$$\frac{dx_1}{dt} = v_{set} - v$$

$$\frac{dv}{dt} = \frac{k_{\text{motor}}\left[K_c\left(v_{\text{isct}} - v\right) + \frac{x_1}{\tau_l} + \tau_D\frac{d(v_{o=t} - v)}{dt} + C_{offset}\right] - \alpha v^2 - \beta v - \gamma(\theta).}{mr}$$

If all changes in v_{set} are gradual then $\frac{dv_{set}}{dt} \approx 0$. Applying this simplification results in the following expression:

$$\frac{dv}{dt} = \frac{k_{\text{motor}}\left[K_c\left(v_{\text{set}} - v\right) + \frac{x_1}{\tau_1} - \tau_D\frac{dv}{dt} + C_{\text{offset}}\right] - \alpha v^2 - \beta v - \gamma(\theta)}{mr}$$

Using a little algebra, an expression for $\frac{dv}{dt}$ can be obtained as follows:

$$\frac{dv}{dt} = \frac{k_{motor}K_c(v_{set} \quad v) \mid \frac{k_{motor}}{\tau_I}x_1 \mid k_{motor}C_{offset} \quad \alpha v^2 \quad \beta v \quad \gamma(\theta)}{k_{motor}t_D + mr}$$

$$\frac{dv}{dt} = \frac{k_{\text{motor}}K_c\left(v_{\text{set}}^{k_{\text{motor}}}x_1|k_{\text{motor}}C_{\text{cffses}}^2\right)^{\alpha}(\theta)}{k_{\text{motor}}t_L + mr}$$

Find Fixed Point

The fixed point can be obtained by setting the derivatives to zero and solving the system of equation.

System of Equations:

$$\frac{dv}{dt} = \frac{k_{\text{meclor}}\,K_c\left(v_{\text{sel}} - v\right) + \frac{k_{\text{motor}}}{\tau_I}x_1 + k_{\text{melur}}\,C_{v\int js\leq t} - \alpha v^2 - \beta v - \gamma(\theta)}{k_{\text{motor}}\,t_D + mr} = 0$$

$$\frac{x_1}{dt} = v_{set} - v = 0$$

Solution

$$x_1 = \frac{-\tau_I \left(k_{\text{motor}} C_{\text{offset}} - \alpha v_{\text{set}}^2 - \beta v_{\text{set}} - \gamma(\theta)]\right)}{k_{\text{motor}}}$$

$$v = v_{set}$$

As expected, a fixed point exists when the set velocity equals the actual velocity of the vehicle.

Linearize System of ODEs

Before the stability of the system can be (easily) examined, the system must be linearized around a fixed point.

Note: To simply the matrix expressions in section the following notation will be used:

$$y_i' = \frac{dy_i}{dt}$$

Overall a linearized system of ODEs has the following form:

$$\begin{bmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{bmatrix} = \mathbf{J} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}$$

The first step of linearizing any system of ODEs to calculate the Jacobian. For this particular system, the Jacobian can be calculated as follows:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial v'}{\partial v} & \frac{\partial v'}{\partial x_1} \\ \frac{\partial x_1'}{\partial v} & \frac{\partial x_1'}{\partial x_1} \end{bmatrix} = \begin{bmatrix} \frac{-k_{motor} K_c - mr[2\alpha v + \beta]}{mr + k_{motor}\tau_D} & \frac{k_{\text{motor}}}{t_I(mr + k_{\text{motor}}\,\tau_D)} \\ -1 & 0 \end{bmatrix}$$

The Jacobian is then evaluated at the fixed point:

$$\mathbf{J} = \begin{bmatrix} \frac{-k_{\text{motor}}\, K_c - mr[2\alpha v_{\text{set}} + \beta]}{mr + k_{\text{motor}}\,\tau_D} & \frac{k_{\text{motor}}}{t_I(mr + k_{\text{motor}}\,\tau_D)} \\ -1 & 0 \end{bmatrix}$$

The next step if to calculate the vector of constants. For this particular system, said vector can be calculated as follows:

$$\begin{bmatrix} k_v \\ k_{x_1} \end{bmatrix} = -\mathbf{J} \begin{bmatrix} v \\ x_1 \end{bmatrix}_{\text{fixed point}} = -\begin{bmatrix} \frac{k_{motor} K_c - 2\alpha v_{set} - \beta}{mr + k_{motor}\tau_D} & \frac{k_{motor}}{t_I(mr + k_{motor}\tau_D)} \\ -1 & 0 \end{bmatrix} \begin{bmatrix} v_{set} \\ \frac{-\tau_I(k_{motor} C_{offset} - \alpha v_{set}^2 - \beta v_{set} - \gamma(\theta))}{k_{motor}} \end{bmatrix}$$

$$\begin{bmatrix} k_v \\ k_{x_1} \end{bmatrix} = \begin{bmatrix} \frac{-(-k_{\text{motor}} K_c - 2\alpha v_{set} - \beta])v_{set}}{(mr + k_{\text{motor}}\tau_D)} + \frac{(k_{\text{motor}} C_{\text{offset}} - \alpha v_{\text{set}}^2 - \beta v_{\text{set}} - \gamma(\theta)])}{(mr + k_{\text{motor}}\tau_D)} \\ v_{\text{set}} \end{bmatrix}$$

Combining the Jacobian and vector of constants results in the following linearized system:

$$\begin{bmatrix} v' \\ x_1' \end{bmatrix} = \mathbf{J} \begin{bmatrix} v \\ x_1 \end{bmatrix} + \begin{bmatrix} k_v \\ k_{x_1} \end{bmatrix}$$

$$\begin{bmatrix} v' \\ x_1' \end{bmatrix} = \begin{bmatrix} \frac{-k_{motor} K_c - 2\alpha v_{set} - \beta}{mr + k_{motor}\tau_D} & \frac{k_{motor}}{t_I(mr + k_{motor}\tau_D)} \\ 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ x_1 \end{bmatrix} + \begin{bmatrix} \left(\frac{k_{motor} K_c - 2\alpha v_{set} - \beta]}{(mr + k_{motor}\tau_D)}\right) v_{set} + \frac{(k_{motor} C_{offset} - \alpha v_{set}^2 - \beta v_{set} - \gamma(\theta))}{(mr + k_{motor}\tau_D)} \\ v_{set} \end{bmatrix}$$

Stability Analysis

To assess the stability of the controller, the eigenvalues of the Jacobian in the linearized systems of ODEs can be examined. In

general, an eigenvalue (*lambda*) is the solution to the following equation:

$$|\mathbf{J} - \lambda \mathbf{I}| = 0$$

Using a computer to solve said equation, the eigenvalues of this particular system can be found to be the following:

$$\lambda = \frac{k_{motor}K_c\tau_I + \beta\tau_I - 2\alpha v_{set}\tau_I \pm \sqrt{\tau_I(k_{motor}^2K_c^2\tau_I + 2k_{motor}K_c\tau_I\beta + 4k_{motor}K_c\tau_I\alpha v_{set} - \beta^2\tau_I - 4\beta\tau_I\alpha v_{set} + 4\alpha^2v_{set}^2\tau_I - 4k_{motor}mr - 4k_{motor}^2t_D)}}{2\tau_I(\alpha v_{set} - k_{motor}r)}$$

For the system to be stable, the real component of all eigenvalues must be non-positive. The following inequality must be true for a stable controller:

$$k_{motor}K_c\tau_I + \beta\tau_I - 2\alpha v_{set}\tau_I < Real(\tau_I(k_{motor}^2K_c^2\tau_I + 2k_{motor}K_c\tau_I\beta + 4k_{motor}K_c\tau_I\alpha v_{set} + \beta^2\tau_I - 4\beta\tau_I\alpha v_{set} + 4\alpha^2v_{set}^2\tau_I - 4k_{motor}mr - 4k_{motor}^2t_D))$$

For the system to not oscillate, the imaginary component of all eigenvalues must be zero. The following inequality must be true for a non-oscillating controller:

$$\tau_I(k_{motor}^2K_c^2\tau_I + 2k_{motor}K_c\tau_I\beta + 4k_{motor}K_c\tau_I\alpha v_{set} + \beta^2\tau_I + 4\beta\tau_I\alpha v_{set} + 4\alpha^2v_{set}^2\tau_I - 4k_{motor}mr - 4k_{motor}^2t_D)$$

Interestingly, neither of these criteria depend on the grade of the road (θ). However, during the analysis, it was assumed that θ is constant. For most roads, this is not the case; θ is actually a function of vehicle position. In order to add this additional level of detail, the original system of ODEs needs to be revised:

$$\frac{dv}{dt} = \frac{k_{\text{motor}}K_c\left(v_{\text{set}}, v\right)|k_{\text{motor}}x_1|\, k_{\text{motor}}C_{\text{offset}}}{k_{\text{motor}}l_D + \pi ur}$$

$$\frac{dx_1}{dt} = v_{set} - v$$

$$\frac{ds}{dt} = v$$

$$\theta = f(s)$$

Unfortunately for any normal road, the grade is not a simple (or even explicate) function of position (s). This prevents an in depth analytical analysis of stability. However for a very smooth road with very gradual changes in grade, the stability of the controller should be unaffected.

Example Electric Vehicle

Simulating the system also for other properties of the controller to be examined. This section presents the results from simulating a specified fictitious electric vehicle.

Parameters

For the fictitious electric vehicle simulated in this analysis, the following parameters were used. These parameters are roughly based on the parameters one would expect to see in a typical electric automobile.

$$m = 1400\ kg$$

$$CdA = 0.58\ m^2$$

$$ClA = 0\ m^2$$

$$C_{rr1} = 2.75\ \frac{Ns}{m}0.018$$

$$C_{rr2} = 0.018$$

$$k_{motor} = 1.2 \frac{Nm}{A_{rms}}$$

Root Locus Plots

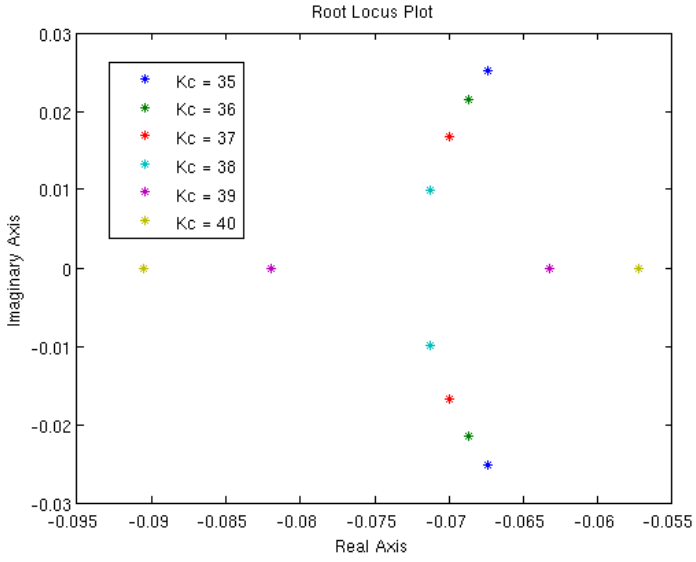
Using the parameters above, Root Locus Plots of the system were constructed to numerically explore the stability. The following controller constants were used when constructing the plots:

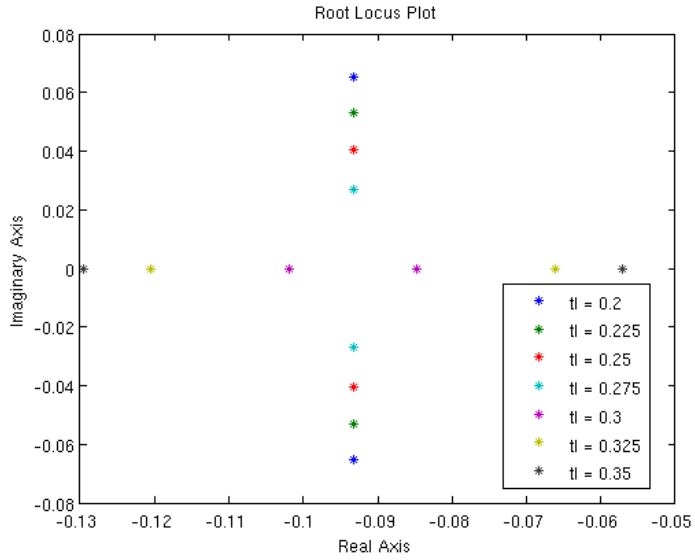
$$K_c = 55$$

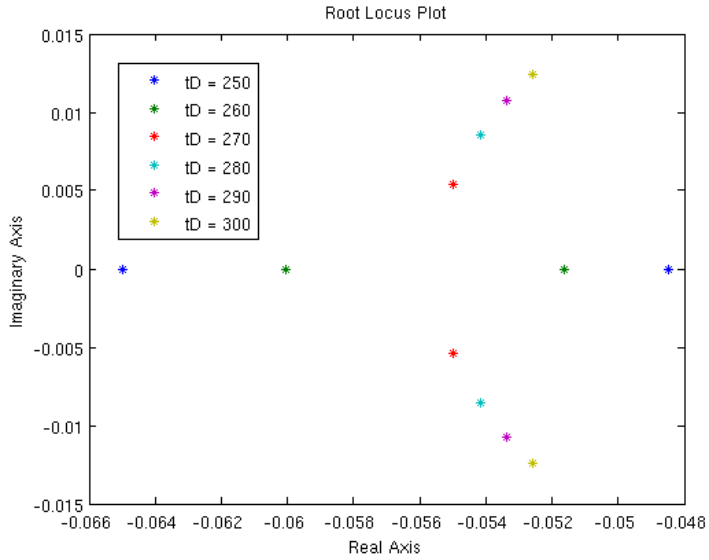
$$\tau_I = 0.5$$

$$\tau_D = 1$$

$$C_{offset} = 0$$







These Root Locus plots show that the controller with said constants is both stable and does not oscillate.

Phase Portrait

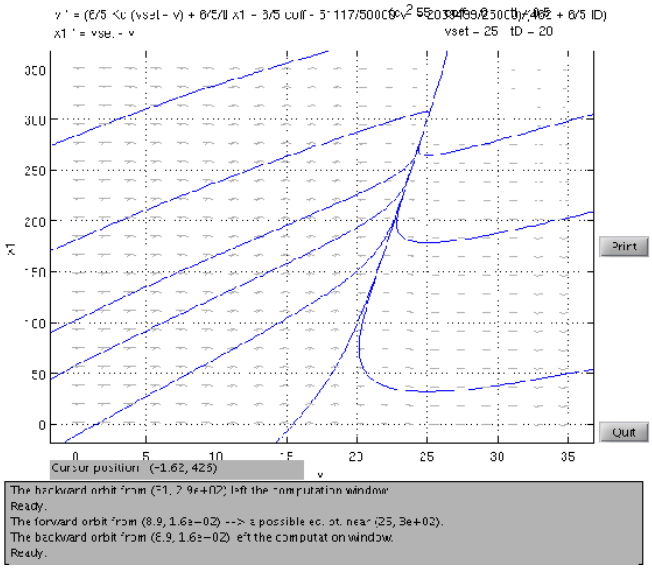
A phase portrait of the system with the following parameters was also constructed:

$$K_c = 55$$

$$\tau_I = 0.5$$

$$\tau_D = 20$$

$$C_{offset} = 0$$



The phase portrait shows that the system is both stable and does not oscillate, as predicted by the Root Locus plots.

Driving Simulation on Level Terrain

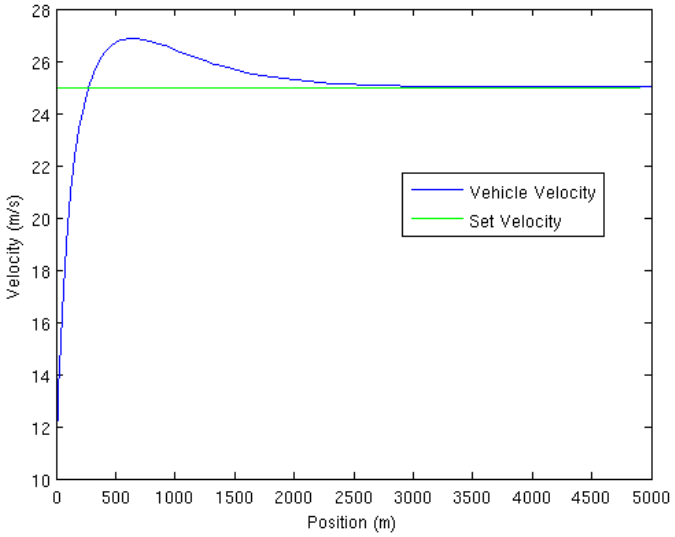
The vehicle was simulated starting at 10 m/s and accelerating to 25 m/s via cruise control on level terrain ($\theta = 0$). For this simulation, the following constants were used:

$$K_c = 55$$

$$\tau_I = 0.5$$

$$\tau_D = 20$$

$$C_{offset} = 146 \approx \frac{\alpha v_{set}^2 + \beta v_{set} + \gamma(\theta = 0)}{k_{motor}}$$



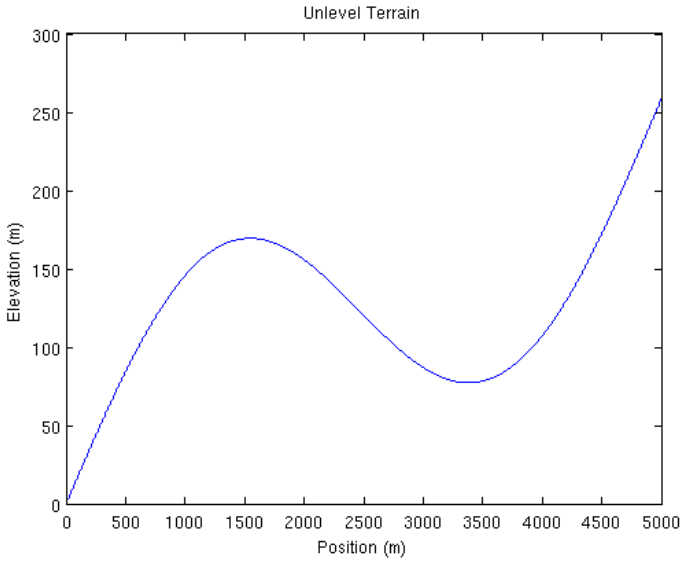
Interestingly this graph shows oscillation, despite the Root Locus plots and phase diagrams. It is important to remember, however, that the Root Locus plot and stability methods involve linearizing the system. It is possible the linearized system is not a good approximation of the system.

It is also important to remember that this particular example involves a large set point change, which can induce oscillations in certain systems.

Driving Simulation on Unlevel Terrain

In order to explore how the controller behaves on a road with a non-zero grade, a route with hills was constructed from the following equation, where *h* is elevation and *s* is position.:

$$h = 100 \sin\left(\frac{s}{250\pi}\right) + \frac{s}{2000}$$



The vehicle was simulated driving said road starting with the following initial conditions and controller constants:

Initial Conditions

$$s(t=0) = 0$$

$$x_1(t=0) = 0$$

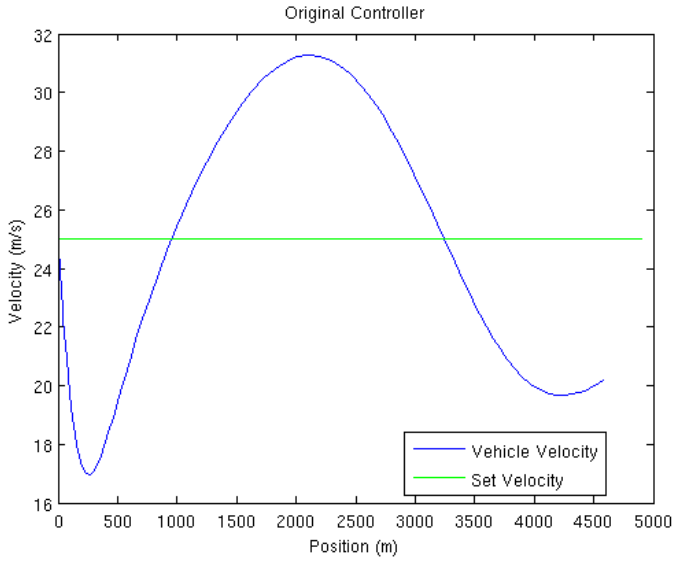
Controller Constants

$$k_c = 55$$

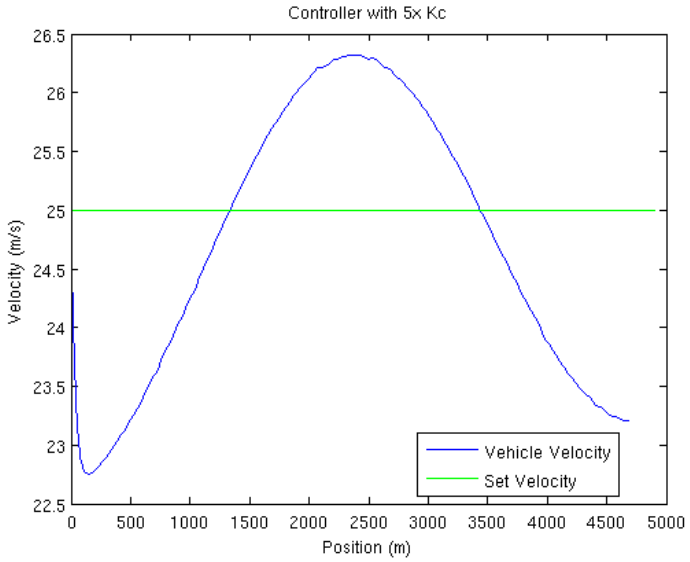
$$\tau_I = 0.5$$

$$\tau_D = 20$$

$$C_{offset} = 146$$



The current controller tuning is inadequate for this road. There are vary large variations in velocity. In order to reduce these variation, the propartional gain K_c was increased by a factor of 5. Below are the results:



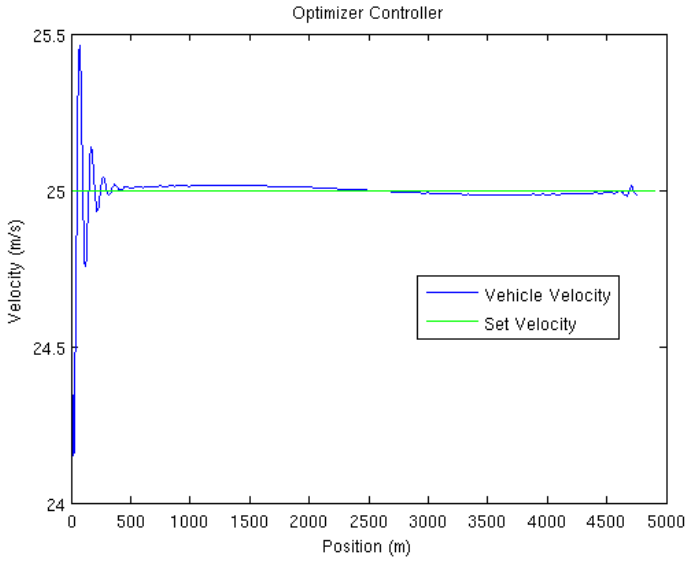
Using the Optimization Toolbox in MATLAB, the controller was optimized by minimizing the sum of the errors (difference between vehicle velocity and set velocity) on this particular segment of road. Below are the optimized controller constants and a plot of the velocity profile:

$$k_c = 232.58$$

$$\tau_I = 0.001$$

$$\tau_D = 0$$

$$C_{offset} = 114.62$$



This example shows the power of optimization and model predictive control.

Summary

This example demonstrates the following:

1. Modeling a simple, non-chemical engineering system.
2. Develop a PID controller for said physical dynamical system.
3. Manipulate said system to develop an system of differential equations.
4. Find fixed point(s) for said system.
5. Linearize said system of differential equations.
6. Find the eigenvalues of said linearized system.
7. Construct Root Locus plots for said linearized system.
8. Construct a phase portrait for said dynamical system.
9. Simulate said dynamical system under various conditions.
10. Demonstrate the idea of model predictive control by optimize said controller under a specific scenario.

This page titled 6.8: Modeling and PID Controller Example - Cruise Control for an Electric Vehicle is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Alex Dowling via source content that was edited to the style and standards of the LibreTexts platform.

6.9: Blood Glucose Control in Diabetic Patients

Background of why insulin control is important for diabetic patients

Diabetes mellitus is a disease of the endocrine system where the body cannot control blood glucose levels. There are two general classifications of diabetes:

Type I (also known as juvenile diabetes)

- Genetic predisposition and/or an autoimmune attack destroys T-cells of pancreas
- Body cannot produce insulin to regulate blood glucose

Type II

- Most common form of diabetes and has reached epidemic status in the United States
- Usually caused by lifestyle
- Obesity reduces body's responsiveness to insulin

Treatment for both types of diabetes may include exercise, dieting, oral medications, or insulin injections. Most insulin dependent diabetics follow a management plan that requires frequent testing of blood glucose levels and then injection of a prescribed dose of insulin based on the blood glucose level. However, the downside of this treatment method is that there is no predictive control. If blood glucose levels are falling and insulin is administered, a hypoglycemic episode may occur. Recent biomedical advancements have resulted in continuous blood glucose monitoring devices as well as insulin pumps. Continuous monitoring allows for finer blood glucose control and can help predict fluctuations in the blood glucose level. Insulin pumps replace the need to administer insulin injections by automatically injecting a prescribed dose, however it requires blood glucose level input from the patient. In the future, insulin pumps and continuous blood glucose monitors may be integrated forming a closed loop control system which will can replace the body's own faulty control system.

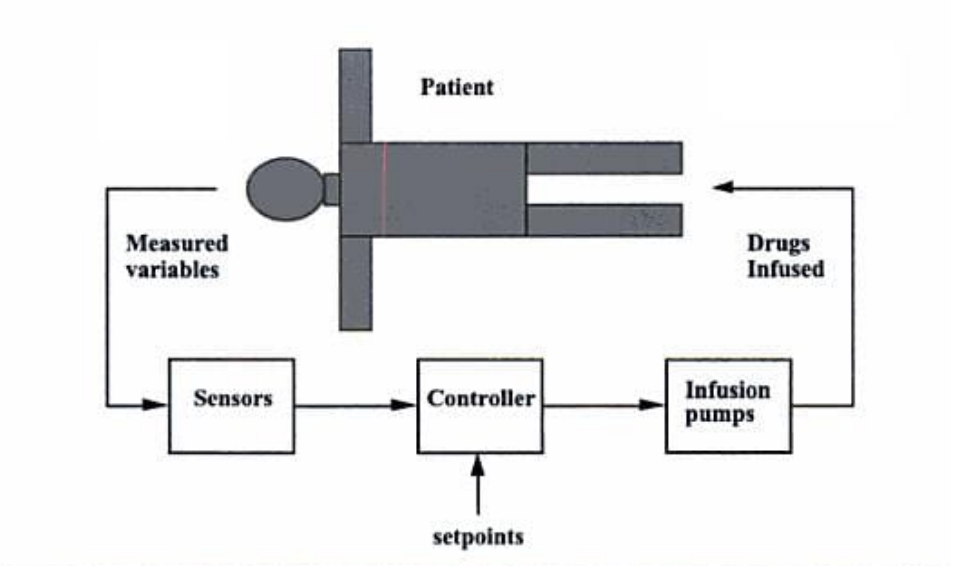


Figure 1: Control Schematic of Insulin Infusion

Mathematical model for a closed loop insulin delivery system

The following set of differential equations are known as the Bergman "minimal model":

$$\frac{dG}{dt} = -p_1 G - X\,(G - G_b) + \frac{G_{meal}}{V_1}$$

$$\frac{dX}{dt} = -p_2 X + p_3 I$$

$$\frac{dI}{dt} = -n\left(I + I_b\right) + \frac{U}{V_1}$$

where:

- G = deviation variable for blood glucose concentration
- X = deviation variable for insulin concentration in a "remote" compartment
- I = deviation variable for blood insulin concentration
- G_{meal} = a meal disturbance input in glucose
- U = the manipulation insulin infusion rate
- G_b = steady state value of blood glucose concentration
- I_b = steady state value of blood insulin concentration

Blood parameters include p_1, p_2, p_3, n, V_1(blood volume). These are specific to the blood specimen and must be predetermined.

A linear state space model can be used to express the Bergman equations seen above. The general form for a state space model can be seen below:

$$\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nm} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$$

and

$$\begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{r1} & \cdots & c_{rn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} d_{11} & \cdots & d_{1m} \\ \vdots & \ddots & \vdots \\ d_{r1} & \cdots & d_{rm} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$$

In general:

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

and

$$y = Cx + Du$$

where:

- x = states
- u = inputs
- y = outputs

$$A = \begin{bmatrix} -p_1 & -G_b & 0 \\ 0 & -P_2 & P_3 \\ 0 & 0 & -n \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & \frac{1}{V_1} \\ 0 & 0 \\ \frac{1}{V_1} & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Using this general formula, we can deconstruct the Bergman equations as a linear state space model. The first input is the insulin infusion and the second input represents the meal glucose disturbance.

First Input:

$$\begin{bmatrix} \dot{G} \\ \dot{X} \\ \dot{I} \end{bmatrix} = \begin{bmatrix} -p_1 & -G_b & 0 \\ 0 & -P_2 & P_3 \\ 0 & 0 & -n \end{bmatrix} \begin{bmatrix} G \\ X \\ I \end{bmatrix} + \begin{bmatrix} 0 & \frac{1}{V_1} \\ 0 & 0 \\ \frac{1}{V_1} & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

where

- $\dot{G}$ = differential blood glucose concentration
- $\dot{X}$ = differential insulin concentration in a "remote" compartment
- $\dot{I}$ = differential blood insulin concentration

Second Input:

$$y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} G + \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

where

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} U - U_b \\ G_{meal} - 0 \end{bmatrix}$$

Example Set of Parameters

- Gb = 4.5 mmol/liter
- Ib = 4.5 mU/liter
- V1 = 12 liters
- p1 = 0/min
- p2 = 0.025/min
- p3 = 0.000013 mU/liter
- n = 5/54 min-1

It is important to keep track of units when using these parameters. The concentrations listed above are in mmol/liter, but the glucose disturbance has units of grams. Therefore, it is necessary for us to apply a conversion factor of 5.5556 mmol/grams to the G_{meal} term. Using these variables solve one can solve for the steady states, and calculate the basal insulin infusion rate (Ub) such that is is equal to 16.1667 mU/min. For these parameters, the resulting state space model is

$$A = \begin{bmatrix} 0 & -4.5 & 0 \\ 0 & -0.025 & 0.000013 \\ 0 & 0 & \frac{-5}{54} \end{bmatrix}$$

and

$$B = \begin{bmatrix} 0 & 0.4630 \\ 0 & 0 \\ \frac{1}{12} & 0 \end{bmatrix}$$

It is common practice in the U.S. to describe glucose concentration in units of mg/deciliter as opposed to mmol/liter. Therefor, the units will be converted from mmol/liter to mg/deciliter. The molecular weight of glucose is 180g/mol, and therefore one it is necessary to multiply the glucose state (mmol/liter) by 18, so that the measured glucose output obtained will be in units of mg/deciliter. The following state-output relationship will handle that:

$$C = \begin{bmatrix} 18 & 0 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

Through Laplace transforms it will be found that the process transfer function is :

$$G_p(s) = \frac{-3.79}{(40s+1)(10.8s+1)s}$$

The disturbance transfer function due to pole/zero cancelation is simply :

$$G_d(s) = \frac{8.334}{s}$$

However, in reality glucose does not directly enter the blood stream. There is a "lag time" associated with the processing of glucose in the gut. It must first be processed here before entering the blood. However, it can be modeled as a first-order function, with a 20-minute time constant. This modifies the above equation for the disturbance transfer function to include the lag in the gut such that:

$$G_d(s) = \frac{8.334}{s(20s+1)}$$

Desired Control Performance

The steady-state glucose concentration of 4.5 mmol/liter corresponds to the glucose concentration of 81 mg/deciliter. A diabetic patient, in order to stay healthy, must keep their blood glucose concentration above 70 mg\deciliter. If the blood glucose concentration falls below 70 mg/deciliter the patient may be likely to faint, as this is a very typical short-term symptom of hypoglycemia. The insulin infusion rate (manipulated input) cannot fall below 0, and this constraint must be set in simulations. Numerous different control strategies can be used. The equation given for the process function above can be simplified down to the following form and a PD or PID controller can be used.

$$G_p(s) = \frac{k_p}{(\tau_p + 1)\, s}$$

A diabetic will know when they are consuming a meal, and therefore when their blood glucose concentration may rise. Therefore a feed-forward control system may be desired.

References

- Bequette, B. W. Process Control: Modeling, Design, and Simulation New Jersey: Prentice Hall, 2003. pp 81-83, 694-697.

Contributors and Attributions

- Written by: Robbert Appel, Jessica Rilly, Jordan Talia

This page titled 6.9: Blood Glucose Control in Diabetic Patients is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

7: MATHEMATICS FOR CONTROL SYSTEMS

7.1: Dirac delta (impulse) function
7.2: First-order Differential Equations
7.3: Second-order Differential Equations
7.4: Taylor Series
7.5: Laplace Transforms

This page titled 7: Mathematics for Control Systems is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

7.1: Dirac delta (impulse) function

Introduction

The Dirac delta function $\delta(t-t_0)$ is a mathematical idealization of an impulse or a very fast burst of substance at $t=t_0$. (Here we are considering time but the delta function can involve any variable.) The delta function is properly defined through a limiting process. One such definition is as a thin, tall rectangle, of width ε:

$$\delta\left(t-t_0\right)=\frac{1}{\epsilon}$$

for

\[t_{0}-\frac{\epsilon}{2}

and zero otherwise, in the limit that $\epsilon \to 0$.

Then, we have

$$\int_a^b \delta\left(t-t_0\right) dt=\lim_{\epsilon \to 0}\left[\epsilon \cdot\left(\frac{1}{\epsilon}\right)\right]=1$$

as long as $a<t_0<b$. When t_0 is outside the range of (a,b), then the integral is zero.

Likewise, for any function $f(t)$ that is continuous and differentiable (analytic) at t_0,

$$\int_a^b \delta\left(t-t_0\right) f(t) dt=\lim_{\epsilon \to 0}\left[\frac{1}{\epsilon} \int_{t_0-\epsilon/2}^{t_0+\epsilon/2} f(t) dt\right]=f\left(t_0\right)$$

where the quantity in the square brackets above is just the average value of $f(t)$ in the interval \(t_{0}-\frac{\epsilon}{2}

$$\int_0^\infty \delta(t-2) t^2 dt=2^2=4$$

$$\int_3^\infty \delta(t-2) t^2 dt=0$$

Delta function at the initial time

Note: if one of the limits of the integral coincides exactly with t_0, then the result is usually cut in half:

$$\int_{t_0}^b \delta\left(t-t_0\right) f(t) dt=f\left(t_0\right)/2$$

for $b>t_0$. For example:

$$\int_2^\infty \delta(t-2) t^2 dt=(1/2) 2^2=2$$

However, when we think of an impulse to a system at the initial time t_0, then we really consider that the entire delta function is added to the system - that is, the actual time is an infinitesimal amount beyond t_0; that is, $t=t_0^+$. In that case

$$\int_{t_0}^b \delta\left(t-t_0^+\right) f(t) dt=f\left(t_0\right)$$

Definition in terms of a Gaussian

Another equivalent definition of the delta functions are as a Gaussian function:

$$\delta\left(t-t_0\right)=\lim_{\sigma\to 0}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(t-t_0)^2}{2\sigma^2}}$$

All the properties work out to be the same where $a=\sqrt{2}\sigma$ is used.

Physical applications

In Control, the delta function is used an idealization of a very fast perturbation to the system. For example, if you dump a bucket of water into a tank, then the "flow rate" is essentially a delta function - a very highly peaked function, but with a net integral (the total amount of water in the bucket).

In mechanics, and example of the delta function is the force when hitting an object by a hammer. Say you hit a steel ball with a hammer. It moves with a certain velocity representing the total momentum transferred by the hammer. Rather than talk about the force x time (the net momentum transfer), one talks about an "impulse" which is the net momentum transferred in an infinitesimally short amount of time.

Relation to the step function

The **step function**, $\Theta(t-t_0)$, is the integral of the delta function or alternatively, the delta function is the derivative of the theta function, where $\Theta(t-t_0)$ is defined at 1 for $t>t_0$ and 0 for $t<t_0$:

$$\delta\left(t-t_0\right)=(d/dt)\Theta\left(t-t_0\right)$$

\[\int_{a}^{b} \delta\left(t-t_{0}\right) d t=\int_{a}^{b}(d / d t) \Theta\left(t-t_{0}\right) d t=\Theta\left(a-t_{0}\right)-\Theta\left(b-t_{0}\right)=1-0=1 \quad \text { for } a

Here, the smooth or gaussian definition of the delta function corresponds to a smooth representation of the Θ function as the integral of a gaussian or equivalently, the error function.

This page titled 7.1: Dirac delta (impulse) function is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

7.2: First-order Differential Equations

Introduction

We consider the general first-order differential equation:

$$\tau\frac{dy(t)}{dt}+y(t)=x(t)$$

The general solution is given by:

$$y(t)=y_0e^{-(t-t_0)/\tau}+\frac{e^{-(t-t_0)/\tau}}{\tau}\int_{t_0}^{t}x\left(t'\right)e^{(t'-t_0)/\tau}dt'$$

where $y_0=y\left(t-t_0\right)$. Note that t' is used to be distinguished from the upper limit t of the integral.

To obtain the general solution, begin with the first order differential equation:

$$\tau\frac{dy(t)}{dt}+y(t)=x(t)$$

Divide both sides by τ:

$$\frac{dy(t)}{dt}+\frac{1}{\tau}y(t)=\frac{1}{\tau}x(t)$$

Rewrite the LHS in condensed form using the integrating factor $e^{-t/\tau}$:

$$e^{-t/\tau}\frac{d}{dt}\left[e^{t/\tau}y(t)\right]=\frac{1}{\tau}x(t)$$

Notice how a chain differentiation will return the LHS to the previous form

Simplify:

$$\frac{d}{dt}\left[e^{t/\tau}y(t)\right]=\frac{1}{\tau}x(t)e^{t/\tau}$$

Integrate both sides:

$$e^{t/\tau}y(t)-e^{t_0/\tau}y\left(t_0\right)=\frac{1}{\tau}\int_{t_0}^{t}x\left(t'\right)e^{t'/\tau}dt'$$

Solve for $y(t)$:

$$y(t)=y_0e^{-(t-t_0)/\tau}+\frac{e^{-t/\tau}}{\tau}\int_{t_0}^{t}x\left(t'\right)e^{t'/\tau}dt'$$

Example 7.2.1: Solutions of First Order Differential Equations

Consider:

$$\frac{dy(t)}{dt}=x(t)$$

Multiplying both sides by dt gives:

$$\int_{t_0}^{t} \frac{dydt}{dt} = \int_{t_0}^{t} x\left(t'\right) dt'$$

The general solution is given as:

$$y(t) = y_0 + \int_{t_0}^{t} x\left(t'\right) dt'$$

Now Consider:

$$\frac{dy(t)}{dt} = -ay(t)$$

Dividing both sides by $y(t)$ gives:

$$\frac{1}{y(t)} \frac{dy(t)}{dt} = -a$$

which can be rewritten as:

$$\frac{d}{dt}[\ln(y(t))] = -a$$

Multiplying both sides by dt, integrating, and setting both sides of the equation as exponents to the exponential function gives the general solution:

$$y(t) = y_0 e^{-a(t-t_0)}$$

Now Consider:

$$\frac{dy(t)}{dt} = -ay(t) + x(t)$$

The detailed solution steps are as follows:

1. Separate $y(t)$ and $x(t)$ terms

$$\frac{dy(t)}{dt} + ay(t) = x(t)$$

2. Rewrite the LHS in condensed form using the "integrating factor" e^{-at}

$$e^{-at} \frac{d\left(e^{at} y(t)\right)}{dt} = x(t)$$

Notice how a chain differentiation will return the LHS to the form written in step 1

3. Divide both sides by e^{-at}

$$\frac{d\left(e^{at} y(t)\right)}{dt} = e^{at} x(t)$$

4. Multiply both sides by dt and integrate

$$e^{at} y(t) - e^{at_0} y_0 = \int_{t_0}^{t} x\left(t'\right) e^{at'} dt'$$

The general solution is given as:

$$y(t) = y_0 e^{-a(t-t_0)} + e^{-at} \int_{t_0}^{t} x\left(t'\right) e^{a(t'-t_0)} dt'$$

Step Function Simplification

$$\tau\frac{dy(t)}{dt}+y(t)=x(t)$$

For $x(t)=\Theta\left(t-t_0\right)$, which is the step function and $y_0=0$

That is, $x(t)=1$ for $t_0>t_o$ and $x(t)=0$ otherwise:

The previously derived general solution:

$$y(t)=y_0e^{-(t-t_0)/\tau}+\frac{e^{-(t-t_0)/\tau}}{\tau}\int_{t_0}^{t}x\left(t'\right)e^{(t'-t_0)/\tau}dt'$$

reduces to:

$$y(t)=e^{-(t-t_0)/\tau}$$

since the integral equals 1.

The quantity τ can be seen to be the time constant whereby $y(t)$ drops to $1/e$ of its original value

Example 7.2.2

Consider the differential equation

$$\frac{dy(t)}{dt}=-0.5y(t)+x(t)$$

where $x(t)=2+0.01t$

Assuming $y_o=0$, what is the behavior of $y(t)$?

Solution

General solution was derived previously as:

$$y(t)=y_0e^{-a(t-t_0)}+e^{-at}\int_{t_0}^{t}x\left(t'\right)e^{a(t'-t_0)}dt'$$

For $a=0.5$ and $y_o=0$, while setting $t_o=0$, the solution reduces to:

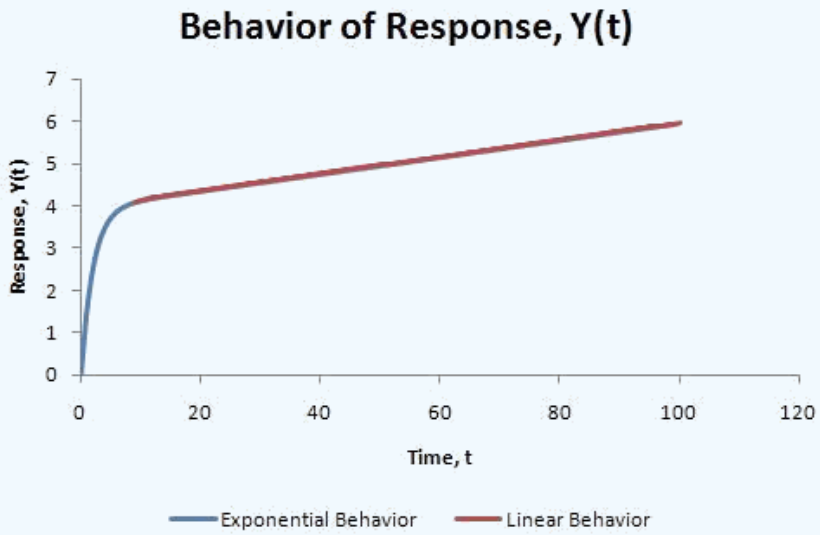
$$y(t)=e^{-0.5t}\int_{0}^{t}\left(2+0.01t'\right)e^{0.5(t')}dt'$$

The following link may be referred to for integral tables: S.O.S. Math

Simplifying the solution gives the following:

$$y(t)=\frac{\left(e^{-0.5t}\right)\left(\left(x+198\right)\left(e^{0.5t}\right)-198\right)\right)}{50}$$

Plotting this function displays the following behavior:



As can be seen clearly from the graph, initially the systemic response shows more exponential characteristics. However, as time progresses, linear behavior dominates.

This page titled 7.2: First-order Differential Equations is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

7.3: Second-order Differential Equations

Introduction

We consider the general Second-order differential equation:

$$\tau^2 \frac{d^2Y(t)}{dt^2} + 2\zeta\tau\frac{dY(t)}{dt} + Y(t) = X(t)$$

If you expand the previous Second-order differential equation:

$$\tau_1\tau_2\frac{d^2Y(t)}{dt^2} + (\tau_1 + \tau_2)\frac{dY(t)}{dt} + Y(t) = X(t)$$

$$\left(\tau_1\frac{d}{dt} + 1\right)\left(\tau_2\frac{d}{dt} + 1\right)Y(t) = X(t)$$

where:

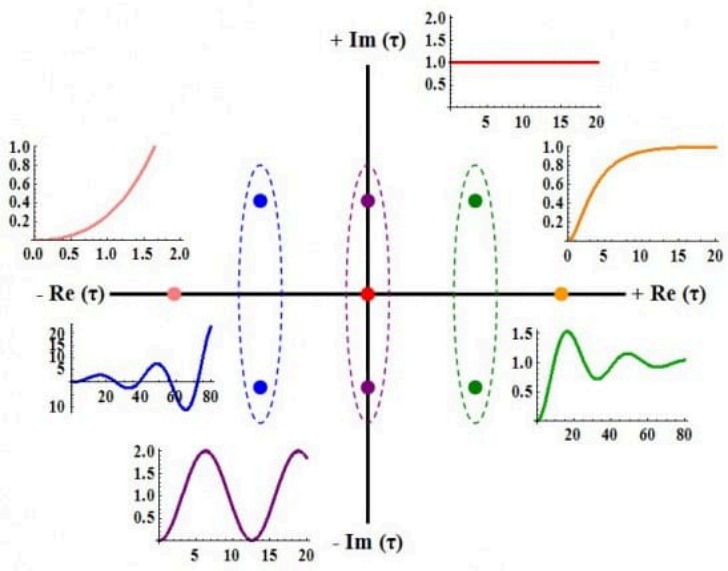
$$\tau = \sqrt{\tau_1\tau_2}$$

$$\zeta = \frac{\tau_1 + \tau_2}{2\sqrt{\tau_1\tau_2}}$$

Expansion of the differential equation allows you to guess what the shape of the solution ($Y(t)$) will look like when $X(t) = 1$.

The following rules apply when $\tau_1 = Re(\tau_1) + i * Im(\tau_1)$ and $\tau_2 = Re(\tau_2) + i * Im(\tau_2)$:

- If Re(τ_1) and Re(τ_2) are both positive, and there are NO imaginary parts, $Y(t)$ will exponentially decay (overdamped).
- If Re(τ_1) and Re(τ_2) are both positive, and there ARE imaginary parts, $Y(t)$ will oscillate until it reaches steady state (underdamped).
- If Re(τ_1) and Re(τ_2) are both negative, and there are NO imaginary parts, $Y(t)$ will exponentially grow (unstable).
- If Re(τ_1) and Re(τ_2) are both negative, and there ARE imaginary parts, $Y(t)$ will oscillate and grow exponentially (unstable).
- If Re(τ_1) and Re(τ_2) are both zero, and there ARE imaginary parts, $Y(t)$ will oscillate and neither grow nor decay.
- If τ_1 and τ_2 are both zero, $Y(t)$ is equal to X(t).



Solution of the General Second-Order System (When X(t)= θ(t))

The solution for the output of the system, $Y(t)$, can be found in the following section, if we assume that the input, $X(t)$, is a step function $\theta(t)$. The solution will depend on the value of ζ. If ζ is less than one, $Y(t)$ will be underdamped. This means that the output will overshoot and oscillate. If ζ is equal to one, $Y(t)$ will be critically damped. This means that the output will reach the steady state value quickly, without overshoot or oscillation. If ζ is greater than one, $Y(t)$ will be overdamped. This means that the output will not reach the steady state value as quickly as a critically damped system, but there will be no overshoot or oscillation.

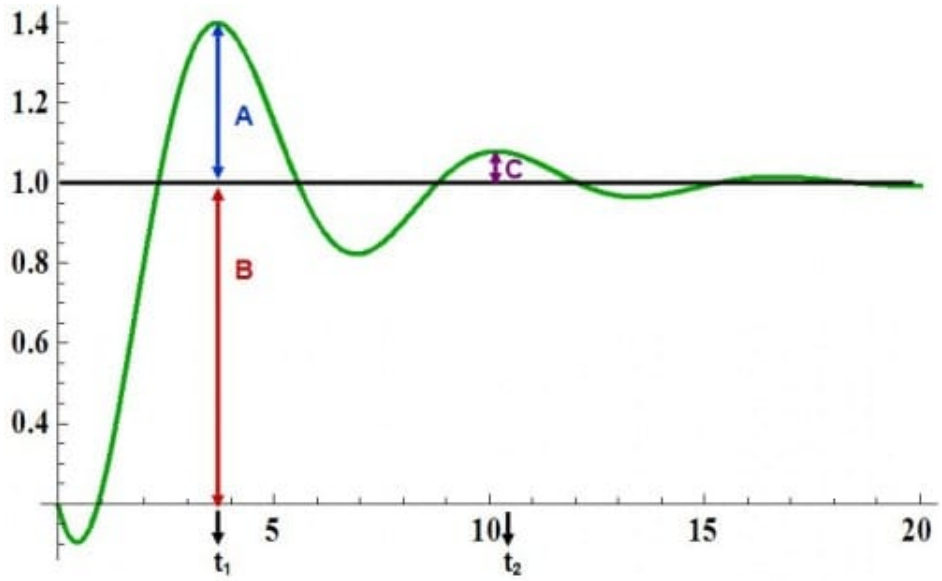
Underdamped (ζ<1)

If $\zeta < 1$, the solution is:

$$Y(t) = 1 - \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta t/\tau} \sin\left(\sqrt{1-\zeta^2}\frac{t}{\tau} + \phi\right)$$

where:

$$\phi = -\tan^{-1}\left(\frac{\sqrt{1-\zeta^2}}{\zeta}\right)$$



The decay ratio (C/A) can be calculated using the following equation:

$$\frac{C}{A} = e^{-2\pi\zeta/\sqrt{1-\zeta^2}}$$

The overshoot (A/B) can be calculated using the following equation:

$$\frac{A}{B} = e^{-\pi\zeta/\sqrt{1-\zeta^2}}$$

The period (T) and the frequency (ω) are the following:

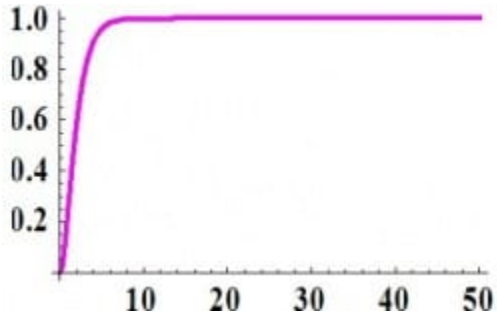
$$T = t_2 - t_1 = \frac{2\pi\tau}{\sqrt{1-\zeta^2}}$$

$$\omega = \frac{2\pi}{T} = \frac{\sqrt{1-\zeta^2}}{\tau}$$

Critically Damped (ζ=1)

If $\zeta = 1$, the solution is:

$$Y(t) = 1 - \left(1 + \frac{t}{\tau}\right) e^{-t/\tau}$$



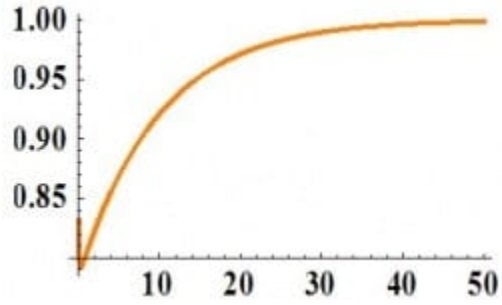
Overdamped (ζ>1)

If $\zeta > 1$, the solution is:

$$Y(t) = 1 - \frac{1}{\sqrt{\zeta^2 - 1}} e^{-\zeta t/\tau} \sinh\left(\sqrt{\zeta^2 - 1}\frac{t}{\tau} + \phi\right)$$

where:

$$\phi = -\tanh^{-1}\left(\frac{\sqrt{\zeta^2 - 1}}{\zeta}\right)$$



Example 7.3.1

Given:

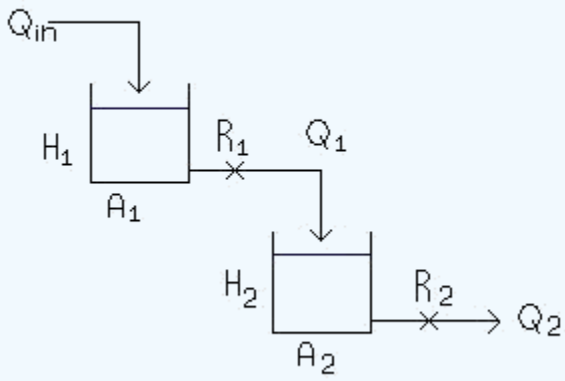
- $A_1 = 1\ m^2$
- $A_2 = 1.5\ m^2$
- $R_1 = 0.25\ s/m^2$
- $R_2 = 0.75\ s/m^2$

where:

- A is the area of the tank
- Q is the volumetric flowrate
- R is the resistance to the flow of the stream exiting the tank

- H is the height of liquid in the tank

develop an expression describing the response of H_2 to Q_{in}. Determine if the system is over, under or critically damped and determine what the graph of the expression would look like using the complex τ plane above. A diagram of the system is shown below:



Solution

Performing a mass balance on each tank:

$$A_1 \frac{dH_1}{dt} = Q_{in} - \frac{H_1}{R_1} \tag{7.3.1}$$

$$A_2 \frac{dH_2}{dt} = \frac{H_1}{R_1} - \frac{H_2}{R_2} \tag{7.3.2}$$

where the left hand terms account for the accumulation in the tank and the right hand terms account for the flow in the entering and exiting streams

Let $\tau_1 = R_1A_1$ and $\tau_2 = R_2A_2$

Equations 7.3.1 and 7.3.2 now become

$$\tau_1 \frac{dH_1}{dt} = R_1 Q_{in} - H_1 \tag{7.3.3}$$

$$\tau_2 \frac{dH_2}{dt} = \frac{R_2}{R_1} H_1 - H_2 \tag{7.3.4}$$

Put like terms on the same side and factor

$$\left(\tau_1 \frac{d}{dt} + 1\right) H_1 = R_1 Q_{in} \tag{7.3.5}$$

$$\left(\tau_2 \frac{d}{dt} + 1\right) H_2 = \frac{R_2}{R_1} H_1 \tag{7.3.6}$$

Apply $\left(\tau_1 \frac{d}{dt} + 1\right)$ operator from Equation 7.3.5 to Equation 7.3.6

$$\left(\tau_1 \frac{d}{dt} + 1\right)\left(\tau_2 \frac{d}{dt} + 1\right) H_2 = \left(\tau_1 \frac{d}{dt} + 1\right) \frac{R_2}{R_1} H_1 \tag{7.3.7}$$

The $\left(\tau_1 \frac{d}{dt} + 1\right) H_1$ term from the left hand portion of Equation 7.3.5 can be substituted into the right hand side of Equation 7.3.7

$$\left(\tau_1 \frac{d}{dt} + 1\right)\left(\tau_2 \frac{d}{dt} + 1\right) H_2 = R_1 Q_{in} \frac{R_2}{R_1}$$

$$\left(\tau_1 \frac{d}{dt} + 1\right)\left(\tau_2 \frac{d}{dt} + 1\right) H_2 = Q_{in} R_2$$

This expression shows the response of H_2 to Q_{in} as a second order solution like those pictured above. Here $Y(t)=H_2$ and $X(t)=R_2\ Q_{in}$

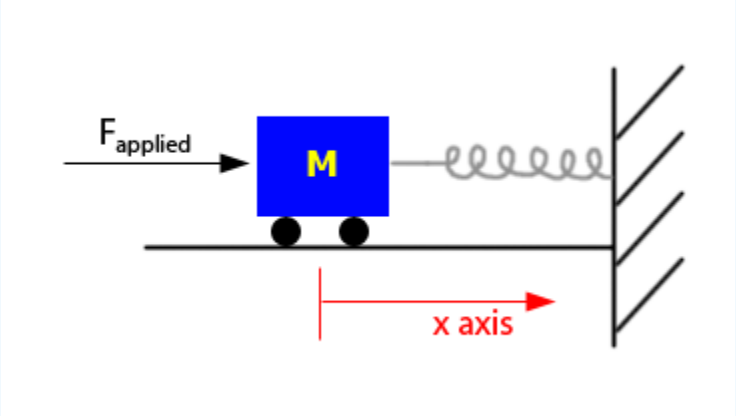
$$\zeta = \frac{\tau_1 + \tau_2}{2\sqrt{\tau_1 \tau_2}} = \frac{(0.25 * 1) + (0.75 * 1.5)}{2\sqrt{(0.25 * 1)(0.75 * 1.5)}} = 1.296$$

Overdamped.

Both values of τ are positive real numbers, and the behavior of the graph of the equation can be found on the complex τ plane above.

Example 7.3.2: Analogy to Physics - Spring System

It might be helpful to use a spring system as an analogy for our second order systems.



From Newton's second law of motion,

$$F = ma$$

where:

- F is Force
- m is mass
- a is acceleration

For the spring system, this equation can be written as:

$$F_{\text{applied}} - F_{\text{friction}} - F_{\text{restoring}} = mx''$$

where \(*x*'\) is the acceleration of the car in the x-direction

$$F_{\text{applied}} - fx - kx = mx''$$

where:

- *k* is the spring constant, which relates displacement of the object to the force applied
- *f* is the frequency of oscillation

$$\frac{m}{k}x'' + \frac{f}{k}x' + x = F_{applied}$$

As you can see, this equation resembles the form of a second order equation. The equation can be then thought of as:

$$\mathrm{T}^2 X'' + 2\zeta \mathrm{T} X' + X = F_{\text{applied}}$$

$$\tau = \sqrt{\frac{m}{k}}$$

$$\zeta = \frac{f}{2\sqrt{mk}}$$

Because of this, the spring exhibits behavior like second order differential equations:

- If $\zeta > 1$ or $f > 2\sqrt{mk}$ it is overdamped
- If $\zeta = 1$ or $f = 2\sqrt{mk}$ it is critically damped
- If $\zeta < 1$ or $f < 2\sqrt{mk}$ it is underdamped

This page titled 7.3: Second-order Differential Equations is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

7.4: Taylor Series

A **Taylor Series** is a representation of a function in the form of an infinite sum. Each term is calculated from using a derivative of the function as well as a factorial. The following equation is the defining equation of a Taylor series:

$$f(x) = \int (a) - \frac{f'(a)}{1!}(x-a) + \frac{f'''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots - \frac{f^{(n)}(a)}{n!}(x-a)^n \tag{7.4.1}$$

where a is the point, around which the function is defined.

The Taylor approximation is most accurate at a, and becomes less precise as the distance from a increases. The following equation is the same as Equation 7.4.1 mentioned above, in summation notation:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

We will first consider a simple example, the function, e^x, defined around the point, 0:

$$e^x = e^0 + \frac{1}{1!}(x-0) + \frac{1}{2!}(x-0)^2 + \frac{1}{3!}(x-0)^3 + \cdots + \frac{1}{n!}(x-0)^n$$

The numerators of each equation are all 1, because any level derivative of e^x is e^x, and e^x, defined at 0 is always 1. The Taylor expansion simplifies to the following equation:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$$

Also note that a Taylor series centered around 0 is called a Maclaurin series. In addition to the e^x Maclaurin series, some other common, simplified Maclaurin series are listed here:

$$\cos(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \cdots$$

$$\sin(x) = \frac{1}{1!}x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \cdots$$

$$\cosh(x) = 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \frac{1}{6!}x^6 + \cdots$$

$$\sinh(x) = \frac{1}{1!}x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \frac{1}{7!}x^7 + \cdots$$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots + x^n$$

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots + \frac{(-1)^{n+1}}{n}x^n$$

These Maclaurin series can be modified to become Taylor series when the equation is not centered around zero.

Expansion of functions using Taylor series may be useful in derivations. Maclaurin series will be useful for when the function is being approximated for small values of x. Using an infinite number of terms is generally unnecessary to model the function around the center point. The first few terms of either a Taylor or Maclaurin series usually sufficiently approximate the function.

A stimulating discussion of Taylor series may be found in Comtet's "Calcul pratique des coefficients de Taylor d'une fonction algébrique" (Enseign. Math. 10, 267-270, 1964) as well as Whittaker and Watson's landmark treatise, "Forms of the Remainder in Taylor's Series." found in A Course in Modern Analysis, 4th ed.

This page titled 7.4: Taylor Series is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

7.5: Laplace Transforms

Introduction

Laplace Transform are frequently used in determining solutions of a wide class of partial diffferential equations. Like other transforms, Laplace transforms are used to determine particular solutions. In solving partial differential equations, the general solutions are difficult, if not impossible, to obtain. Thus, the transform techniques sometimes offer a useful tool for attaining particular solutions. The Laplace transform is closely related to the complex Fourier transform, so the Fourier integral formula can be used to define the Laplace transform and its inverse[3]. Integral transforms are one of many tools that are very useful for solving linear differential equations[1]. An integral transform is a relation of the form:

$$F(s) = \int_a^b K(s,t) f(t) dt$$

where $K(s,t)$ is the kernel of the transformation

The relation given above transforms the function *f* into another function *F*, which is called the transform of *f*. The general idea in using an integral transform to solve a differential equation is as follows:

1. Use the relation to transform a problem for an unknown function *f* into a simpler problem for *F*
2. Solve this simpler problem to find *F*
3. Recover the desired function *f* from its transform *F*

Several integral transforms are known to be useful in applied mathematics, but in this chapter we consider only the Laplace transform. This transform is defined in the following way Let *f*(*t*) be given for $t \leq 0$. Then the Laplace transform of *f*, which we will denote by $\mathcal{L}\{f(t)\}$, is defined by the equation:

$$\mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t) dt$$

Whenever this improper integral converges. The Laplace transform makes use of the kernel *K*(s,*t*) = e^{-}(s*t*). Since the solutions of linear differential equations with constant coefficients are bsed on teh exponential function, the Laplace transform is particularly useful for such equations.

Table 1 - Table of Laplace Transforms

Signal	Laplace Transform
1. $\delta^{(n)}(t)$	s^n
2. 1 or $u(t)$	$\frac{1}{s}$
3. $\frac{t^n \exp(-\alpha t)u(t)}{n!}$	$\frac{1}{(s+\alpha)^{n+1}}$
4. $\cos \omega_0 t\ u(t)$	$\frac{s}{s^2+\omega_0^2}$
5. $\sin \omega_0 t\ u(t)$	$\frac{\omega_0}{s^2+\omega_0^2}$
6. $\exp(-\alpha t) \cos \omega_0 t\ u(t)$	$\frac{s+\alpha}{(s+\alpha)^2+\omega_0^2}$
7. $\exp(-\alpha t) \sin \omega_0 t\ u(t)$	$\frac{\omega_0}{(s+\alpha)^2+\omega_0^2}$
8. Square wave: $u(t) - 2u\left(t - \frac{T_0}{2}\right) + 2u(t - T_0) - \cdots$	$\frac{1}{s}\frac{1-e^{-sT_0/2}}{1+e^{-sT_0/2}}$
9. $(\sin \omega_0 t - \omega_0 t \cos \omega_0 t)u(t)$	$\frac{2\omega_0^3}{(s^2+\omega_0^2)^2}$
10. $(\omega_0 t \sin \omega_0 t)u(t)$	$\frac{2\omega_0^2 s}{(s^2+\omega_0^2)^2}$
11. $\omega_0 t \exp(-\alpha t) \sin \omega_0 t\ u(t)$	$\frac{2\omega_0^2(s+\alpha)}{[(s+\alpha)^2+\omega_0^2]^2}$
12. $\exp(-\alpha t)(\sin \omega_0 t - \omega_0 t \cos \omega_0 t)u(t)$	$\frac{2\omega_0^3}{[(s+\alpha)^2+\omega_0^2]^2}$

Example Problems

Elementary Functions Using Laplace Transforms

Example 1. Given $f(c) = c$, c is a constant

$$mathcalL[c] = \int_0^\infty e^{-st} c\, dt$$

$$= \left[-\frac{ce^{-st}}{s} \right]_0^\infty$$

$$= \frac{c}{s}$$

Example 2. Given $f(c) = e^{at}$, a is a constant

$$mathcalL[e^{at}] = \int_0^\infty e^{-st} e^{at}\, dt$$

$$= \left[-\frac{e^{-(s-a)t}}{(s-a)} \right]_0^\infty$$

$$= \frac{1}{s-a}, \text{ where } s \geq a$$

Example 3. Given $f(t) = t^2$, Then

$$\mathcal{L}\left[t^2\right] = \int_0^\infty e^{-st} t^2 dt$$

Integration by parts yields

$$= \left[-\frac{t^2 e^{-st}}{s}\right]_0^\infty + \int_0^\infty \frac{e^{-st}}{s} 2t dt$$

Since $t^2 e^{-st} \to 0$ as $t \to \infty$, we have

$$mathcalL[t^2] = \frac{2}{s}\left[-\frac{e^{-st}}{s} t\right]_0^\infty + \frac{2}{s}\int_0^\infty \frac{e^{-st}}{s}\, dt$$

$$\mathcal{L}\left[t^2\right] = \frac{2}{s}\left[-\frac{e^{-st}}{s} t\right]_0^\infty + \frac{2}{s}\int_0^\infty \frac{e^{-st}}{s} dt$$

$$= \frac{2}{s^3}$$

Example 4. Given $f(t) = \sin \omega t$. Then

$$F(s) = \mathcal{L}[\sin \omega t] = \int_0^\infty e^{-st} \sin \omega t dt$$

$$= \left[-\frac{e^{-st}}{s} \sin \omega t\right]_0^\infty + \int_0^\infty \frac{e^{-st}}{s} \omega \cos \omega t\, dt$$

$$= \frac{\omega}{s}\left[-\frac{e^{-st}}{s} \cos \omega t\right]_0^\infty - \frac{\omega}{s}\int_0^\infty \frac{e^{-st}}{s} \omega \sin \omega t\, dt$$

$$F(s) = \frac{\omega}{s} - \frac{\omega^2}{s^2} F(s)$$

Thus, solving for $F(s)$, we obtain

$$\mathcal{L}[\sin \omega t] = \frac{\omega}{(s^2 + \omega^2)}$$

In Class Example Using Laplace Transforms

Solving a differential equation:

$$\frac{dy(t)}{dt} = -ay(t)$$

Rearranging the equation to one side, we have:

$$\frac{dy(t)}{dt} + ay(t) = 0$$

Next, we take the Laplace transform

$$(sy(s) - y_o) + ay(s) = 0$$

where $y(s) = L[y(t)]$ and $y_o = y(0)$ Solving for y(s), we find:

$$y(s) = \frac{y_o}{s+a}$$

$$y(t) = \mathcal{L}^{-1}\{y(s)\} = \mathcal{L}^{-1}\left\{\frac{y_o}{s+a}\right\} = y_o e^{-at}$$

References

1. William E. Boyce, Elementary Differential Equations and Boundary Value Problems (2005) Chapter 6 pp 307-440.
2. Dr. Ali Muqaibel, EE207 Signals & Systems (2009)
3. Tyn Myint-U, Partial Differential Equations for Scientists and Engineers (2005) pp 337 -341.

Contributors and Attributions

Written By: Paul Rhee, Eric Chuang, and David Mui

This page titled 7.5: Laplace Transforms is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

8: OPTIMIZATION

8.1: Introduction to Optimization
8.2: Linear Optimization
8.3: Non-linear Optimization

This page titled 8: Optimization is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

8.1: Introduction to Optimization

Optimization aims to obtain the best results in a given situation, or to minimize input to maximize benefit [1].

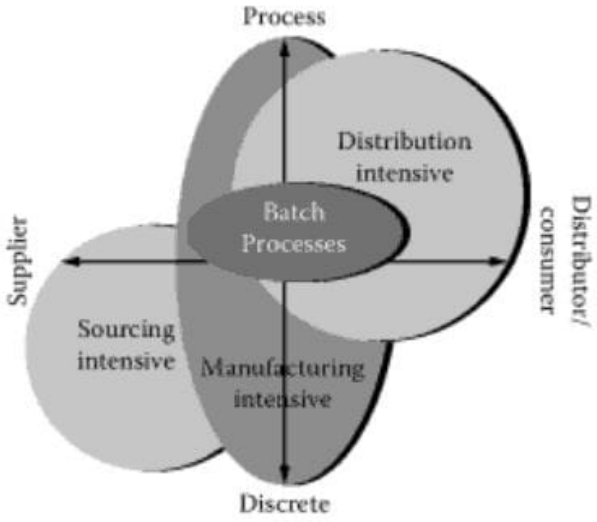
Engineering Application for Optimization

The following are examples of optimization across all engineering disciplines [1].

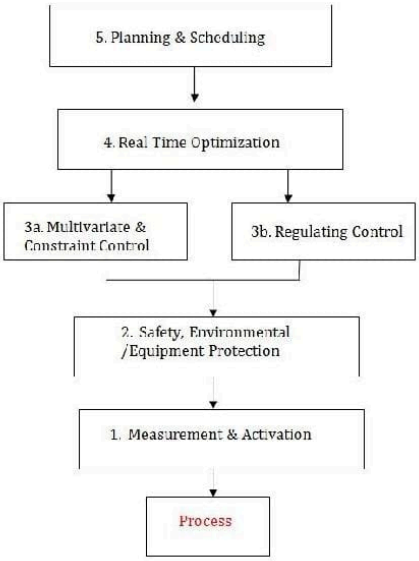
- Design of aircraft and aerospace structures for minimum weight
- Finding the optimal trajectories of space vehicles
- Design of civil engineering structures such as frames, foundations, bridges, towers, chimneys, and dams for minimum cost
- Minimum-weight design of structures for earthquake, wind, and other types of random loading
- Design of water resources systems for maximum benefit
- Optimal plastic design of structures
- Optimum design of linkages, cams, gears, machine tools, and other
- mechanical components
- Selection of machining conditions in metal-cutting processes for minimum production cost
- Design of material handling equipment such as conveyors, trucks, and cranes for minimum cost
- Design of pumps, turbines, and heat transfer equipment for maximum efficiency
- Optimum design of electrical machinery such as motors, generators,and transformers
- Optimum design of electrical networks
- Shortest route taken by a salesperson visiting various cities during one tour
- Optimal production planning, controlling, and scheduling
- Analysis of statistical data and building empirical models from experimental
- results to obtain the most accurate representation of the physical phenomenon
- Optimum design of chemical processing equipment and plants
- Design of optimum pipeline networks for process industries
- Selection of a site for an industry
- Planning of maintenance and replacement of equipment to reduce operating costs
- Inventory control
- Allocation of resources or services among several activities to maximize the benefit
- Controlling the waiting and idle times and queueing in production lines to reduce the costs
- Planning the best strategy to obtain maximum profit in the presence of a competitor
- Optimum design of control systems

Optimization across an Organization

The highest level of optimization includes the optimization of the raw material supply chain and the optimization of the packaging and product distribution chain. The plant wide optimization must consider documentation, maintenance, scheduling, and quality management considerations. Plant wide optimization resolves the conflict of objectives between the unit operations and the envelope strategies required to optimize the entire plant. Within the unit operations level multivariable optimization cannot be achieved when individual processing equipment is defective or when the control loops are not properly tuned. It is important that measurements be sampled fast enough, that controls loops be tuned for fast rates of recovery, and loop cycling be eliminated. When no mathematical model can describe a process, the process can only be optimized experimentally and empirical optimization is required [2].



Optimization Design



Constraints

Design constraints are physical limitations or restrictions that must be satisfied to produce an acceptable design [1].

- Operation conditions – safety, environmental
- Equipment constraints – e.g. pump rates
- Storage capacities
- Product quality and impurities

Optimization Situations

The following list describes common reasons for optimization in an industrial plant.

- Sales limited by production (e.g. reduce costs by minimizing downtime)

- Sales limited by market (e.g. be the “low cost producer”)
- Plants with high throughput
- High raw material or energy consumption
- Product quality better than specifications
- Loss of valuable or hazardous components through waste stream

Real-Time Optimization

The following describes the steps in order to optimize a chemical engineering process.

1. Identify process variables
2. Select objective function (e.g. profit $$$)
3. Develop process model and constraints
4. Simplify model to objective function (e.g linearization)
5. Compute the optimum
6. Perform sensitivity study

Industry Experience

Optimization can be applied to every aspect of a process. For example, at a refinery there are operators that work out in the units. Optimization can be applied to increase the operator/engineer communication by implementing "real-time" computer programs that allow the process engineers to see what is actually happening in the plant. This will help optimize process by allowing engineers to see what the conditions in the plant are in real time.

Pilot Plant Experience

Optimization of processes in a pilot plant will allow for more efficient scale-up to commercial size. In the ChE 460 course, four unit operations are optimized to produce soybean biodiesel. In order to optimize the reaction conversion, the project engineers vary catalyst concentration, agitation rate, and temperature. A design of experiment (DOE) is used to find the best set of input parameters.

References

1. Rao, Singiresu S. *Engineering Optimization - Theory and Practice (3rd Edition)*. (pp: 5). John Wiley & Sons. Online version available at: http://knovel.com/web/portal/browse/display?_EXT_KNOVEL_DISPLAY_bookid=1100&VerticalID=0
2. Lipták, Béla G. *Instrument Engineers' Handbook: Process control and optimization*. Boca Raten, FL: CRC Press, 2006.

This page titled 8.1: Introduction to Optimization is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Brittany Flaherty, Christine Curran, & Lauren Pakan via source content that was edited to the style and standards of the LibreTexts platform.

8.2: Linear Optimization

Linear optimization is a method applicable for the solution of problems in which the objective function and the constraints appear as linear functions of the decision variables. The constraint equations may be in the form of equalities or inequalities[1]. In other words, linear optimization determines the way to achieve the best outcome (for example, to maximize profit or to minimize cost) in a given mathematical model and given some lists of requirements represented as linear equations [2].

Applications

Linear optimization can be applied to numerous fields, in business or economics situations, and also in solving engineering problems. It is useful in modeling diverse types of problems in planning, routing, scheduling, assignment and design [2].

Some examples of applications in different industries

Petroleum refineries

One of the early industrial applications of linear optimization has been made in the petroleum refineries. An oil refinery has a choice of buying crude oil from different sources with different compositions at different prices. It can manufacture different products, such as diesel fuel, gasoline and aviation fuel, in varying quantities. A mix of the purchased crude oil and the manufactured products is sought that gives the maximum profit.

Manufacturing firms

The sales of a firm often fluctuate, therefore a company has various options. It can either build up an inventory of the manufactured products to carry it through the period of peak sales, or to pay overtime rates to achieve higher production during periods of high demand. Linear optimization takes into account the various cost and loss factors and arrive at the most profitable production plan.

Food-processing industry

Linear optimization has been used to determine the optimal shipping plan for the distribution of a particular product from different manufacturing plants to various warehouses.

Telecommunications

The optimal routing of messages in a communication network and the routing of aircraft and ships can also be determined by linear optimization method.

Characteristics of Linear Optimization

The characteristics of a linear optimization problem are:

1. *The objective function is of the minimization type*
2. *All the constraints are of the equality type*
3. *All the decision variables are non-negative*

Any linear optimization problem can be expressed in the standard form by using the following transformation:

1) The maximization of a function $f(x_1, x_2, \ldots, x_n)$ is equivalent to the minimization of the negative of the same function.

For example:

Minimize

$$f = c_1x_1 + c_2x_2 + \ldots + c_nx_n$$

is equivalent to

Maximize

$$f' = -f = -c_1x_1 - c_2x_2 - ... - c_nx_n$$

Consequently, the objective function can be stated in the minimization form in any linear optimization problem.

2) If a constraint appears in the form of a "less than or equal to" type of inequality as

$$a_{k1}x_1 + a_{k2}x_2 + ... + a_{kn}x_n \leq b_k$$

it can be converted into the equality form by **adding** a non-negative slack variable x_{n+1} as follows:

$$a_{k1}x_1 + a_{k2}x_2 + ... + a_{kn}x_n + x_{n+1} = b_k$$

Similarly, if the constraint is in the form of a "greater than or equal to" type of inequality, it can be converted into the equality form by **subtracting** the surplus variable x_{n+1}.

3) In most engineering optimization problems, the decision variables represent some physical dimensions, hence the variables x_j will be non-negative.

However, a variable may be unrestricted in sign in some problems. In such cases, an unrestricted variable (which can take a positive, negative or zero value) can be written as the difference of two non-negative variables.

Thus, if x_j is unrestricted in sign, it can be written as $x_j = x_j' - x_j''$, where

$$0 \leq x_j' \text{ and } 0 \leq x_j''$$

It can be seen that x_j will be negative, zero or positive, depending on whether x_j'' is greater than, equal to, or less than x_j'

Example 8.2.1

This example comes from Seborg, Edgar, and Mellinchamp (with some corrections). A chemical company is mixing chemical component, A and B to produce product, E and F. The chemical reaction is listed below:

$$A + B = E$$

$$A + 2B = F$$

The conditions of this production is listed as below:

	Cost of Materials	Maximum Available per day	Fixed Cost
A	$0.15/lbs	40,000lbs	350
B	$0.20/lbs	30,000lbs	200
	Value of Products	**Maximum Production Rate**	
E	$0.40/lbs	30,000lbs	
F	$0.33/lbs	30,000lbs	

The **profit** of this production can be simply described as the function below:

Profit

=	∑	F_sV_s –	∑	F_rV_r – C.P. – F.C.
	s		r	

- F_s = Flow Rates of Products
- F_r = Flow Rates of Raw Materials
- V_s = Values of Products
- V_r = Values of Raw Materials
- $C.P.$ = Cost of Production
- $F.C.$ = Fixed Cost

Constraints

- $0 < F_A < 40,000$
- $0 < F_B < 30,000$
- $0 < F_E < 30,000$
- $0 < F_F < 30,000$
- $\frac{1}{3}F_F + \frac{2}{3}F_E \geq 0$

Solution

$$F_A = \frac{1}{2}F_E + \frac{1}{3}F_F$$

$$F_B = \frac{1}{2}F_E + \frac{2}{3}F_F$$

$$\text{Profit } = \sum_s F_sV_s - \sum_r F_rV_r - C.P. - F.C.$$

$$= (0.4\psi_{F_j} + 0.33F_{F^-}) - (0.15F'_A - 0.2F'_H) - (0.15F_{F_i} + 0.05F'_F) - (350 + 200i$$

$$= (0.4F'_E + 0.33F'_F) - (0.15F'_A - 0.2F'_H) - (0.15F'_E + 0.05F'_F) - (350 + 200)$$

F_A0 " src="/@api/deki/files/18014/image-717.png">

\frac{1}{2} F_E + \frac{1}{3} F_F0 " src="/@api/deki/files/18016/image-718.png">

F_F-\frac{3}{2} F_E .....(1) " src="/@api/deki/files/18018/image-719.png">

$$F_A < 40,000$$

$$\frac{1}{2}F_E + \frac{1}{3}F_F < 40,000$$

$$F_F < 120,000 - \frac{3}{2}F_E.....(2)$$

F_B0 " src="/@api/deki/files/18025/image-723.png">

\frac{1}{2} F_E + \frac{2}{3} F_F0 " src="/@api/deki/files/18027/image-724.png">

F_F-\frac{3}{4} F_E .....(3) " src="/@api/deki/files/18029/image-725.png">

$$F_A < 30,000$$

$$\frac{1}{2}F_E + \frac{2}{3}F_F < 30,000$$

$$F_F < 45,000 - \frac{3}{4}F_E.....(4)$$

Solution using Mathematica

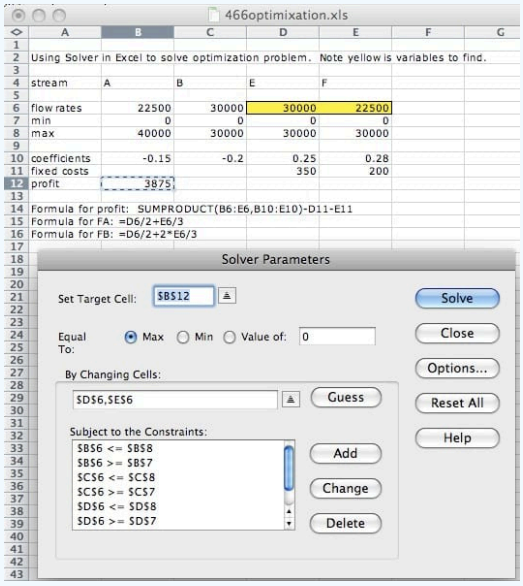
INPUT:

profit = 0.25 FE + 0.28 FF - 0.15 FA - 0.2 FB - 350 - 200

sol = Maximize[profit, {FA > 0, FB > 0, FA < 40000, FB < 30000, FE > 0, FE < 30000, FF > 0, FF < 30000, FA == (1/2) FE + (1/3) FF, FB == (1/2) FE + (2/3) FF}, {FE, FF}]

OUTPUT: {3875., {FE -> 30000., FF -> 22500.}}

Solution using "Solver" in Excel.

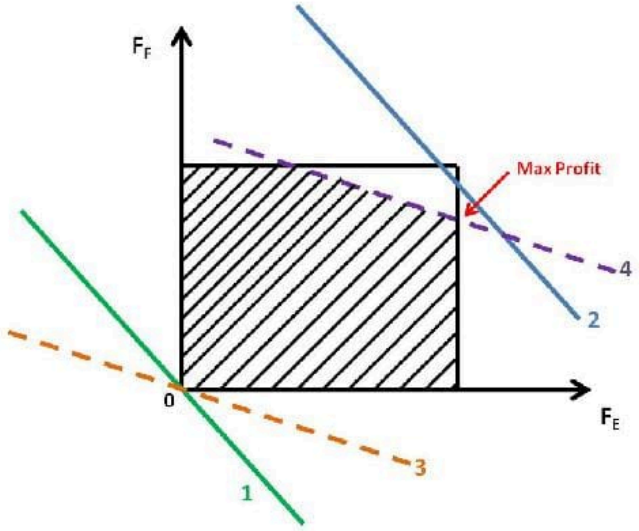


Result is FE = 30,000, FF = 22500.

If only process ! or process 2 were running a full capacity, the profit would be less.

Linear Optimization

The above is an example of linear optimization. It is often used in oil refinery to figure out maximal profit in response to market competition.



2.4.4 Graph

Example 2

Example of Linear Optimization Problem in Excel

Written by: Jennifer Campbell, Katherine Koterba, MaryAnn Winsemius

Part 1: Organize Given Information

As stated in the Linear Optimization section example above, there are three categories of information needed for solving an optimization problem in Excel: **an objective function, constraints, and decision variables**.

We will use the following example to demonstrate another application of linear optimization. We will be optimizing the profit for Company X's trucking business.

To reach capacity, Company X must move 100 tons of cargo per day by truck. Company X's trucking fee is $250/ton. Besides the weight constraint, the company can only move 50,000 ft^3 of cargo per day due to limited volume trucking capacity. The following amounts of cargo are available for shipping each day:

Cargo Weight (tons)	Weight (tons)	Volume (ft^3/ton)
1	30	550
2	40	800
3	50	400

Maximize the profit for Company X. Set up this problem:

Objective Function ($/week):

Profit = 250*(Cargo1 + Cargo2+ Cargo3)

Decision Variables (weight in tons):

Cargo 1

Cargo 2

Cargo 3

Constraints:

Weight: Cargo1 + Cargo2 + Cargo3 ≤ 100

Volume: 550*Cargo1 + 800*Cargo2 +400*Cargo3 ≤ 50000

Amount 1: Cargo1 ≤ 30

Cargo1 ≥ 0

Amount 2: Cargo2 ≤ 40

Cargo2 ≥ 0

Amount 3: Cargo3 ≤ 50

Cargo3 ≥ 0

Part 2: Set Up the Problem Using Excel

Solver is an Add-in for Microsoft Excel. It will be used to optimize Company X's profit. If 'Solver' is not on the 'Tools' menu in Excel, then use the following steps to enable it:

For Windows 2007:

- Click on the Office button at the top left corner of the screen. Click on the "Excel Options" button on the bottom right of the menu.
- Select "Add-ins." Make sure that "Excel Add-ins" is selected in the "Manage" drop down list. Click "Go."
- A new window will appear entitled "Add-ins." Select "Solver Add-in" by checking the box. Click "Go."
- A Configuration window will appear. Allow Office to install the Add-in.
- The solver has been successfully installed. (See Windows Help for more instruction.)

Use the figure below to set up your Excel worksheet.

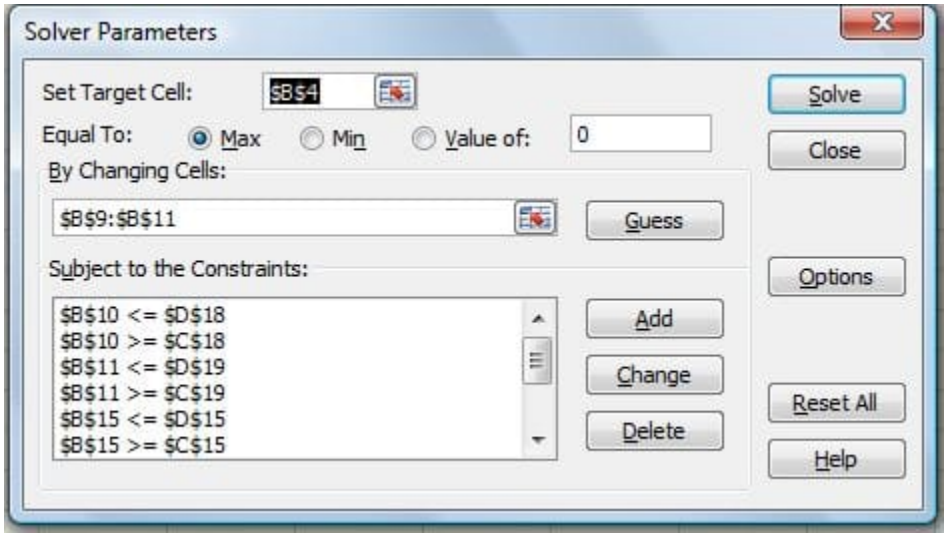
	A	B	C	D
1				
2				
3	**Objective**			
4	**Profit**	0		
5				
6	**Variables**			
7				
8				
9	**Cargo1**	0		
10	**Cargo2**	0		
11	**Cargo3**	0		
12				
13				
14	**Constraints**	**Formula**	**Min**	**Max**
15	Weight	0	0	100
16	Volume	0	0	50000
17	Amount1	0	0	30
18	Amount2	0	0	40
19	Amount3	0	0	50

Enter in the following formulas to the cells as shown below:

	A	B	C	D
1				
2				
3	**Objective**			
4	**Profit**	=250*SUM(B9:B11)		
5				
6	**Variables**			
7				
8				
9	**Cargo1**	0		
10	**Cargo2**	0		
11	**Cargo3**	0		
12				
13				
14	**Constraints**	**Formula**	**Min**	**Max**
15	Weight	=SUM(B9:B11)	0	100
16	Volume	=SUM(B9:B11)	0	50000
17	Amount1	=SUM(B9:B11)	0	30
18	Amount2	=SUM(B9:B11)	0	40
19	Amount3	=SUM(B9:B11)	0	50

Part 3: Running Solver

- Click on the “Data” tab and select “Solver”. A dialog box will appear.
- Enter the parameters as shown in the figure below.



Detailed steps are as follows:

- In “Set Target Cell,” enter the cell corresponding to the company’s profit (B4).
- Select “Max” under “Equal To.”
- Click on the “Options” tab and check the “Assume Linear Model” box.
- For “By Changing Cells:” select the cells in column B corresponding the cargo amounts (B9, B10, B11).

- To add constraints, select “Add” under “Subject to the Constraints” A dialog box will open.
- In the “Cell Reference:” field, enter the cell location of the decision value that is subject to constraint (i.e. B9).
- Use the pull-down menu in the middle to select the appropriate inequality relation (i.e. <=)
- In the “Constraint:” field, enter the cell location of the constraint value (i.e. D17).
- Continue to click the “Add.”
- Repeat above steps until all of the constraints are entered. Then click “OK.”

- When all the proper settings have been entered, click “Solve.”

- A “Solver Results” box will appear. Select “Keep Solver Solution” and click “OK.”.

The solved worksheet is below.

	A	B	C	D
1				
2				
3	**Objective**			
4	**Profit**	24218.75		
5				
6	**Variables**			
7				
8				
9	**Cargo1**	30		
10	**Cargo2**	16.875		
11	**Cargo3**	50		
12				
13				
14	**Constraints**	**Formula**	**Min**	**Max**
15	Weight	96.875	0	100
16	Volume	50000	0	50000
17	Amount1	0	0	30
18	Amount2	0	0	40
19	Amount3	0	0	50

Sensitivity Report

Written by Michael Chisholm and Doug Sutherland, Dec. 2009

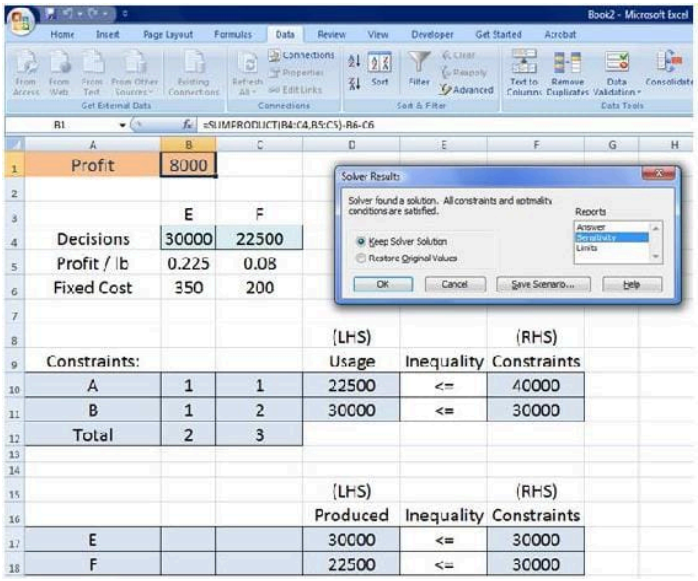
Excel's solver program allows us to analyze how our profit would change if we had an alteration in our constraint values. These values can change due to a variety of reasons such as more readily available resources, technology advancements, natural disasters limiting resources, etc.

First, it analyzes whether the constraints are binding or non-binding. The binding constraints limit the profit output where the non-binding constraints do not limit the overall process. If the non-binding constraints were changed, the profit would not be effected as long the change in these constraints lies within the allowable increase and decrease that is indicated within the sensitivity report. If the binding constraints are changed, the profit will be directly affected. The affect on the profit is shown with shadow price values, also displayed in the sensitivity report. The shadow price is the resulting increase or decrease in profit per unit increase or decrease in the constraint. This applies as long as the change in constraint remains within the allowable increase or decrease where a linear relationship can be assumed.

The shadow price only analyzes the change in one variable at a time. In order to do two, you must plug the new constraint value for one of the variables and solve using solver. Using the new sensitivity report, analyze the effect that changing the second variable would have with the change being made in the first constraint.

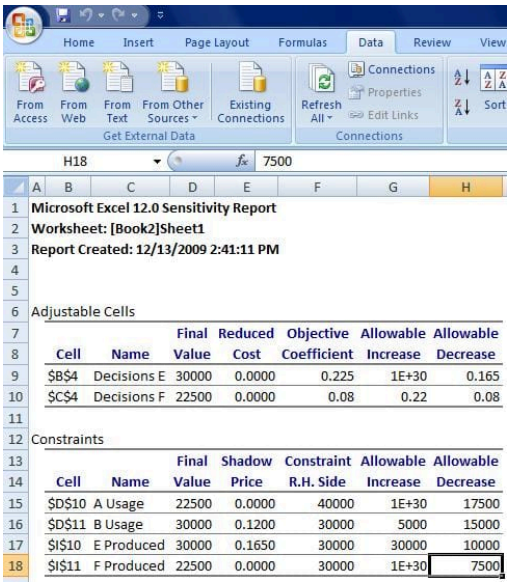
Looking at Example 1 above, we will now walk through the steps on how to create a sensitivity report.

After clicking "solve" in excel, a solver results dialogue box appears as seen below.



Profit	8000					
		E	F			
Decisions		30000	22500			
Profit / lb		0.225	0.08			
Fixed Cost		350	200			
				(LHS)		(RHS)
Constraints:				Usage	Inequality	Constraints
A		1	1	22500	<=	40000
B		1	2	30000	<=	30000
Total		2	3			
				(LHS)		(RHS)
				Produced	Inequality	Constraints
E				30000	<=	30000
F				22500	<=	30000

There is a list of three options on the right; answer, sensitivity, and limits. Select the sensitivity option before clicking ok. A new tab will be generated in the worksheet titled "sensitivity 1." A view of the sensitivity report within the tab is seen below. As you can see, two tables are generated. For this example, resource A and product F are non-binding as shown with a shadow price of 0 and an infinite allowable increase. The allowable decrease is the amount the capacity changes until the final value is reached. Past this point the constraint would become a binding constraint. For the constraining variables (resource B and product E), their constraints are binding. Regarding resource B, if its constraint was increased by up to 5000 or decreased by up to 15000, this would have a linear effect on profit within this range. For each unit increase or decrease, the profit will change by 12 cents per unit, respectively. The same is true if our capacity for product E changes with it's allowable values and shadow price on the table.



Microsoft Excel 12.0 Sensitivity Report
Worksheet: [Book2]Sheet1
Report Created: 12/13/2009 2:41:11 PM

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
B4	Decisions E	30000	0.0000	0.225	1E+30	0.165
C4	Decisions F	22500	0.0000	0.08	0.22	0.08

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
D10	A Usage	22500	0.0000	40000	1E+30	17500
D11	B Usage	30000	0.1200	30000	5000	15000
I10	E Produced	30000	0.1650	30000	30000	10000
I11	F Produced	22500	0.0000	30000	1E+30	7500

If the constraint on B increased by 5,000 lbs our new profit would be $8,600/day (8,000+.12*5,000). Instead, if our facility could increase production of E by 30,000 lb/day the resulting profit would be $12,950/day (8,000+.165*30,000).

Solving Linear Optimization Problems Using The Primal Simplex Algorithm

Written by: Tejas Kapadia and Dan Hassing [Note: needs specific reference, and also solution to the preceding problem by this method would be good -- R.Z.]

Instead of solving linear optimization problems using graphical or computer methods, we can also solve these problems using a process called the **Primal Simplex Algorithm**. The Primal Simplex Algorithm starts at a Basic Feasible Solution (BFS), which is a solution that lies on a vertex of the subspace contained by the constraints of the problem. In the Graph in Example 1, this subspace refers to the shaded region of the plot. Essentially, after determining an initial BFS, the Primal Simplex Algorithm moves through the boundaries from vertex to vertex until an optimal point is determined.

The basic procedure is the following:

1. Find a unit basis.
2. Set-up the problem in standard form using a canonical tableau.
3. Check optimality criterion.

1. If criterion passes, then stop, solution has been found.

4. Select an entering variable among the eligible variables.
5. Perform pivot step.
6. Go back to 1.

For simplicity, we will make the following assumptions:

1. The optimum lies on a vertex and is not unbounded on an extreme half-line.
2. The constraints are equations and not also inequalities.

1. In the case that the constraints are inequalities, slack variables will need to be introduced. Although the process is not very different in this case, we will ignore this to make the algorithm slightly less confusing.

3. Decision variables are required to be nonnegative.
4. The problem is a minimization problem. To turn a maximization problem into a minimization problem, multiply the objective function by -1 and follow the process to solve a minimization problem.

We will begin with the following example:

Objective Function: Minimize $z = -x_5 - 8x_6$

Subject to the constraints:

$$x_1 - x_5 + x_6 = 2 \tag{8.2.1}$$
$$x_2 + x_5 + x_6 = 1 \tag{8.2.2}$$
$$x_3 + 2x_5 + x_6 = 5 \tag{8.2.3}$$
$$x_4 + x_6 = 0 \tag{8.2.4}$$
$$x_i \geq 0 \tag{8.2.5}$$

First we begin by finding a unit basis:

A shortcut method to finding this unit basis is putting numbers in for each variable so that every constraint equation is satisfied.

In this case, setting $x_4 = 0$, $x_5 = 0$, and $x_6 = 0$ will satisfy the final equation and also set the values for x_1, x_2, and x_3 to 2, 1, and 5, respectively. Remember, these decision variables must be nonnegative.

Set up the canonical tableau in the following form:

x_1	x_2	x_3	x_4	x_5	x_6	-z	b
1	0	0	0	-1	1	0	2
0	1	0	0	1	1	0	1
0	0	1	0	2	1	0	5
0	0	0	1	0	1	0	0
0	0	0	0	-1	-8	1	0

As you can see, the first four rows correspond to the constraints, while the final row corresponds to the objective function. The "b" column corresponds to the right hand side (RHS) of the constraints. As you can see, the "-z" column is on the left hand side (LHS) of the equation, rather than the RHS.

First, we should perform pivot steps so that the tableau corresponds to the unit basis we found earlier. By performing pivot steps on x_1, x_2, x_3, and x_4, we will reach the feasible point where (x_1, x_2, x_3, x_4, x_5, and x_6) = $(2, 1, 5, 0, 0, 0)$. Because x_4, x_5, and x_6 all equal zero, the pivot step on x_4 can actually be done on x_5 or x_6, but in this example, we used x_4. These pivot steps can be performed on any row as long as they are all different rows. In this example, we performed pivot steps on $(x_1, 1)$, $(x_2, 2)$, $(x_3, 3)$, $(x_4, 4)$ using the Pivot and Gauss-Jordan Tool at people.hofstra.edu/Stefan_Waner/RealWorld/tutorialsf1/scriptpivot2.html. To use this tool, place the cursor on the cell that you wish to pivot on, and press "pivot".

After four pivot steps, the tableau will look like this:

x1	x2	x3	x4	x5	x6	x7	x8
1	0	0	0	-1	1	0	2
0	1	0	0	1	1	0	1
0	0	1	0	2	1	0	5
0	0	0	1	0	1	0	0
0	0	0	0	-1	-8	1	0

As you can see, this is identical to the initial tableau, as x_1, x_2, x_3, and x_4 were set up such that an initial feasible point was already chosen.

The optimality criterion states that if the vector in the bottom left of the tableau is all positive, then an optimal solution exists in the "b" column vector, with the value at the bottom of the "b" column vector as the negative of the value of the objective function at that optimal solution. If this is not true, then a pivot step must be performed. In this example, clearly, a pivot step must be performed.

Next, we need to choose an entering variable. We want to choose an entering variable that has a negative element in the bottom row, meaning that the objective value could be improved if that variable was nonzero in the solution. So, we will choose x_5 in this example. Now, we must calculate ratios of each RHS coefficient divided by the coefficient of the entering variable in that row. In this case, the vector corresponding to this calculation would equal $(2/-1, 1/1, 5/2, 0/0)$. We cannot pivot on a zero element, so we cannot pivot on the fourth row. We want to keep the RHS positive, so we cannot pivot on the first row. We must choose the minimum nonnegative ratio to remain at a feasible solution, so we choose the second row in the x_5 column, which has a ratio of 1/1.

After the pivot step:

x1	x2	x3	x4	x5	x6	x7	x8
1	1	0	0	0	2	0	3
0	1	0	0	1	1	0	1
0	-2	1	0	0	-1	0	3
0	0	0	1	0	1	0	0
0	1	0	0	0	-7	1	1

As we can see, x_6 has a negative coefficient in the bottom row, indicating the same step must be repeated on that column. We calculate ratios for that column, and get: $(3/2, 1/1, 3/-1, 0/1)$. Consequently, we choose to pivot on the fourth row because it corresponds to the minimum nonnegative ratio of 0.

After another pivot step:

x1	x2	x3	x4	x5	x6	x7	x8
1	1	0	-2	0	0	0	3
0	1	0	-1	1	0	0	1
0	-2	1	1	0	0	0	3
0	0	0	1	0	1	0	0
0	1	0	7	0	0	1	1

Because the bottom row is all positive, we are now at an optimal solution. To understand this final tableau, we look at each column for variables that only have one "1" in the column. If the column has only one "1", the RHS value in that row is the value of that variable. In this case, $x_1 = 3$, $x_3 = 3$, and $x_5 = 1$. Any variable that does not have just a single "1" in the column is equal to zero. So, the optimal solution is (x_1, x_2, x_3, x_4, x_5, and x_6) = $(3, 0, 3, 0, 1, 0)$, and the optimal value is $z = -1$ (z was on the LHS in the tableau).

Now, we have successfully solved a linear optimization problem using the Primal Simplex Algorithm. Verification of the solution can be easily performed in Microsoft Excel.

References

1. D. E. Seborg, T. F. Edgar, D. A. Mellichamp: Process Dynamics and Control, 2nd Edition, John Wiley & Sons.
2. Rao, Singiresu S. Engineering Optimization - Theory and Practice, 3rd Edition, 129-135, John Wiley & Sons.

This page titled 8.2: Linear Optimization is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Danny Hsiao, Huey Shann Sue, & Jenny Ou via source content that was edited to the style and standards of the LibreTexts platform.

8.3: Non-linear Optimization

Various conditions and situations are not adequately described using linear systems. In this case, nonlinear optimization may be applied. Unlike linear optimization, the optimal operating condition does not exist at the boundaries.

Quadratic Optimization

$$f(x) = c - x^T b + \frac{1}{2} x^T A x$$

To optimize, it is necessary to find when the gradient of f is equal to zero.

$$\nabla f(x) = 0$$

$$\nabla f(x) = b - Ax$$

$$x_* = A^{-1} b$$

It may be possible to solve the optimal x_* by a linear equation, approximated by a Taylor series.

$$f(x_*) = f(x) + (x_* - x)' \nabla f(x) + \frac{1}{2}(x_* - x)' \nabla\nabla f(x)(x_* - x) + \ldots$$

Iterative Methods

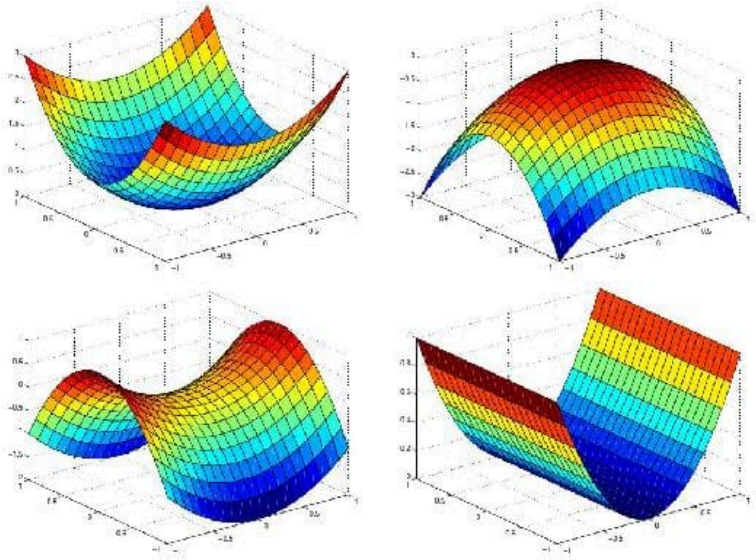
When direct methods cannot solve the equation (i.e. A is not symmetric positive definite), iterative methods are possible [1].

By starting with an initial guess of x_i , an algorithm may lead to a x_{i+1} that better satisfies the equation. Through iteration, theoretically, $x_\infty = x$.

Applications

- Finance: Portfolio optimization
- Businesses: Optimize inventory
- Engineering: Rigid body dynamics
- Biochemistry: Kinetic modeling [2]

Example: Typical Nonlinear 3d Curves



(Image from [1])

As observed, the optimal condition does not necessarily exist at the boundary of the curve.

Example: Quadratic Optimization

$$f(x) = \vec{c}^T\vec{x} + \frac{1}{2}\vec{x}^T Q\vec{x}$$

where

$$\vec{c}^T = (c_1, c_2, \dots, c_n)$$

$$\vec{x}^T = (x_1, x_2, \dots, x_n)$$

For a quadratic system, $n = 2$, thus, Q (the quadratic term constant) is defined as a symmetric matrix as follows.

$$Q = \begin{bmatrix} Q_1 & Q_3 \\ Q_3 & Q_2 \end{bmatrix}$$

Thus, multiplying out the f,

$$f(x) = (c_1x_1 + c_2x_2) + \frac{1}{2}\left(Q_1x_1^2 + 2Q_3x_1x_2 + Q_2x_2^2\right)$$

References

1. Lippert, Ross A. "Introduction to non-linear optimization." D.E. Shaw Research, February 25, 2008. http://www.mit.edu/~9.520/spring08/Classes/optlecture.pdf
2. Mendes, Pedro and Kell, Douglas B. "Non-linear optimization of biochemical pathways: application to metabolic engineering and parameter estimation." Journal of Bioinformatics, Volume 14, 869-883. 1998.
3. "Introduction to Non-linear optimization." Georgia Institute of Technology Systems Realization Laboratory. www.srl.gatech.edu/education/ME6103/NLP-intro.ppt

This page titled 8.3: Non-linear Optimization is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Andrew King, Edwin Yik, & Matthew Goh via source content that was edited to the style and standards of the LibreTexts platform.

9: PROPORTIONAL-INTEGRAL-DERIVATIVE (PID) CONTROL

9.1: Constructing Block Diagrams- Visualizing control measurements

9.2: P, I, D, PI, PD, and PID control

9.3: PID Tuning via Classical Methods

9.4: PID tuning via Frequency Responses with Bode Plots

9.5: PID tuning via Optimization

9.6: PID Downsides and Solutions

This page titled 9: Proportional-Integral-Derivative (PID) Control is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

9.1: Constructing Block Diagrams- Visualizing control measurements

It is often convenient to express a control system with a block diagram. These diagrams help to visualize the relationships between each part of the control system (Figure 9.1.1).

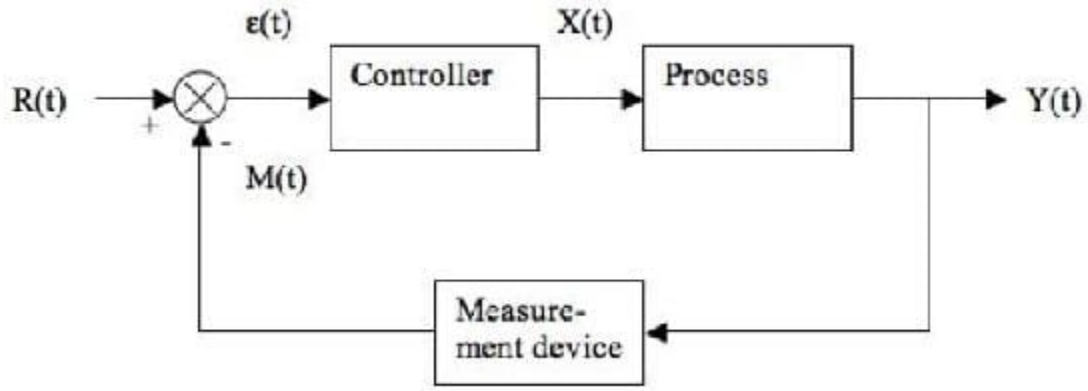


Figure 9.1.1: Block diagram representing a control system.

The control system shown is called a closed-loop system or a feedback system because the measured value of the controlled variable is returned or "fed back" to the comparator. In the comparator the controlled variable is compared with the desired value or set point. If there is any difference between the measured variable and the set point, an error is generated. This error enters a controller, which in turns adjusts the final control element to return the controlled variable to the set point [1].

- *R(t)* = "set point", is a synonym for the desired value of the controlled variable
- ε(*t*) = process error or comparator
- *X(t)* = process variable
- *M(t)* = measurement variable
- *Y(t)* = output variable

The "hat" above the G indicates that it is an operator (not a simple function) and may include integration and differentiation. Thus, you cannot divide by a $\hat{G}$, but have to multiply both sides by its inverse $\hat{G}^{-1}$ (see below). This is a more general representation that the standard Laplace transform form; here it is not necessary that the operators be linear.

From the block diagram the equations for a system can be constructed. Define the following:

- $\hat{G}_c$ =Controller operator
- $\hat{G}_p$ =Process operator
- $\hat{G}_m$ =Measurement delay operator

The equations for the system in Figure 9.1.1 are given below. Note that these equations are always written as "output is equal to operator applied to input."

- Process: $Y(t) = \hat{G}_p X(t)$
- Controller: $X(t) = \hat{G}_c \epsilon(t)$

- Comparator: ε(*t*) = *R*(*t*) – *M*(*t*)
- Measurement delay: $M(t) = \hat{G}_m Y(t)$

Typically to plot the response of the controller, $Y(t)$, it is necessary to solve these equations to obtain a single equation for $Y(t)$. This can be done by first plugging the measurement delay equation into the compartor equation, or.

$$\epsilon(t) = R(t) - \hat{G}_m Y(t).$$

This equation can be put into the controller equation, or

$$X(t) = \hat{G}_c \left[R(t) - G_m Y(t)\right].$$

Finally, the equation for $Y(t)$ can be obtained by putting the above equation into the equation for the process, or

$$Y(t) = \hat{G}_p \hat{G}_c \left[R(t) - \hat{G}_m Y(t)\right].$$

Frequently we have a differential equation for $X(t)$ in terms of derivatives of $Y(t)$ so we write

$$X(t) = \hat{G}_p^{-1} Y(t)$$

which implies the form of the general equation

$$\hat{G}_p^{-1} Y(t) = \hat{G}_c \left[R(t) - G_m Y(t)\right]$$

If the process, controller, and measurement delay operators are known, $Y(t)$ can be plotted and the response of the system can be observed (i.e. using Mathematica).

That is, we have

1. $Y(t) = \hat{G}_p X(t)$
2. $X(t) = \hat{G}_c \epsilon(t)]$
3. $\epsilon(t) = R(t) - M(t)$
4. $M(t) = \hat{G}_m Y(t)$
5. $\hat{G}_p^{-1} Y(t) = \hat{G}_c \left[R(t) - G_m Y(t)\right]$

Assume no measurement lag:

$$\hat{G}_m = 1$$

Assume first-order system:

$$\hat{G}_p^{-1} = \frac{1}{K_p}\left(\tau_p \frac{d}{dt} + 1\right)$$

Assume PI controller:

$$\hat{G}_c^{-1} = K_c + \frac{K_c}{\tau_I}\int_0^t dt$$

Substituting equations (2) and (3) into (5) gives:

$$\frac{1}{K_p}\left(\tau_p \frac{d}{dt} \quad 1\right) Y(t) = \left(K_z, \frac{K_e}{\eta_1}\int_0^t at\right) c(t) - \left(K_z | \frac{K_c}{7_1}\int_0^t dt\right) \{H(t) \quad Y(t)$$

Simplifying further gives:

$$\tau_p Y'(t) + Y(t) = K_c(R(t) - Y(t)) + \frac{K_c}{\tau_I}(R(t) - Y(t))$$

Differentiating the above equation gives:

$$\tau_p Y''(t) + Y'(t) = K_c R'(t) - K_c Y'(t) + \frac{K_c}{\tau_I} R(t) - \frac{K_c}{\tau_I} Y(t)$$

Rearranging the above equation:

$$\tau_p Y''(t) + (1 + K_c) Y'(t) + \frac{K_c}{\tau_I} Y(t) = K_c R'(t) + \frac{K_c}{\tau_I} R(t)$$

Multiply the above equation by $\frac{\tau_I}{K_c}$ to eliminate the co-efficient of $Y'(t)$ gives:

$$\frac{\tau_I \tau_p}{K_c} Y''(t) + \tau_I \frac{1 + K_c}{K_c} Y'(t) + Y(t) = \tau_I R'(t) + R(t)$$

References

1. Coughanowr, D.R. and S.E. LeBlanc (2009). Process Systems Analysis And Control. Third Edition. Mc-Graw Hill.

This page titled 9.1: Constructing Block Diagrams- Visualizing control measurements is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

9.2: P, I, D, PI, PD, and PID control

Introduction

Process controls are necessary for designing safe and productive plants. A variety of process controls are used to manipulate processes, however the most simple and often most effective is the **PID controller**. The controller attempts to correct the error between a measured process variable and desired setpoint by calculating the difference and then performing a corrective action to adjust the process accordingly. A PID controller controls a process through three parameters: Proportional (P), Integral (I), and Derivative (D). These parameters can be weighted, or tuned, to adjust their effect on the process.

Much more practical than the typical on/off controller, PID controllers allow for much better adjustments to be made in the system. While this is true, there are some advantages to using an on/off controller including that they are (1) relatively simple to design and execute and (2) binary sensors and actuators (such as an on/off controller) are generally more reliable and less expensive.

Although there are some advantages, there are large disadvantages to using an on/off controller scheme. They are (1) inefficient (using this control is like driving with full gas and full breaks), (2) can generate noise when seeking stability (can dramatically overshoot or undershoot a set-point), and (3) physically wearing on valves and switches (continuously turning valves/switches fully on and fully off causes them to become worn out much quicker).

To allow for much better control and fine-tuning adjustments, most industrial processes use a PID controller scheme. This section will provide a brief introduction on PID controllers as well as methods to model a controlled system in Excel.

The **Process Gain** (K) is the ratio of change of the output variable (responding variable) to the change of the input variable (forcing function). It specifically defines the sensitivity of the output variable to a given change in the input variable.

$$K = \frac{\Delta \text{ Output}}{\Delta \text{input}}$$

Gain can only be described as a steady state parameter and give no knowledge about the dynamics of the process and is independent of the design and operating variables. A gain has three components that include the sign, the value, the units. The sign indicates how the output responds to the process input. A positive sign shows that the output variable increases with an increase in the input variable and a negative sign shows that the output variable decreases with an increase in the input variable. The units depend on the process considered that depend on the variables mentioned.

For example, the pressure was increased from 21 psi to 29 psi. This change increased the valve position from 30% vp to 22% vp.

$$K = \frac{(29-21)\,psi}{(22-30)\%vp} = -1.0\,psi/(\%vp)$$

Dead Time (t_0) is the between the change in an input variable and when the output variable begins. Dead time is important because it effects the controllability of the control system. A change in set point is not immediate because of this parameter. Dead time must be considered in tuning and modeling processes.

Types of Control

Process controls are instruments used to control a parameter, such as temperature, level, and pressure. PID controllers are a type of continuous controller because they continually adjust the output vs. an on/off controller, when looking at feed forward or feed backward conditions. An example of a temperature controller is shown in Figure $9.2.1$.

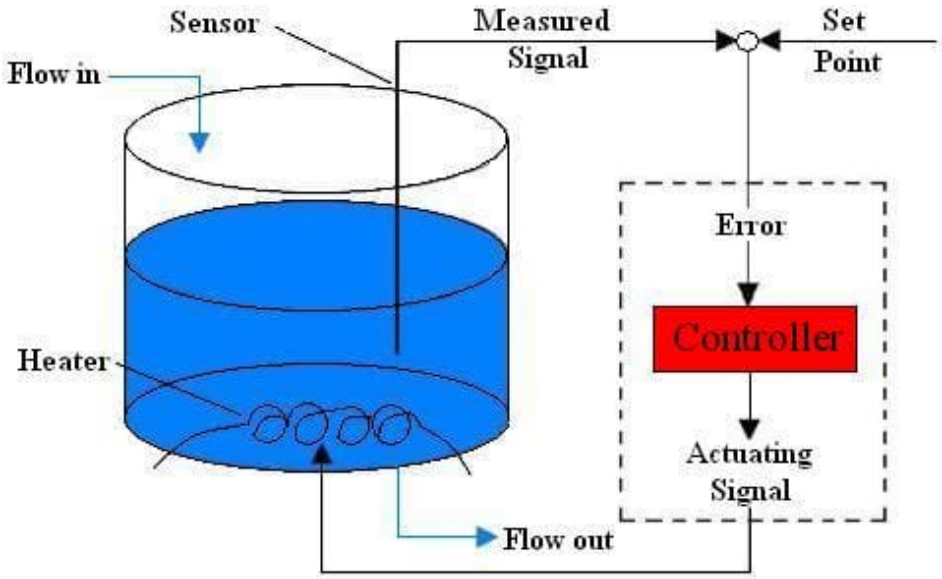


Figure 9.2.1: Temperature controller in a CSTR

As shown in Figure 1, the temperature controller controls the temperature of a fluid within a CSTR (Continuous Stirred Tank Reactor). A temperature sensor first measures the temperature of the fluid. This measurement produces a measurement signal. The measurement signal is then compared to the set point, or desired temperature setting, of the controller. The difference between the measured signal and set point is the error. Based on this error, the controller sends an actuating signal to the heating coil, which adjusts the temperature accordingly. This type of process control is known as error-based control because the actuating signal is determined from the error between the actual and desired setting. The different types of error-based controls vary in the mathematical way they translate the error into an actuating signal, the most common of which are the PID controllers. Additionally, it is critical to understand feed-forward and feed-back control before exploring P, I, and D controls. Feed Forward Control Feedback Control

I, D, PI, PD, PID Control

As previously mentioned, controllers vary in the way they correlate the controller input (error) to the controller output (actuating signal). The most commonly used controllers are the **proportional-integral-derivative (PID) controllers**. PID controllers relate the error to the actuating signal either in a proportional (P), integral (I), or derivative (D) manner. PID controllers can also relate the error to the actuating signal using a combination of these controls.

Proportional (P) Control

One type of action used in PID controllers is the proportional control. Proportional control is a form of feedback control. It is the simplest form of continuous control that can be used in a closed-looped system. P-only control minimizes the fluctuation in the process variable, but it does not always bring the system to the desired set point. It provides a faster response than most other controllers, initially allowing the P-only controller to respond a few seconds faster. However, as the system becomes more complex (i.e. more complex algorithm) the response time difference could accumulate, allowing the P-controller to possibly respond even a few minutes faster. Although the P-only controller does offer the advantage of faster response time, it produces deviation from the set point. This deviation is known as the offset, and it is usually not desired in a process. The existence of an offset implies that the system could not be maintained at the desired set point at steady state. It is analogous to the systematic error in a calibration curve, where there is always a set, constant error that prevents the line from crossing the origin. The offset can be minimized by combining P-only control with another form of control, such as I- or D- control. It is important to note, however, that it is impossible to completely eliminate the offset, which is implicitly included within each equation.

P-control linearly correlates the controller output (actuating signal) to the error (diference between measured signal and set point). This P-control behavior is mathematically illustrated in Equation 9.2.1 (Scrcek, *et. al*).

$$c(t) = K_c e(t) + b \tag{9.2.1}$$

where

- $c(t)$ = controller output
- K_c = controller gain
- $e(t)$ = error
- b = bias

In this equation, the bias and controller gain are constants specific to each controller. The bias is simply the controller output when the error is zero. The controller gain is the change in the output of the controller per change in the input to the controller. In PID controllers, where signals are usually electronically transmitted, controller gain relates the change in output voltage to the change in input voltage. These voltage changes are then directly related to the property being changed (i.e. temperature, pressure, level, etc.). Therefore, the gain ultimately relates the change in the input and output properties. If the output changes more than the input, K_C will be greater than 1. If the change in the input is greater than the change in the output, K_C will be less than 1. Ideally, if K_C is equal to infinity, the error can be reduced to zero. However, this infinitesimal nature of K_C increases the instability of the loop because zero error would imply that the the measured signal is *exactly* equal to the set point. As mentioned in lecture, exact equality is never achieved in control logic; instead, in control logic, error is allowed to vary within a certain range. Therefore, there are limits to the size of K_C, and these limits are defined by the system. Graphical representations of the effects of these variables on the system is shown in PID Tuning via Classical Methods.

As can be seen from the above equation, P-only control provides a linear relationship between the error of a system and the controller output of the system. This type of control provides a response, based on the signal that adjusts the system so that any oscillations are removed, and the system returns to steady-state. The inputs to the controller are the set point, the signal, and the bias. The controller calculates the difference between the set point and the signal, which is the error, and sends this value to an algorithm. Combined with the bias, this algorithm determines the action that the controller should take. A graphical representation of the P-controller output for a step increase in input at time t0 is shown below in Figure 2. This graph is exactly similar to the step input graph itself.

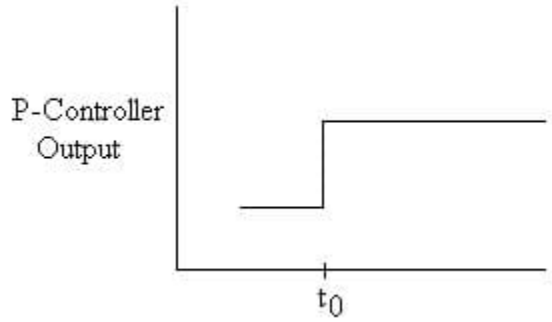


Figure 2. P-controller output for step input.

To illustrate this linear P-control relationship, consider the P-only control that controls the level of a fluid in a tank. Initially, the flow into the tank is equal to the flow out of the tank. However, if the flow out of the tank decreases, the level in the tank will increase because more fluid is entering than is leaving. The P-only control system will adjust the flow out of the tank so that it is again equal to the flow into the tank, and the level will once again be constant. However, this level is no longer equal to the initial level in the tank. The system is at steady-state, but there is a difference between the initial set point and the current position in the tank. This difference is the P-control offset.

Integral (I) Control

Another type of action used in PID controllers is the integral control. Integral control is a second form of feedback control. It is often used because it is able to remove any deviations that may exist. Thus, the system returns to both steady state and its original setting. A negative error will cause the signal to the system to decrease, while a positive error will cause the signal to increase. However, I-only controllers are much slower in their response time than P-only controllers because they are dependent on more parameters. If it is essential to have no offset in the system, then an I-only controller should be used, but it will require a slower response time. This slower response time can be reduced by combining I-only control with another form, such as P or PD control. I-only controls are often used when measured variables need to remain within a very narrow range and require fine-tuning control. I controls affect the system by responding to accumulated past error. The philosophy behind the integral control is that deviations will be affected in proportion to the cumulative sum of their magnitude. The key advantage of adding a I-control

to your controller is that it will eliminate the offset. The disadvantages are that it can destabilize the controller, and there is an integrator windup, which increases the time it takes for the controller to make changes.

I-control correlates the controller output to the integral of the error. The integral of the error is taken with respect to time. It is the total error associated over a specified amount of time. This I-control behavior is mathematically illustrated in Equation 9.2.2 (Scrcek, *et. al*).

$$c(t) = \frac{1}{T_i} \int e(t)dt + c\left(t_0\right) \tag{9.2.2}$$

where

- $c(t)$ is the controller output
- T_i is the integral time
- $e(t)$ is the error
- $c(t_0)$ is the controller output before integration

In Equation 9.2.2, the integral time is the amount of time that it takes for the controller to change its output by a value equal to the error. The controller output before integration is equal to either the initial output at time t=0, or the controller output at the time one step before the measurement. Graphical representations of the effects of these variables on the system is shown in PID Tuning via Classical Methods.

The rate of change in controller output for I-only control is determined by a number of parameters. While the P-only controller was determined by e, the rate of change for I-only depends on both e and T_i. Because of the inverse relationship between c(t) and T_i, this decreases the rate of change for an I-only controller.

The I-only controller operates in essentially the same way as a P-only controller. The inputs are again the set point, the signal, and the bias. Once again, the error is calculated, and this value is sent to the algorithm. However, instead of just using a linear relationship to calculate the response, the algorithm now uses an integral to determine the response that should be taken. Once the integral is evaluated, the response is sent and the system adjusts accordingly. Because of the dependence on T_i, it takes longer for the algorithm to determine the proper response. A graphical representation of the I-controller output for a step increase in input at time t0 is shown below in Figure 3. As expected, this graph represents the area under the step input graph.

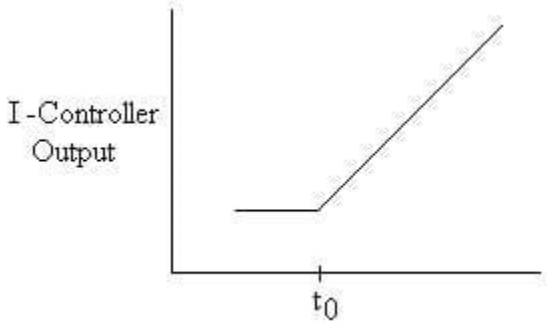


Figure 3. I-controller output for step input.

Derivative (D) Control

Another type of action used in PID controllers is the derivative control. Unlike P-only and I-only controls, D-control is a form of feed forward control. D-control anticipates the process conditions by analyzing the change in error. It functions to minimize the change of error, thus keeping the system at a consistent setting. The primary benefit of D controllers is to resist change in the system, the most important of these being oscillations. The control output is calculated based on the rate of change of the error with time. The larger the rate of the change in error, the more pronounced the controller response will be.

Unlike proportional and integral controllers, derivative controllers do not guide the system to a steady state. Because of this property, D controllers must be coupled with P, I or PI controllers to properly control the system.

D-control correlates the controller output to the derivative of the error. The derivative of the error is taken with respect to time.

It is the change in error associated with change in time. This D-control behavior is mathematically illustrated in Equation 9.2.3 (Scrcek, *et. al*).

$$c(t) = T_d \frac{de}{dt} \tag{9.2.3}$$

where

- $c(t)$ is the controller output
- T_d is the derivative time constant
- de is the differential change in error
- dt is the differential change in time

Graphical representations of the effects of these variables on the system is shown in PID Tuning via Classical Methods.

Mathematically, derivative control is the opposite of integral control. Although I-only controls exist, D-only controls do not exist. D-controls measure only the change in error. D-controls do not know where the setpoint is, so it is usually used in conjunction with another method of control, such as P-only or a PI combination control. D-control is usually used for processes with rapidly changing process outputs. However, like the I-control, the D control is mathematically more complex than the P-control. Since it will take a computer algorithm longer to calculate a derivative or an integral than to simply linearly relate the input and output variables, adding a D-control slows down the controller's response time. A graphical representation of the D-controller output for a step increase in input at time t_0 is shown below in Figure 4. As expected, this graph represents the derivative of the step input graph.

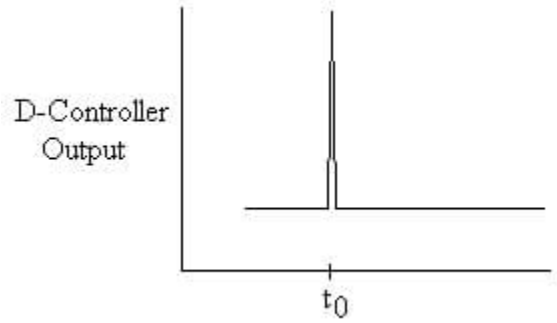
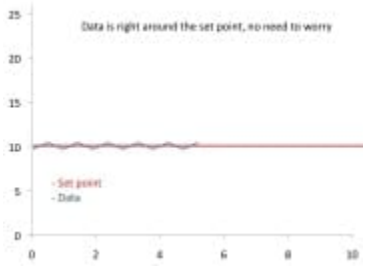


Figure 4. D-controller output for step input.

Controller Effects on a System

The following images are intended to give a visual representation of how P, I, and D controllers will affect a system.



Description

Figure 5. Stable data sample.

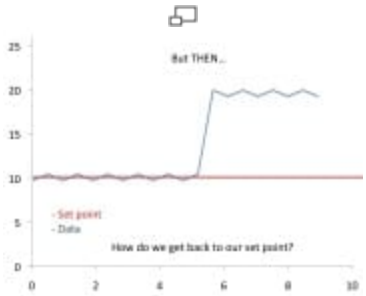


Figure 6. Data disturbance.

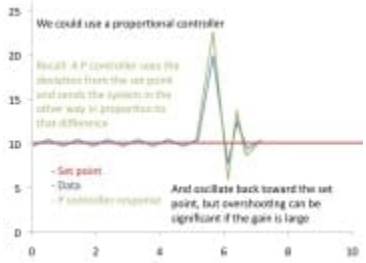


Figure 7. P-controller effect on data.

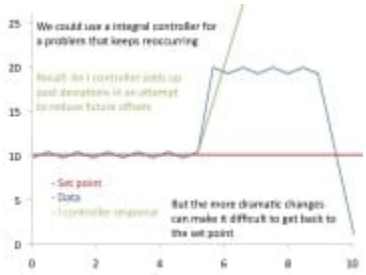


Figure 8. I-controller effect on data.

Figure 9. D-controller effect on data.

Continue reading to see the results of combining controllers.

Proportional-Integral (PI) Control

One combination is the **PI-control**, which lacks the D-control of the PID system. PI control is a form of feedback control. It provides a faster response time than I-only control due to the addition of the proportional action. PI control stops the system from fluctuating, and it is also able to return the system to its set point. Although the response time for PI-control is faster than I-only control, it is still up to 50% slower than P-only control. Therefore, in order to increase response time, PI control is often combined with D-only control.

PI-control correlates the controller output to the error and the integral of the error. This PI-control behavior is mathematically illustrated in Equation 9.2.4 (Scrcek, *et. al*).

$$c(t) = K_c\left(e(t) + \frac{1}{Ti}\int e(t)dt\right) + C \tag{9.2.4}$$

where

- $c(t)$ is the controller output,
- K_c is the controller gain,
- T_i is the integral time,
- $e(t)$ is the error, and
- C is the initial value of controller

In this equation, the integral time is the time required for the I-only portion of the controller to match the control provided by the P-only part of the controller.

The equation indicates that the PI-controller operates like a simplified PID-controller with a zero derivative term. Alternatively, the PI-controller can also be seen as a combination of the P-only and I-only control equations. The bias term in the P-only control is equal to the integral action of the I-only control. The P-only control is only in action when the system is not at the set point. When the system is at the set point, the error is equal to zero, and the first term drops out of the equation. The system is then being controlled only by the I-only portion of the controller. Should the system deviate from the set point again, P-only control will be enacted. A graphical representation of the PI-controller output for a step increase in input at time t0 is shown below in Figure 5. As expected, this graph resembles the qualitative combination of the P-only and I-only graphs.

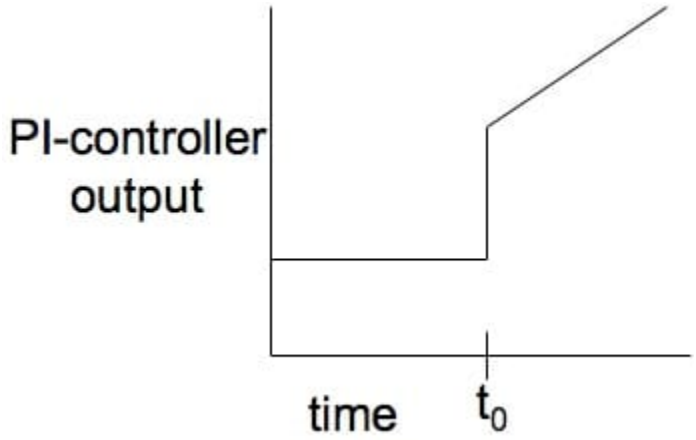


Figure 10. PI-controller output for step input.

Effects of K_c and T_i

With a PI control system, controller activity (aggressiveness) increases as K_c and T_i decreases, however they can act individually on the aggressiveness of a controller's response. Consider Figure 11 below with the center graph being a linear second order system base case.

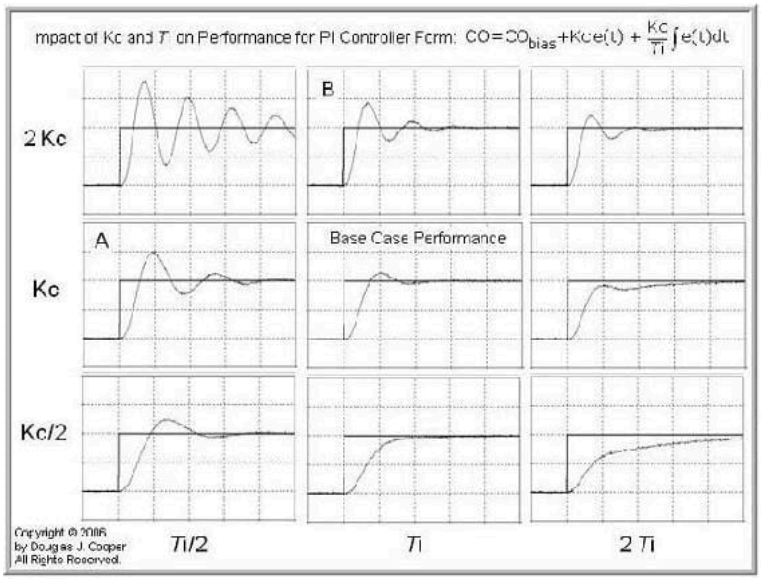


Figure 11. Effects of K_C and T_I [2]

The plot depicts how T_I and K_C both affect the performance of a system, whether they are both affecting it or each one is independently doing so. Regardless of integral time, increasing controller gain (moving form bottom to top on the plot) will increase controller activity. Similarly, decreasing integral time (moving right to left on the plot) will increase controller activity independent of controller gain. As expected, increasing K_C and decreasing T_I would compound sensitivity and create the most aggressive controller scenario.

With only two interacting parameters in PI control systems, similar performance plots can still cause confusion. For example, plots A and B from the figure both look very similar despite different parameters being affected in each of them. This could cause further problems and create a wildly aggressive system if the wrong parameter is being corrected. While trial and error may be feasible for a PI system, it becomes cumbersome in PID where a third parameter is introduced and plots become increasingly similar.

Another noteworthy observation is the plot with a normal K_C and double T_I. The plot depicts how the proportional term is practical but the integral is not receiving enough weight initially, causing the slight oscillation before the integral term can finally catch up and help the system towards the set point.

Proportional-Derivative (PD) Control

Another combination of controls is the PD-control, which lacks the I-control of the PID system. PD-control is combination of feedforward and feedback control, because it operates on both the current process conditions and predicted process conditions. In PD-control, the control output is a linear combination of the error signal and its derivative. PD-control contains the proportional control's damping of the fluctuation and the derivative control's prediction of process error.

As mentioned, PD-control correlates the controller output to the error and the derivative of the error. This PD-control behavior is mathematically illustrated in Equation 5 (Scrcek, *et. al*).

$$c(t) = K_c \left(e(t) + T_d \frac{de}{dt} \right) + C \tag{9.2.5}$$

where

- c(t) = controller output
- K_C = proportional gain
- e = error
- C = initial value of controller

The equation indicates that the PD-controller operates like a simplified PID-controller with a zero integral term. Alternatively, the

PD-controller can also be seen as a combination of the P-only and D-only control equations. In this control, the purpose of the D-only control is to predict the error in order to increase stability of the closed loop system. P-D control is not commonly used because of the lack of the integral term. Without the integral term, the error in steady state operation is not minimized. P-D control is usually used in batch pH control loops, where error in steady state operation does not need to be minimized. In this application, the error is related to the actuating signal both through the proportional and derivative term. A graphical representation of the PD-controller output for a step increase in input at time t0 is shown below in Figure 6. Again, this graph is a combination of the P-only and D-only graphs, as expected.

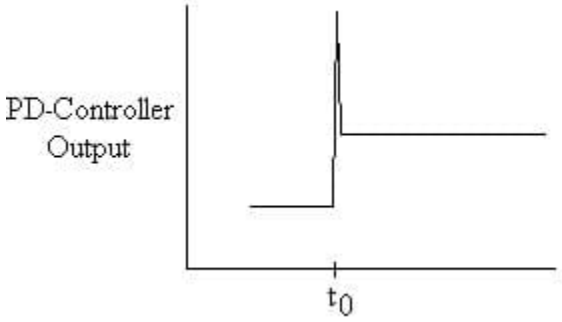


Figure 12. PD-controller output for step input.

Proportional-Integral-Derivative (PID) Control

Proportional-integral-derivative control is a combination of all three types of control methods. PID-control is most commonly used because it combines the advantages of each type of control. This includes a quicker response time because of the P-only control, along with the decreased/zero offset from the combined derivative and integral controllers. This offset was removed by additionally using the I-control. The addition of D-control greatly increases the controller's response when used in combination because it predicts disturbances to the system by measuring the change in error. On the contrary, as mentioned previously, when used individually, it has a slower response time compared to the quicker P-only control. However, although the PID controller seems to be the most adequate controller, it is also the most expensive controller. Therefore, it is not used unless the process requires the accuracy and stability provided by the PID controller.

PID-control correlates the controller output to the error, integral of the error, and derivative of the error. This PID-control behavior is mathematically illustrated in Equation 6 (Scrcek, *et. al*).

$$c(t) = K_c \left(e(t) + \frac{1}{T_i} \int e(t)dt + T_d \frac{de}{dt} \right) + C \tag{9.2.6}$$

where

- c(t) = controller output
- K_c = controller gain
- e(t) = error
- T_i = integral time
- T_d = derivative time constant
- C = intitial value of controller

As shown in the above equation, PID control is the combination of all three types of control. In this equation, the gain is multiplied with the integral and derivative terms, along with the proportional term, because in PID combination control, the gain affects the I and D actions as well. Because of the use of derivative control, PID control cannot be used in processes where there is a lot of noise, since the noise would interfere with the predictive, feedforward aspect. However, PID control is used when the process requires no offset and a fast response time. A graphical representation of the PID-controller output for a step increase in input at time t0 is shown below in Figure 7. This graph resembles the qualitative combination of the P-only, I-only, and D-only graphs.

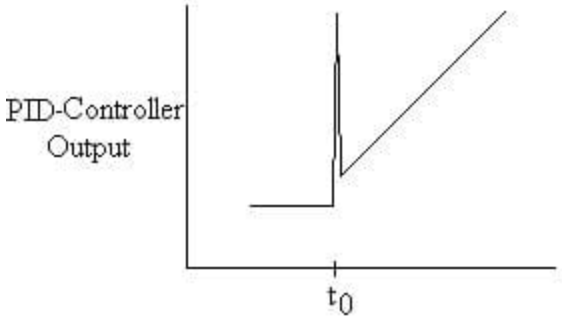


Figure 7. PID-controller output for step input.

In addition to PID-control, the P-, I-, and D- controls can be combined in other ways. These alternative combinations are simplifications of the PID-control.

Note: Order of e(t)

The order of the elements in the e(t) can vary depending on the situation. It could be the fixed element minus the varying element or the other way around. To better illustrate the concept let's go to an example. Let's say you are creating a PID control to control the fluid level in a tank by manipulating the outlet valve. When the fluid level in the tank exceeds your set value, you will want the valve to open up more to allow more flow out of the tank. You are looking for a positive response. Therefore your e(t) should give a positive value when the fluid level is higher than the set. In this case your e(t) will be (V-Vset). The same logic can be used for other systems to determine what the e(t) should be in the PID controls.

Modeling PID Controllers With Euler in Excel

As with many engineering systems, PID controllers can be modeled in Excel via numerical methods such as Euler's Method. First begin with the initial value for a given parameter. Determine the change in that parameter at a certain time-step by summing the three controllers P, I, and D at that step, which are found using the equations listed in the P, I, D, PI, PD, PID Control section above. Take this change, multiply it by the chosen time-step and add that to the previous value of the parameter of interest. For more detailed information see Numerical ODE Solving in Excel. An example of a chemical engineering problem that uses this method can be seen in Example 4 below.

Troubleshooting PID Modeling in Excel

When setting up an Excel spreadsheet to model a PID controller, you may receive an error message saying that you have created a circular reference. Say you are controlling the flow rate of one reactant (B) to a reactor which is dependent upon the concentration of another reactant (A) already inside the reactor. Your PID equations look as follows:

$$FB = 1 + K_c \left[(A - A_{\text{set}}) + \frac{1}{\tau_i} \int (A - A_{\text{set}})\, dt + \tau_d \frac{d\,(A - A_{\text{set}})}{dt} \right]$$

and

$$x_i = \frac{d\,(A - A_{\text{set}})}{dt}$$

After you have set up your columns for $A - A_{set}$, $d(A - A_{set})/dt$, x_i, and the cells for your parameters like K_c, τ_i and τ_d, you will need to set up your PID column with your PID equation in it. After entering your equation into the first cell of the PID column, you may receive the Circular Reference error message when trying to drag the equation down to the other cells in your column.

There are two things you can do:

1. It is likely that you need to start your PID equation in the second or third cell of your PID column. Enter reasonable values into the first couple of cells before you start the PID equation, and you will find that these values shouldn't affect the end

result of your controller.
2. You can also try decreasing the step size (Δt).

Summary Tables

A summary of the advantages and disadvantages of the three controls is shown below is shown in Table 1.

Table 1. Advantages and disadvantages of controls

	Proportional (P)	Integral (I)	Derivative (D)
Advantages	Fast response time – Minimizes fluctuation	+ Contains small offset + Returns system to steady state	+ Keeps system at consistent setting + Controls processes with rapidly changing outputs
Disadvantages	Contains large offset - Does not bring system to desired set point	- Slow response time	Slow response time - Requires combined use with another controller

A guide for the typical uses of the various controllers is shown below in Table 2.

Table 2. Typical uses of P, I, D, PI, and PID controllers

Controller	Estimates	When to use	Examples
P	Present	Systems with slow response, systems tolerant to offset	Float valves, thermostats, humidistat
I	Back	Not often used alone, as is too slow	Used for very noisy systems
D	Forward	Not used alone because it is too sensitive to noise and does not have set point	None
PI	Present & back	Often used	Thermostats, flow control, pressure control
PID	All time	Often used, most robust, but can be noise sensitive	Cases where the system has intertia that could get out of hand: i.e. temperature and concentration measurements on a reactor to avoid runaway.

A summary of definitions of the terms and symbols are shown below in Table 3.

Table 3. Definitions of terms and symbols.

Term or Symbol	Definition
Bias	Controller output when the error is equal to zero
Change in Error, Δe	Differential change in error used in the Derivative Action equation
Change in Time, Δt	Differential change in time used in the Derivative Action equation
Controller	Instrument used to adjust a system properly such as temperature, pressure, and level. Controllers vary in three main aspects: (1) The way they correlate error and actuating signal (such as PID controllers) (2) The way they work relative to the data they gather (feedback Vs. feedforward) (3) The way the adjust the system setting (on-off Vs. continuous)
Controller Gain, K_c	Ratio of the change in output to the change in input
Controller Output, $c(t)$	Voltage output from the controller, usually relating to the actuating signal
Controller Output before Integration, $c'(t)$	Constant used in the Integral Action equation
Derivative Time Constant, T_d	Constant used in the Derivative Action equation
Error, $e(t)$	The difference between the measured signal and the set point
Offset	Systematic error that prevents the system from being at the desired set point during steady state
Integral Time, T_i	The time it takes the controller to change the output by a value equal to the error

Example 9.2.1

Hypothetical Industries has just put you in charge of one of their batch reactors. Your task is to figure out a way to maintain a setpoint level inside of the reactor. Your boss wants to use some type regulator controller, but he is not quite sure which to use. Help your boss find the right type of controller. It is extremely important that the level inside the reactor is at the setpoint. Large fluctuation and error cannot be tolerated.

Solution

You would want to use a PID controller. Because of the action of P control, the system will respond to a change very quickly. Due to the action of I control, the system is able to be returned to the setpoint value. Finally, because it is so critical for the system to remain at a constant setpoint, D control will measure the change in the error, and help to adjust the system accordingly.

Example 9.2.2

You go back to your high school and you notice an oven in your old chemistry class. The oven is used to remove water from solutions. Using your knowledge from ChE 466, you begin to wonder what type of controller the oven uses to maintain its set temperature. You notice some high school students eager to learn, and you decide to share your knowledge with them in hopes of inspiring them to become Chemical Engineers. Explain to them the type of controller that is most likely located within the oven, and how that controller works.

Solution

Since the oven is only used to remove water from a solution, fluctuation, error, and lag between the set point and the actual temperature are all acceptable. Therefore, the easiest and simplest controller to use would be the On-Off controller. The On-Off controller turns on the heating mechanism when the temperature in the oven is below the setpoint temperature. If the temperature of the oven increases above the set temperature, the controller will turn the heating mechanism off.

Example 9.2.3

Having taken your advice, your boss at Hypothetical Industries decides to install a PID controller to control the level in the batch reactor. When you first start up the reactor, the controller initially received a step input. As the reactor achieves steady state, the level in the reactor tends to fluctuate, sending pulse inputs into the controller. For a pulse input, provide a grahical representation of the PID controller output.

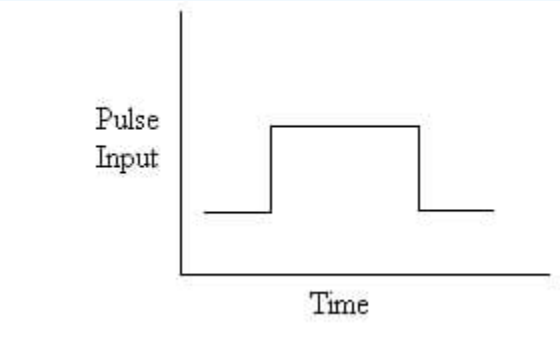


Figure 8. Pulse input.

Solution

The PID-controller output will be a combination of the P-only, I-only and D-only controller outputs. Analogous to the P-controller output for the step input, the P-controller output for the pulse input will exactly resemble the input.

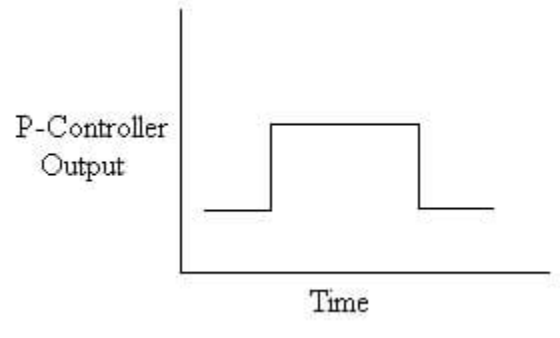


Figure 9. P-controller output for pulse input.

The I-controller output represents the area under the input graph. Unlike the step input, the area under the pulse input graph dropped back down to zero once the pulse has passed. Therefore, rather than continually increase, the I-controller output graph will level off in the end.

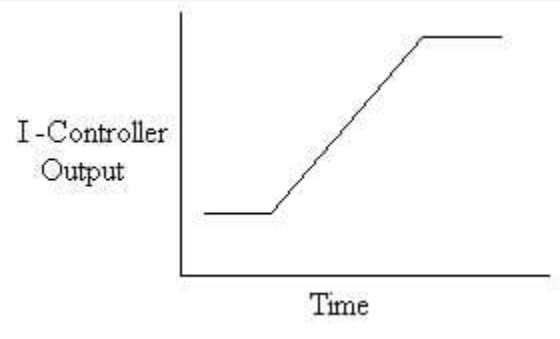


Figure 10. I-controller output for pulse input.

The D-controller output represents the derivative of the input graph. The derivative at the first discontinuity of the graph would be positive infinity. The derivative of the second downward discontinuity is negative infinity.

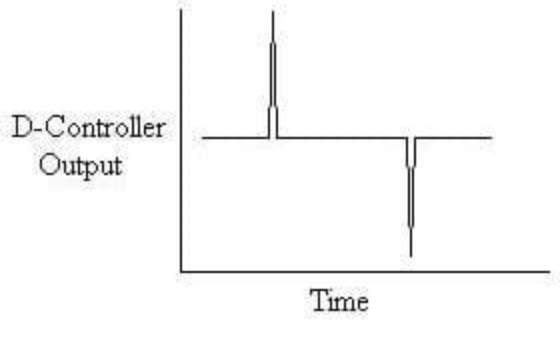


Figure 11. D-controller output for pulse input.

Combining the qualitative characteristics of all three graphs we can determine the PID-controller output for a pulse input.

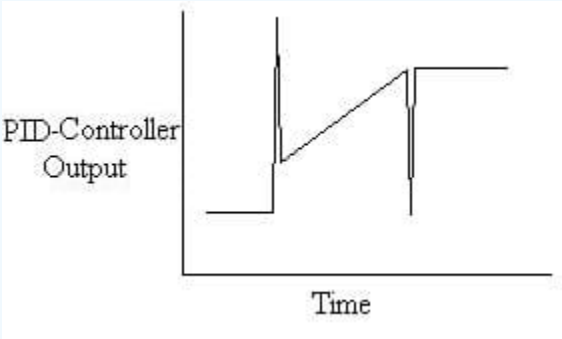


Figure 12. PID-controller output for pulse input.

Example 9.2.3

Different kinds of disturbances are possible when determining the PID controller robustness. These different disturbances are used to simulate changes that might occur within your system. For a CSTR reactor, you decide to heat up your system to account for the cold outside weather. The disturbance in the input temperature is a ramp disturbance, as shown in figure #. If the controller responds to the input temperature, what will be the PID controller output?

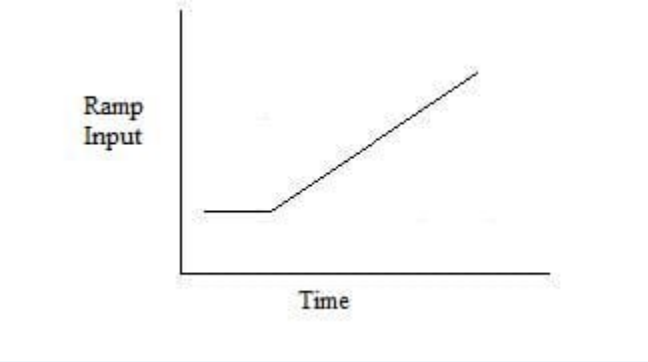


Figure 13. Ramp input.

Solution

Using a controller with a p-only controller only, we will see a proportional change in the controller output corresponding to the input variable change. See figure 14 below

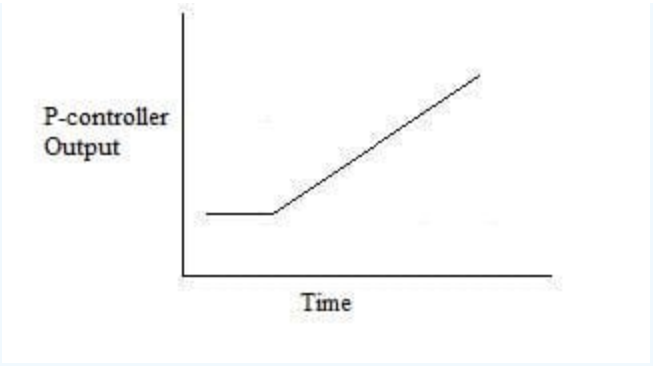


Figure 14. P-controller output for ramp input.

Using an I-only controller, we will see the controller corresponding to the area under the graph, which in this case, seem to increase exponentially with the ramp geometry.

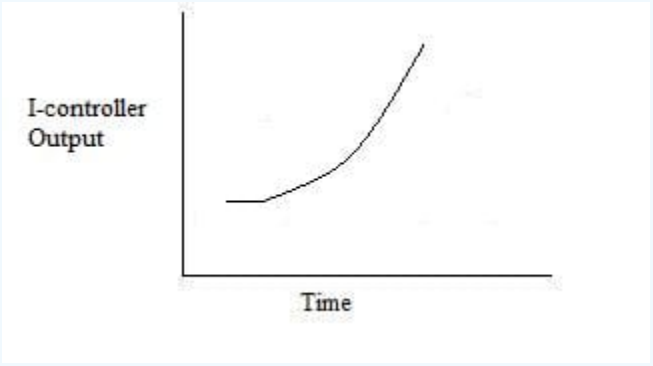


Figure 15. I-controller output for ramp input.

Using a D-only controller, we will see a step response to the ramp disturbance. This is because the D-component corresponds to the derivative, and a ramp input shows a constant slope (positive in this case) which is different than the starting condition slope (zero usually). See figure 16.

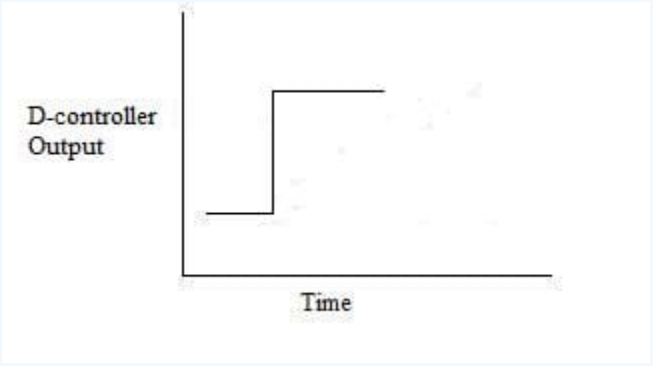


Figure 16. D-controller output for ramp input.

Using a PID controller, the three components all come to play in the controller output. As we would expect, the result will be just a simple addition of the three separate component graphs.

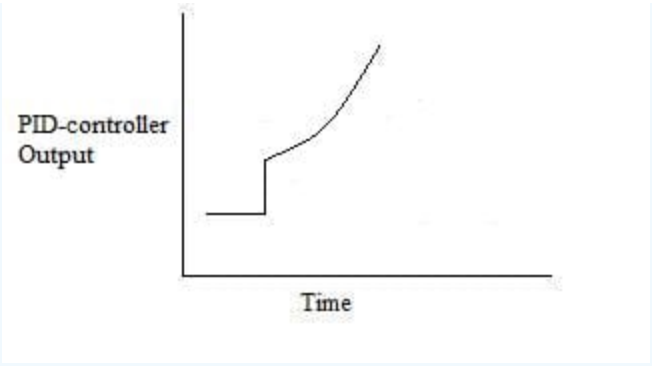


Figure 17. D-controller output for ramp input.

Example 9.2.6

Following is a P&ID of the process A+B-->C.

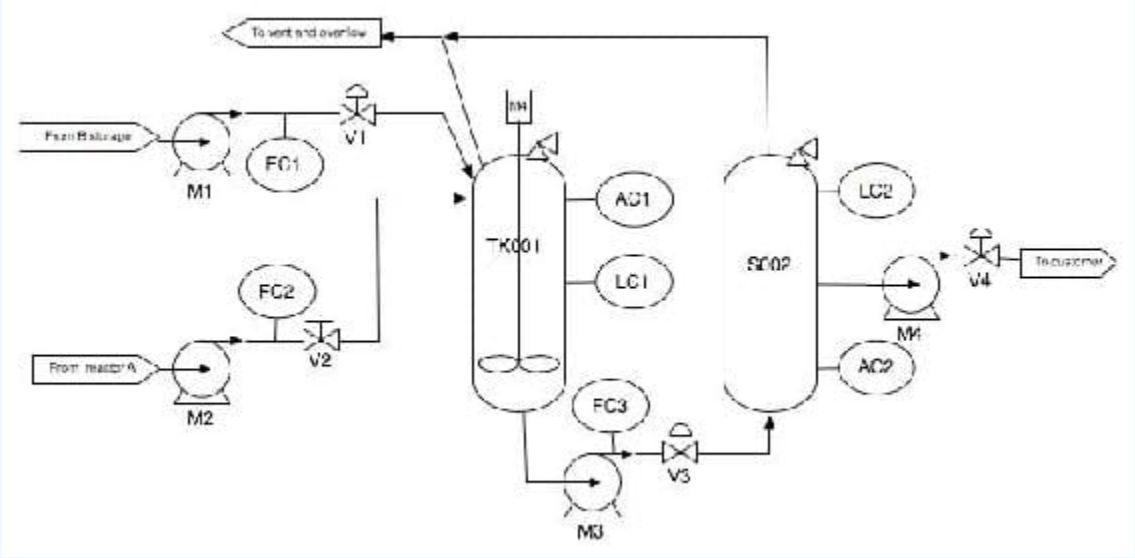


Figure 18. P&ID for a reaction process.

What is the PID controller expression on V3 controlling the volume in TK001 to a setpoint of 50 liters? Note: The PID controller uses LC1 to measure the volume.

Solution

The general equation for a PID controller is:

$$c(t) = K_c \left[e(t) + \frac{1}{T_i} \int e(t)dt + T_d \frac{de}{dt} \right] + C$$

where

- $c(t)$ = controller output
- K_c = controller gain
- $e(t)$ = error
- T_i = integral time
- T_d = derivative time constant
- C = initial value of controller

Therefore, for this example, the solution is:

$$F_{\text{out}} = offset + K_c\left[(V_1 - V_{\text{set}}) + \frac{1}{T_i}\int (V_1 - V_{\text{set}})\,dt + T_d\frac{d\,(V_1 - V_{\text{vet}})}{dt}\right].$$

Example 9.2.6

In this problem, the differential equations describing a particular first-order system with a PID controller will be determined. This is designed to show how well-defined systems can be modeled or explained mathematically. As a supplement to this problem, visit Constructing Block Diagrams. Note that this is an example of solution using integro-differential operators rather than Laplace transforms. Here we have set $K_p = 1$.

Consider a general first-order process:

$$\tau pY'(t) + Y(t) = X(t)$$

Where *Y*(*t*) is the output of the system and *X*(*t*) is the input. Add a PID controller to the system and solve for a single, simple differential equation. The operator or equation for a PID controller is below. Assume there is no dead time in the measurement.

$$G = K_c\left[1 + \frac{1}{\tau_I}\int_0^t d\tau + \tau_D\frac{d}{dt}\right]$$

Solution: Use Constructing Block Diagrams as a reference when solving this problem.
Equations defining system,

- Process: $\tau_p Y'(t) + Y(t) = X(t)$
- Controller: $X(t) = G\varepsilon(t)$
- Comparator: $\varepsilon(t) = R(t) - M(t)$
- Measurement: $M(t) = Y(t)$

When these equations are combined into one equation, the following equation results. This is achieved by adding the measurement to the comparator to the controller to the process equation.

$$\tau_p Y'(t) + Y(t) = G(R(t) - Y(t))$$

Substituting the controller operator and then evaluating yields:

$$\tau_p Y'(t) + Y(t) = K_c\left[1 + \frac{1}{\tau_I}\int_0^t d\tau + \tau_D\frac{d}{dt}(R(t) - Y(t))\right]$$

$$\tau_p Y'(t) + Y(t) = K_c\left[R(t) + \frac{1}{\tau_I}\int_0^t R(t)dt + \tau_D\frac{d}{dt}R(t) - \left[Y(t) + \frac{1}{\tau_I}\int_0^t Y(t)dt + \tau_D\frac{d}{dt}Y(t)\right]\right]$$

Because there is an integral in the differential equation, it is necessary to take the derivative with respect to time.

$$\tau_p Y''(t) + Y'(t) = K_c\left[R'(t) + \frac{1}{\tau_I}R(t) - \tau_D R''(t) - \left[Y'(t) + \frac{1}{\tau_I}Y(t) - \tau_D Y''(t)\right]\right]$$

$$\tau_p Y''(t) + Y'(t) = K_c(R'(t) - Y'(t)) + \frac{K_c}{\tau_I}(R(t) - Y(t)) + K_c\tau_D(R''(t) - Y''(t))$$

To put this in standard notation for solving a second order differential equation, the $Y(t)$ need to be on one side, and the $R(t)$ terms need to be on the opposite side. Also, the coefficient of the $Y(t)$ term needs to be one.

$$\frac{\tau_I\left(\tau_p - K_c\tau_D\right)}{K_c}Y''(t) + \frac{\tau_I\left(1 + K_c\right)}{K_c}Y'(t) - Y(t) = \tau_I\tau_D R''(t) + \tau_I R'(t) + R(t)$$

The above equation can then be solved by hand or using a program such as Mathematica. If using a computer program, different values for the control parameters K_C, τ_I, τ_D can be chosen, and the response to a change in the system can be evaluated graphically.

Exercise 9.2.1

What type of controller is displayed by the equation below?

$$c(t) = K_c\left[e(t) + \frac{1}{T_i}\int e(t)dt\right]$$

a. Feedforward
b. PID
c. Derivative
d. Proportional Integral

Answer available at the end of this section.

Exercise 9.2.2

Which type of controller increases the stability of the system by keeping it at a consistent setting?

a. Derivative
b. Proportional
c. On-Off
d. Integral

Answer available at the end of this section.

Exercise 9.2.3

Which type of controller increases the speed of response to reach the desired set point the fastest while eliminating offset?

a. On-Off
b. Proportional
c. Integral
d. Proportional-Integral

Answer available at the end of this section.

A microbiology laboratory discovered a deadly new strain of bacteria, named P. Woolfi, in the city's water supply. In order to sterilize the water, the bacteria has to be heat killed at a temperature of 105 degrees Celsius. However, this is above the boiling point of water at 1 atm and the bacteria is only susceptible to heat when in liquid due to spore formation when in gas or vapor. To accomplish this sterilization it was suggested that an auto-clave be used to keep the water in the liquid state by keeping it at a pressure of 5 atm while heated for 30 seconds. The auto-clave can only handle up to 7 atm of pressure before exploding, so to ensure that the process is running at the desired specifications, a PID Controller Model must be created in Excel. See figure 18 for a visual of the system.

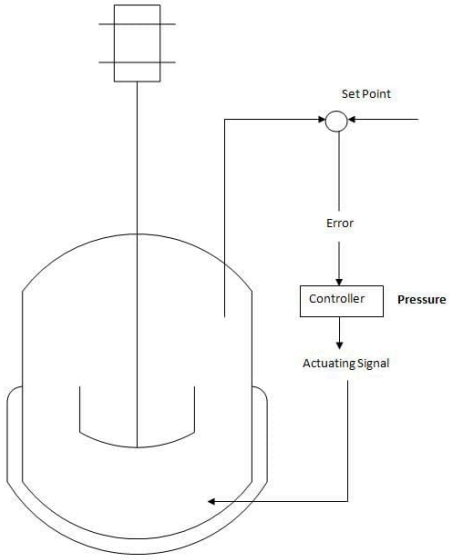


Figure 18.Auto-clave with PID Controls for Temperature and Pressure

Click on this link for the worked out Excel Solution

Explanation:

To simulate the real situation of pressure varying in the system, column B calls an equation to generate a random fluctuation in pressure. Pset is simply the desired specification. Error is the difference between the set pressure and measured signal. du/dt is the sum of the P, I, and D terms. The equations used to calculate each of these can be found in the article, these take into account the error associated with each time-step. dU/dt is the parameter that is varied in order to correct for the difference between the measured pressure and desired pressure.

References

- Astrom, Karl J., Hagglund, Tore., "Advanced PID Control", The Instrumentation, Systems and Automation Society.
- Cooper, Douglas J. "Practical Process Control E-Textbook " http://www.controlguru.com
- Scrcek, William Y., Mahoney, Donald P., Young, Brent R. "A Real Time Approach to Process Control", 2nd Edition. John Wiley & Sons, Ltd.

Answers

Exercise 9.2.1

d

Exercise 9.2.2

a

Exercise 9.2.3

d

This page titled 9.2: P, I, D, PI, PD, and PID control is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Ardemis Boghossian, James Brown, & Sara Zak via source content that was edited to the style and standards of the LibreTexts platform.

9.3: PID Tuning via Classical Methods

Introduction

Currently, more than half of the controllers used in industry are PID controllers. In the past, many of these controllers were analog; however, many of today's controllers use digital signals and computers. When a mathematical model of a system is available, the parameters of the controller can be explicitly determined. However, when a mathematical model is unavailable, the parameters must be determined experimentally. Controller tuning is the process of determining the controller parameters which produce the desired output. Controller tuning allows for optimization of a process and minimizes the error between the variable of the process and its set point.

Types of controller tuning methods include the trial and error method, and process reaction curve methods. The most common classical controller tuning methods are the Ziegler-Nichols and Cohen-Coon methods. These methods are often used when the mathematical model of the system is not available. The Ziegler-Nichols method can be used for both **closed** and **open loop systems**, while Cohen-Coon is typically used for open loop systems. A closed-loop control system is a system which uses feedback control. In an open-loop system, the output is not compared to the input.

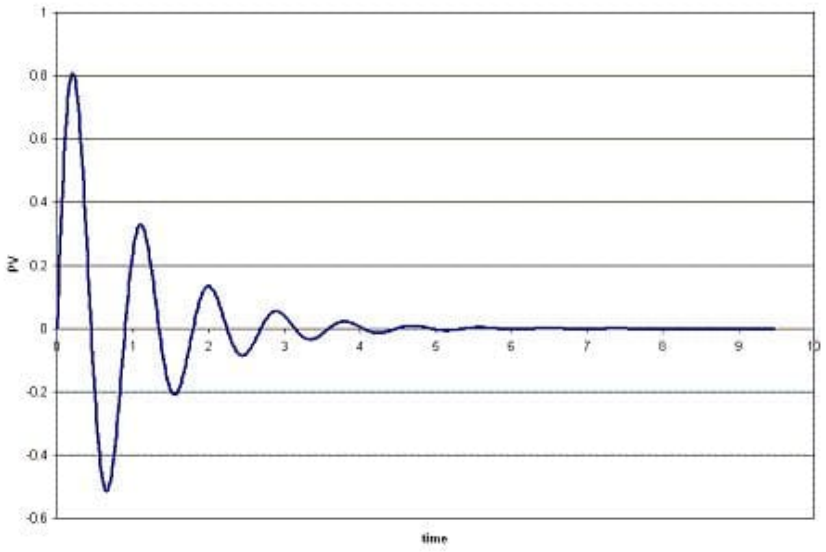
The equation below shows the PID algorithm as discussed in the previous PID Control section.

$$u(t) = K_c \left(\epsilon(t) + \frac{1}{\tau_i} \int_0^t \epsilon\left(t'\right) dt' + \tau_d \frac{d\epsilon(t)}{dt} \right) + b$$

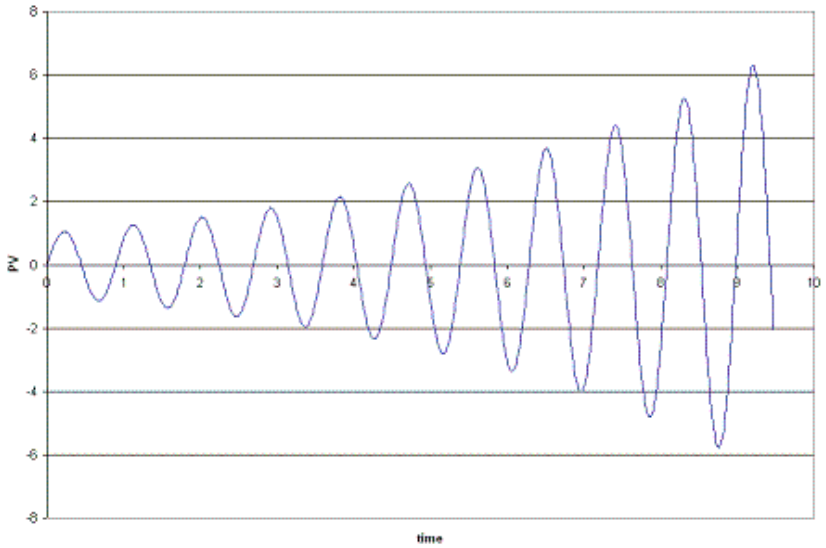
where

- u is the control signal
- ε is the difference between the current value and the set point.
- K_c is the gain for a proportional controller.
- τ_i is the parameter that scales the integral controller.
- τ_d is the parameter that scales the derivative controller.
- t is the time taken for error measurement.
- b is the set point value of the signal, also known as bias or offset.

The experimentally obtained controller **gain** which gives stable and consistent oscillations for closed loop systems, or the **ultimate gain**, is defined as K_u . K_c is the controller gain which has been corrected by the Ziegler-Nichols or Cohen-Coon methods, and can be input into the above equation. K_u is found experimentally by starting from a small value of K_c and adjusting upwards until consistent oscillations are obtained, as shown below. If the gain is too low, the output signal will be damped and attain equilibrium eventually after the disturbance occurs as shown below.



On the other hand, if the gain is too high, the oscillations become unstable and grow larger and larger with time as shown below.



The process reaction curve method section shows the parameters required for open loop system calculations. The Ziegler-Nichols Method section shows how to find K_c, T_i, and T_d for open and closed loop systems, and the Cohen-Coon section shows an alternative way to find K_c, T_i, and T_d.

Open loop systems typically use the quarter decay ratio (QDR) for oscillation dampening. This means that the ratio of the amplitudes of the first overshoot to the second overshoot is 4:1.

Trial and Error

The trial and error tuning method is based on guess-and-check. In this method, the proportional action is the main control, while the integral and derivative actions refine it. The controller gain, K_c, is adjusted with the integral and derivative actions held at a minimum, until a desired output is obtained.

Below are some common values of K_c, T_i, and T_d used in controlling flow, levels, pressure or temperature for trial and error

calculations.

Flow

P or PI control can be used with low controller gain. Use PI control for more accuracy with high integration activity. Derivative control is not considered due to the rapid fluctuations in flow dynamics with lots of noise.

K_c **= 0.4-0.65**

T_i **= 6s**

Level

P or PI control can be used, although PI control is more common due to inaccuracies incurred due to offsets in P-only control. Derivative control is not considered due to the rapid fluctuations in flow dynamics with lots of noise.

The following P only setting is such that the control valve is fully open when the vessel is 75% full and fully closed when 25% full, being half open when 50% filled.

K_c **= 2**

Bias b = 50%

Set point = 50%

For PI control:

K_c **= 2-20**

T_i **= 1-5 min**

Pressure

Tuning here has a large range of possible values of K_c and T_i for use in PI control, depending on if the pressure measurement is in liquid or gas phase.

Liquid
K_c **= 0.5-2**

T_i **= 6-15 s**

Gas
K_c **= 2-10**

T_i **= 2-10 min**

Temperature

Due to the relatively slow response of temperature sensors to dynamic temperature changes, PID controllers are used.

K_c **= 2-10**

T_i **= 2-10 min**

T_d **= 0-5 min**

Process Reaction Curve

In this method, the variables being measured are those of a system that is already in place. A disturbance is introduced into the system and data can then be obtained from this curve. First the system is allowed to reach steady state, and then a disturbance, X_o, is introduced to it. The percentage of disturbance to the system can be introduced by a change in either the set point or process variable. For example, if you have a thermometer in which you can only turn it up or down by 10 degrees, then raising the temperature by 1 degree would be a 10% disturbance to the system. These types of curves are obtained in open loop systems when there is no control of the system, allowing the disturbance to be recorded. The process reaction curve method usually produces a response to a step function change for which several parameters may be measured which include: transportation lag or dead time, τ_{dead}, the time for the response to change, τ, and the ultimate value that the response reaches at steady-state, M_u.

τ_{dead} = transportation lag or dead time: the time taken from the moment the disturbance was introduced to the first sign of change in the output signal

τ = the time for the response to occur

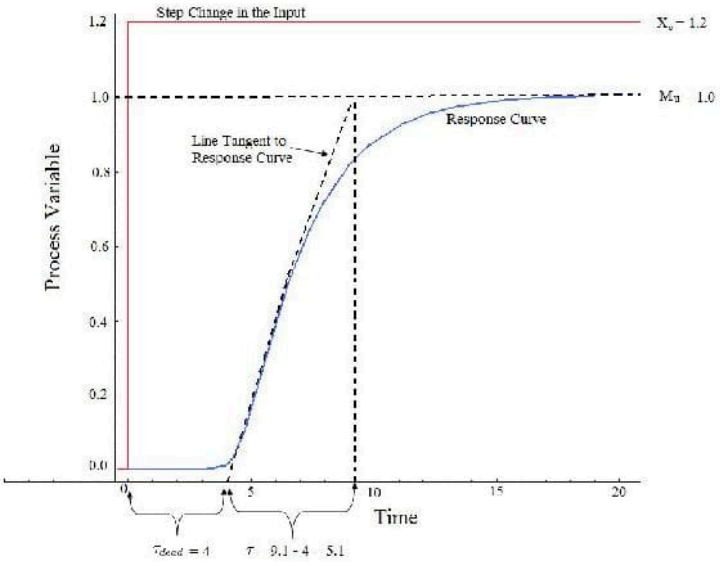
X_o = the size of the step change

M_u = the value that the response goes to as the system returns to steady-state

$$R = \frac{\tau_{\text{dead}}}{\tau}$$

$$K_o = \frac{X_o}{M_u}\frac{\tau}{\tau_{\text{dead}}}$$

An example for determining these parameters for a typical process response curve to a step change is shown below.
In order to find the values for τ_{dead} and τ, a line is drawn at the point of inflection that is tangent to the response curve and then these values are found from the graph.



To map these parameters to P,I, and D control constants, see Table 2 and 3 below in the Z-N and Cohen Coon sections.

Ziegler-Nichols Method

In the 1940's, Ziegler and Nichols devised two empirical methods for obtaining controller parameters. Their methods were used for non-first order plus dead time situations, and involved intense manual calculations. With improved optimization software, most manual methods such as these are no longer used. However, even with computer aids, the following two methods are still employed today, and are considered among the most common:

Ziegler-Nichols closed-loop tuning method

The **Ziegler-Nichols closed-loop tuning method** allows you to use the ultimate gain value, K_u, and the ultimate period of oscillation, P_u, to calculate K_c . It is a simple method of tuning PID controllers and can be refined to give better approximations of the controller. You can obtain the controller constants K_c , T_i , and T_d in a system with feedback. The Ziegler-Nichols closed-loop tuning method is limited to tuning processes that cannot run in an open-loop environment.

Determining the ultimate gain value, K_u, is accomplished by finding the value of the proportional-only gain that causes the control loop to oscillate indefinitely at steady state. This means that the gains from the I and D controller are set to zero so that the influence of P can be determined. It tests the robustness of the K_c value so that it is optimized for the controller. Another important value associated with this proportional-only control tuning method is the ultimate period (P_u). The ultimate period is the time required to complete one full oscillation while the system is at steady state. These two parameters, K_u and P_u, are used to find the loop-tuning constants of the controller (P, PI, or PID). To find the values of these parameters, and to calculate the tuning constants, use the following procedure:

Closed Loop (Feedback Loop)

1. Remove integral and derivative action. Set integral time (T_i) to 999 or its largest value and set the derivative controller (T_d) to zero.
2. Create a small disturbance in the loop by changing the set point. Adjust the proportional, increasing and/or decreasing, the gain until the oscillations have constant amplitude.
3. Record the gain value (K_u) and period of oscillation (P_u).

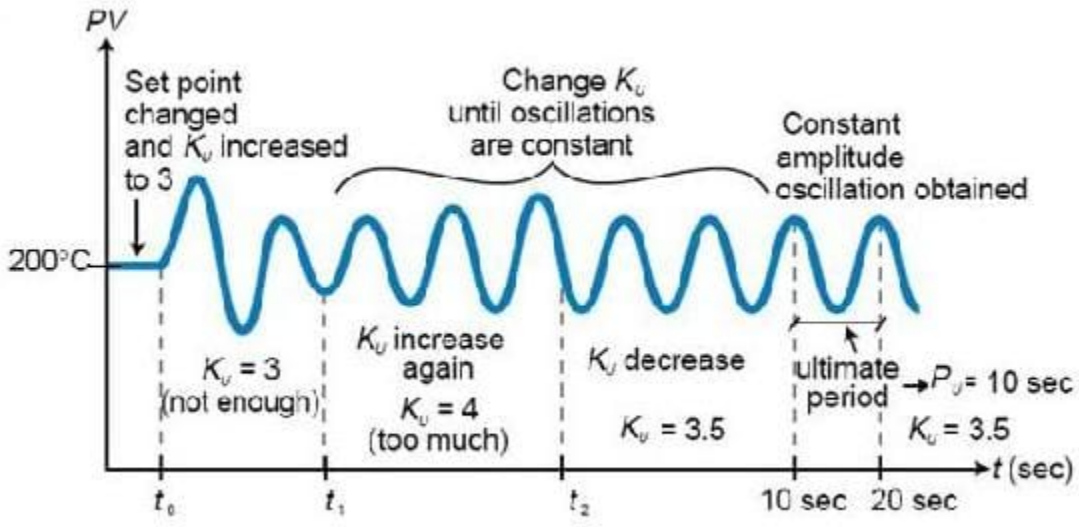


Figure 1. System tuned using the Ziegler-Nichols closed-loop tuning method

4. Plug these values into the Ziegler-Nichols closed loop equations and determine the necessary settings for the controller.

Table 1. Closed-Loop Calculations of K_c, T_i, T_d

	K_c	T_I	T_D
P	$K_u/2$		
PI	$K_u/2.2$	$P_u/1.2$	
PID	$K_u/1.7$	$P_u/2$	$P_u/8$

Advantages

1. Easy experiment; only need to change the P controller

2. Includes dynamics of whole process, which gives a more accurate picture of how the system is behaving

Disadvantages

1. Experiment can be time consuming
2. Can venture into unstable regions while testing the P controller, which could cause the system to become out of control

Ziegler-Nichols Open-Loop Tuning Method or Process Reaction Method:

This method remains a popular technique for tuning controllers that use proportional, integral, and derivative actions. The Ziegler-Nichols open-loop method is also referred to as a process reaction method, because it tests the open-loop reaction of the process to a change in the control variable output. This basic test requires that the response of the system be recorded, preferably by a plotter or computer. Once certain process response values are found, they can be plugged into the Ziegler-Nichols equation with specific multiplier constants for the gains of a controller with either P, PI, or PID actions.

Open Loop (Feed Forward Loop)

To use the Ziegler-Nichols open-loop tuning method, you must perform the following steps:

1. Make an open loop step test

2. From the process reaction curve determine the transportation lag or dead time, τ_{dead}, the time constant or time for the response to change, τ, and the ultimate value that the response reaches at steady-state, M_u, for a step change of Xo.

$$K_o = \frac{X_o}{M_u}\frac{\tau}{\tau_{dead}}$$

3. Determine the loop tuning constants. Plug in the reaction rate and lag time values to the Ziegler-Nichols open-loop tuning equations for the appropriate controller—P, PI, or PID—to calculate the controller constants. Use the table below.

Table 2. Open-Loop Calculations of K_c, T_i, T_d

	K_c	T_i	T_d
P	K_0		
PI	$0.9K_0$	$3.3\tau_{dead}$	
PID	$1.2K_0$	$2\tau_{dead}$	$0.5\tau_{dead}$

Advantages

1. Quick and easier to use than other methods
2. It is a robust and popular method
3. Of these two techniques, the Process Reaction Method is the easiest and least disruptive to implement

Disadvantages

1. It depends upon purely proportional measurement to estimate I and D controllers.
2. Approximations for the K_c, T_i, and T_d values might not be entirely accurate for different systems.
3. It does not hold for I, D and PD controllers

Cohen-Coon Method

The **Cohen-Coon method** of controller tuning corrects the slow, steady-state response given by the Ziegler-Nichols method when there is a large dead time (process delay) relative to the open loop time constant; a large process delay is necessary to make

this method practical because otherwise unreasonably large controller gains will be predicted. This method is only used for first-order models with time delay, due to the fact that the controller does not instantaneously respond to the disturbance (the step disturbance is progressive instead of instantaneous).

The Cohen-Coon method is classified as an 'offline' method for tuning, meaning that a step change can be introduced to the input once it is at steady-state. Then the output can be measured based on the time constant and the time delay and this response can be used to evaluate the initial control parameters.

For the Cohen-Coon method, there are a set of pre-determined settings to get minimum offset and standard decay ratio of 1/4(QDR). A 1/4(QDR) decay ratio refers to a response that has decreasing oscillations in such a manner that the second oscillation will have 1/4 the amplitude of the first oscillation . These settings are shown in Table 3..

Table 3. Standard recommended equations to optimize Cohen Coon predictions

	Kc	Ti	Td
P	(P/NL)*(1+(R/3))		
PI	(P/NL)*(0.9+(R/12))	L*(30+3R) / (9+20R)	
PID	(P/NL)*(1.33+(R/4))	L*(30+3R) / (9+20R)	4L / (11+2R)

where the variables P, N, and L are defined below.

P	Percent Change of Input
N	Percent Change of Output/τ
L	τ_{dead}
R	τ_{dead}/τ

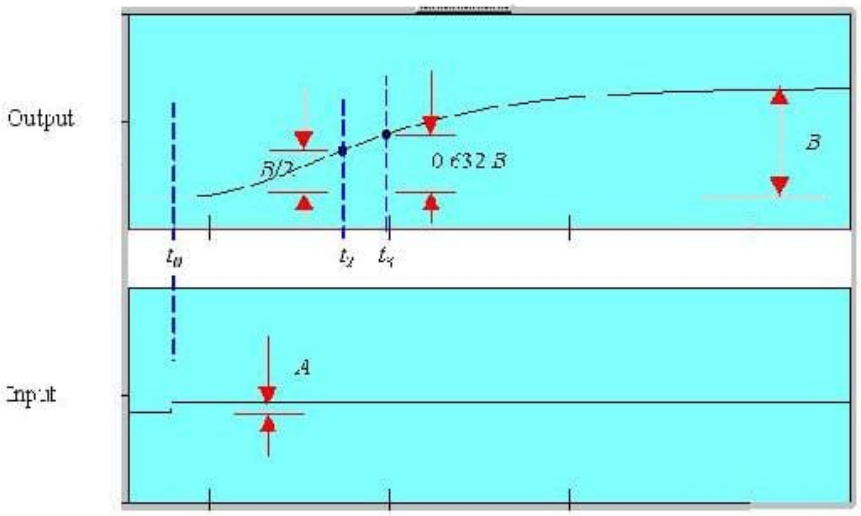
Alternatively, K_0 can be used instead of (P/NL). K_0,τ, and τ_{dead} are defined in process reaction curve section. An example using these parameters is shown here [1].

The process in Cohen-Coon turning method is the following:

1. Wait until the process reaches steady state.
2. Introduce a step change in the input.
3. Based on the output, obtain an approximate first order process with a time constant τ delayed by τ_{dead} units from when the input step was introduced.

The values of τ and τ_{dead} can be obtained by first recording the following time instances:

t0 = time at input step start point t2 = time when reaches half point t3 = time when reaches 63.2% point



4. Using the measurements at t0, t2, t3, A and B, evaluate the process parameters τ, τ_{dead}, and K_o.

5. Find the controller parameters based on τ, τ_{dead}, and K_o.

Advantages

1. Used for systems with time delay.
2. Quicker closed loop response time.

Disadvantages and Limitations

1. Unstable closed loop systems.
2. Can only be used for first order models including large process delays.
3. Offline method.
4. Approximations for the K_c, τ_i, and τ_d values might not be entirely accurate for different systems.

Other Methods

These are other common methods that are used, but they can be complicated and aren't considered classical methods, so they are only briefly discussed.

Internal Model Control

The **Internal Model Control** (IMC) method was developed with robustness in mind. The Ziegler-Nichols open loop and Cohen-Coon methods give large controller gain and short integral time, which isn't conducive to chemical engineering applications. The IMC method relates to closed-loop control and doesn't have overshooting or oscillatory behavior. The IMC methods however are very complicated for systems with first order dead time.

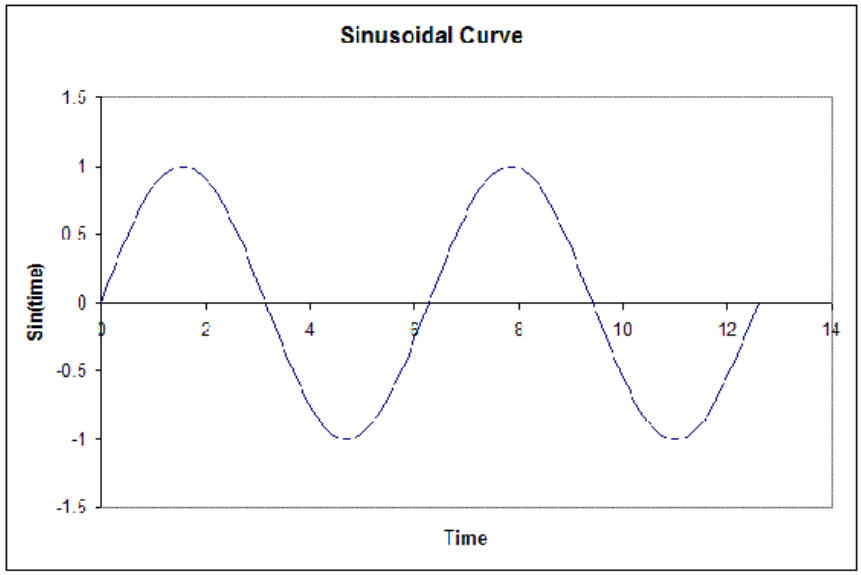
Auto Tune Variation

The **auto-tune variation** (ATV) technique is also a closed loop method and it is used to determine two important system constants (P_u and K_u for example). These values can be determined without disturbing the system and tuning values for PID are obtained from these. The ATV method will only work on systems that have significant dead time or the ultimate period, P_u, will be equal to the sampling period.

Example 9.3.1

You're a controls engineer working for Flawless Design company when your optimal controller breaks down. As a backup, you figure that by using coarse knowledge of a classical method, you may be able to sustain development of the product. After adjusting the gain to one set of data taken from a controller, you find that your ultimate gain is 4.3289.

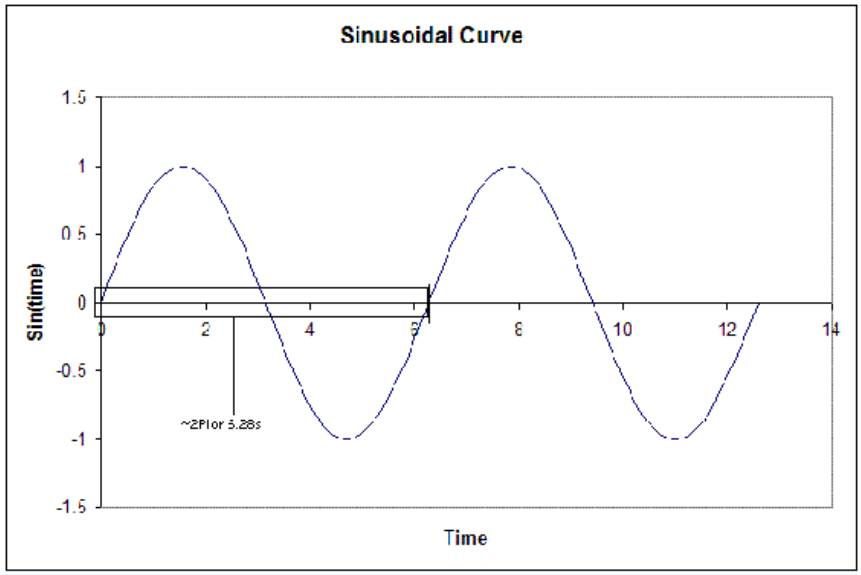
From the adjusted plot below, determine the type of loop this graph represents; then, please calculate K_c, T_i, and T_d for all three types of controllers.



Solution

From the fact that this graph oscillates and is not a step function, we see that this is a closed loop. Thus, the values will be calculated accordingly.

We're given the Ultimate gain, K_u = 4.3289. From the graph below, we see that the ultimate period at this gain is P_u = 6.28

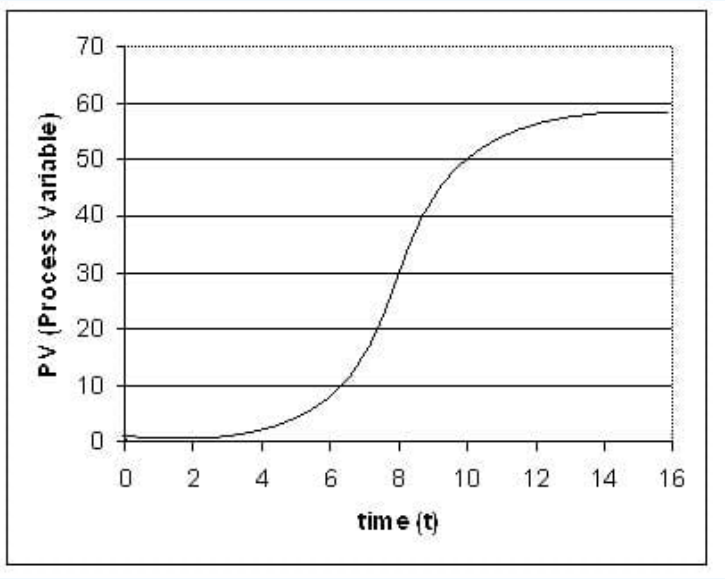


From this, we can calculate the K_c, T_i, and T_d for all three types of controllers. The results are tabulated below. (Results were calculated from the Ziegler-Nichols closed-loop equations.)

				K_c	T_I	T_D
K_U	4.3289		P	2.1645	-	-
P_U	6.28		PI	1.9677	5.2333	-
			PID	2.5464	3.1400	0.7850

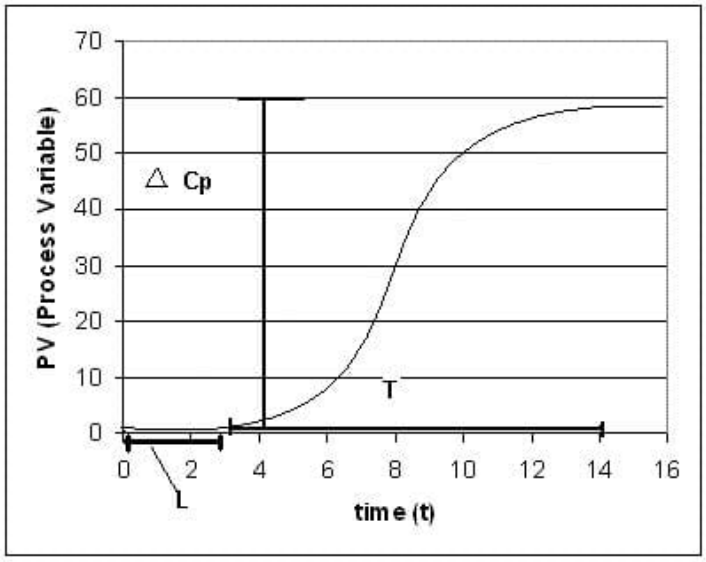
Example 9.3.2

Your partner finds another set of data after the controller breaks down and decides to use the Cohen-Coon method because of the slow response time for the system. They also noticed that the control dial, which goes from 0-8, was set at 3 instead of 1. Luckily the response curve was obtained earlier and is illustrated below. From this data he wanted to calculate K_c, T_i and T_d. Help him to determine these values. Note that the y-axis is percent change in the process variable.



Solution

In order to solve for K_c, T_i and T_d, you must first determine L, ΔC_p, and T. All of these values may be calculated by the reaction curve given.



From the process reaction curve we can find that:
$L = 3$
$T = 11$
\(ΔC_p = 0.55\, (55%)\)

Now that these three values have been found N and R may be calculated using the equations below.

N = $\frac{\Delta C_p}{T}$

R = $\frac{L}{T} = \frac{NL}{\Delta C_p}$

Using these equations you find that
N = .05
R = 0.27
We also know that since the controller was moved from 1 to 3, so a 200% change.
P = 2.00

We use these values to calculate K_c, T_i, and T_d, for the three types of controllers based on the equations found in Table 3.

	Kc	Ti	Td
P	14.53		
PI	12.30	6.42	
PID	18.68	6.65	1.04

Tuning the parameters of a PID controller on an actual system

There are several ways to tune the parameters of a PID controller. They involve the following procedures. For each, name the procedure and explain how the given measured information is used to pick the parameters of the PID controller.

Example 9.3.1

a. The controller is set to P only, and the system is operated in "closed-loop", meaning that the controller is connected and working. The gain is tuned up until a resonance is obtained. The amplitude and frequency of that resonance is measured.
b. The system is kept in "open-loop" mode, and a step-function change is manually made to the system (through a disturbance or through the controller itself). The resulting response of the system is recorded as a function of time.

Solution

a. We will use the Ziegler-Nichols method.

Ki=0.5Ku
Ku is the ultimate gain when the system started oscillating.

b. We will use the Cohen-Coon method.

We can locate the inflection point of the step function and draw a tangent. T_{dead} is located at the crossing of that tangent with t, and T is located at the cross of the tangent with M(t)

Exercise 9.3.1

Which of the following do you RECORD during the Ziegler-Nichols Method?

1. K_c
2. τ_i
3. K_o
4. τ_d

Answer available at the end of this section.

Exercise 9.3.2

For the Ziegler-Nichols Method, it is important to:

1. Find a gain that produces damped oscillation
2. Set P and I controllers to zero
3. Record the period of oscillation
4. Calculate T_c

Answer available at the end of this section.

References

- Svrcek, William Y., Mahoney, Donald P., Young, Brent R. A Real Time Approach to Process Control, 2nd Edition. John Wiley & Sons, Ltd.
- Astrom, Karl J., Hagglund, Tore., Advanced PID Control, ISA, The Instrumentation, Systems and Automation Society.
- "ACT Ziegler-Nichols Tuning," ourworld.compuserve.com/homepages/ACTGMBH/zn.htm
- Ogata, Katsuhiko. System Dynamics, 4th Edition. Pearson Education, Inc.

Answers

Exercise 9.3.1

Answer:C

Exercise 9.3.2

Answer:A,C

This page titled 9.3: PID Tuning via Classical Methods is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Ajay Bhasin, James Bennett, Jamila Grant, & Wen Chung Lim via source content that was edited to the style and standards of the LibreTexts platform.

9.4: PID tuning via Frequency Responses with Bode Plots

Description of Amplitude Ratio, Phase Shift and Frequency Response

A good way to understand the physical meaning of amplitude ratio and phase shift and how they change as a function of the frequency of the input stream is through the analysis of a simple system. A well-insulated **continuously-stirred tank reactor** (CSTR) with an input flow equal to its output flow is a good example (Figure 9.4.1). For this description we consider inlet and outlet temperature, but inlet concentration or feed rate could also be used.

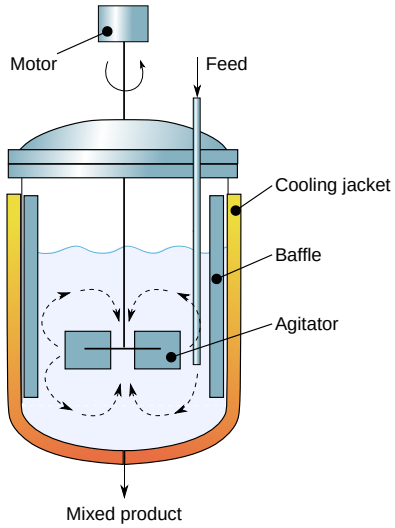


Figure 9.4.1: Schematics of a jacketed agitated CSTR. Agitator geometry is a "Rushton type" (the most common in chemical industry), but many other impeller geometries can be used. (CC BY-SA 3.0; Daniele Pugliesi via Wikipedia)

If the CSTR is initially at 55˚C, and the inlet flow has a temperature that fluctuates sinusoidally between 50˚C and 60˚C, the outlet temperature will also vary sinusoidally. If the fluid has a large residence time in the tank, and the frequency of the temperature variation of the input fluid is high relative to the residence time. Then the outlet temperature will also vary quickly in a sinusoidal fashion, but with a significantly smaller amplitude than that of the inlet. The large hold up in the well mixed tank dampens the fluctuations in the inlet temperature. In contrast, if the the frequency of the inlet temperature sine wave is low relative to the fluid residence time, the output will also be a sine wave with temperatures ranging from 50˚C to 60˚C.

In both cases described above there will be a lag or phase shift between the input and the output of the system (see the figure immediately below). In our example, the phase shift is controlled by the average residence time of fluid in the tank. The time constant of this system is the average residence time of the tank. A system with a small residence time will respond very quickly to changes in input flow rate, and large residence times will result in sizable lag time.

As frequency of the temperature variation increases, the period of the input decreases and the phase shift becomes larger. This can be seen in Figure 9.4.2:

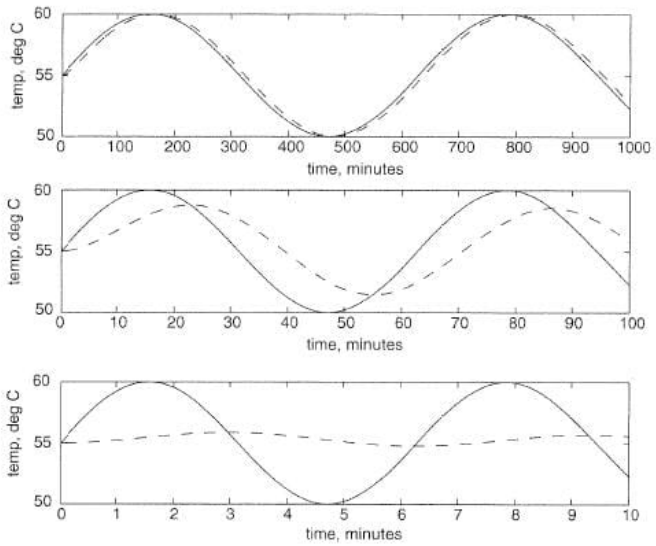


Figure 9.4.2: Solid line = input, dashed line = output (graphs adapted from Bequette)

Imagine that we have a sensitive reaction that must be maintained with 3˚C of 55˚C, it now becomes of utmost importance that we understand the frequency response of this CSTR system. Here in lies the utility of Bode plots; they depict a range of different frequency responses in two plots. This allows a relatively rapid determination of system robustness. In addition, frequency response and Bode plot analysis can be used to tune PID control systems.

Frequency Response

For a given process described by $Y(t) = \hat{G}X(t)$, one considers a sinusoidal input

$$X(t) = \sin(\omega t)$$

After a long time, the solution will also be sinusoidal, with

$$Y(t) \sim AR\sin(\omega t + \phi)$$

where AR is the amplitude ratio (must be a positive number) and $\phi$is the phase lag ($\phi < 0$) or phase lead ($\phi > 0$).

Sometimes, to solve for $Y(t)$ using an ODE solver, it is useful to have the process equation in the form $X(t) = \hat{G}^{-1}Y(t)$ and define $\hat{G}^{-1}$ instead of $\hat{G}$.

Expressions for AR and ϕ

First-Order System

$$AR = \frac{1}{\sqrt{1+\omega^2\tau_p^2}}$$

$$\phi = \tan^{-1}(-\omega\tau_p)$$

Second-Order System

$$AR = \frac{1}{\sqrt{\left(1-\omega^2\tau_p^2\right)^2 + \left(2\zeta\omega\tau_p\right)^2}}$$

$$\phi = \tan^{-1}\left(\frac{-2\zeta_p\omega\tau_p}{1-\omega^2\tau_p^2}\right)$$

Dead Time

$$AR = 1$$

\[A R=1]

$$\phi = -\omega\tau_{\text{dead}}$$

Systems in series

$$AR = AR_1 \cdot AR_2 \cdot ...$$

$$\log AR = \log AR_1 + \log AR_2 + ...$$

$$\phi = \phi_1 + \phi_2 + ...$$

P Controller

$$\hat{G}_c = K_c$$

$$AR = K_c$$

$$\phi = 0$$

PI Controller

$$\hat{G}_c = K_c\left(1 + \frac{1}{\tau_I}\int_0^t dt'\right)$$

$$\phi = \tan^{-1}\left(-\frac{1}{\omega\tau_I}\right)$$

PD Controller

$$AR = K_c\sqrt{1+\omega^2\tau_D^2}$$

$$\phi = \tan^{-1}\left(\omega\tau_D\right)$$

PID Controller

$$\phi = \tan^{-1}\left(\omega\tau_D - \frac{1}{\omega\tau_I}\right)$$

Introduction to Bode Plots

Because it's impossible to perfectly model a real chemical process, control engineers are interested in characterizing the

robustness of the system – that is, they want to tune controllers to withstand a reasonable range of change in process parameters while maintaining a stable feedback system. Bode plots provide an effective means to quantify the system's stability. Bode plots describe an open or closed-loop system as a function of input frequency and give a picture of the system's stability. If all inputs into a system were constant, it would be a relatively simple task to control the system and its output. However, if an input value, such as temperature, varies sinusoidally, then the output should also be describable as a sinusoidal function. In this case Bode plots are a useful tool for predicting the response of the system. There are two Bode plots used to describe a system. The first shows amplitude ratio as a function of frequency. The amplitude ratio is the amplitude of the output sine wave divided by the input sine wave. The second plot graphs phase shift as a function of frequency, where phase shift is the time lag between the output and the input sine curves divided by the period of either of the sine curves and then multiplied by -360° (or -2π radians) to express this shift as an angle.

Constructing Bode Plots

Bode plots concisely display all relevant frequency input and output information on two plots: amplitude ratio as a functions of frequency and phase shift as a function of frequency. The amplitude ratio plot is a log-log plot while the phase angle plot is a semilog (or log-linear) plot.

To construct a Bode plot, an engineer would have empirical data showing input and output values that vary as sinusoidal functions of time. For instance, there might be inlet temperature data that varies sinusoidally and the outlet temperature data that also varies sinusoidally. For example, given a heat exchanger that is used to cool a hot stream, we can vary the flow rate of the hot stream and model it sinusodially. (This can be done using ODEs & an Excel model of a heat exchanger.)

Figure $9.4.3$ includes inlet flowrate (process fluid stream) varying as a sinusoidal function and the hot temperature exiting the heat exchanger likewise varying as a sinusoidal function. Portions of this graph are not yet at steady-state.

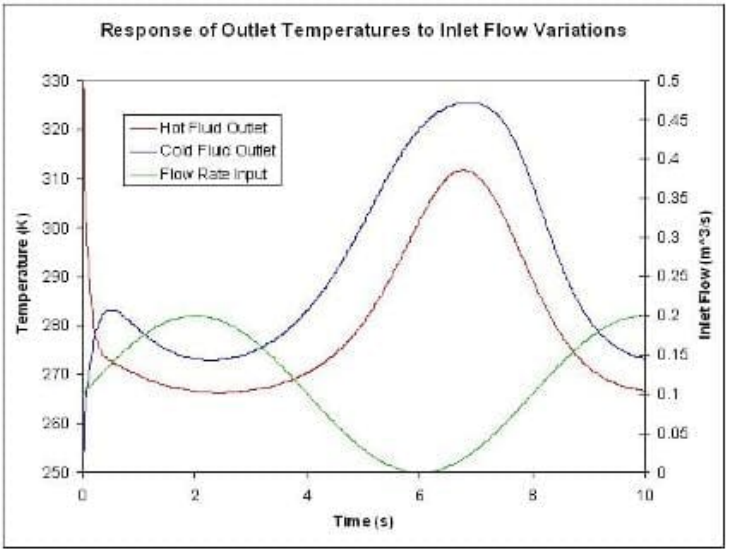


Figure $9.4.3$: Inlet Flowrate modeled by $F_s = 0.1 + 0.1\sin\left(\frac{\pi}{4t}\right)$.

Bode plots are constructed from steady-state data. Figure $9.4.4$ shows part of the steady-state region of the same data used for Figure $9.4.3$.

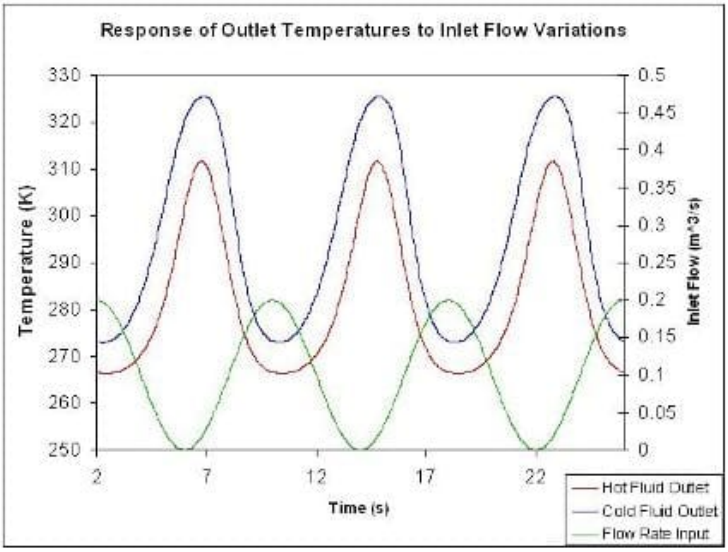


Figure 9.4.4: Steady-state inlet flow rate and outlet temperatures varying as sunusoidal functions.

To collect a single data point for a Bode plot we will use the information from a single period of the inlet flow rate and the corresponding temperature from the hot exiting stream.

The amplitude ratio, AR, is the ratio of the amplitude of the output sinusoidal curve divided by the amplitude of the input sinusoidal curve.

$$AR = \frac{\text{output amplitude}}{\text{input amplitude}}$$

The value of the amplitude ratio should be unitless so if the units of the input frequency and the units of the output frequency are not the same, the frequency data should first be normalized. For example, if the input frequency is in ˚C/min and the output frequency is also in ˚C/min, then AR is unitless and does not need to be normalized. If, instead, the input frequency is in L/min and the output frequency is in ˚C/min, then AR = ˚C/liters. In this case the inlet and outlet frequencies need to be normalized because the ratio AR = ˚C/L doesn't say anything about the physical meaning of the system. The value of AR would be completely different if the units of AR were Kelvin/gal.

To find the phase shift, the periods of the input and output sine curves need to be found. Recall that the period, P, is the length of time from one peak to the next.

$$P = \frac{1}{f} = \frac{2\pi}{\omega}$$

with f is the frequency and ω is the angular frequency (rad/sec)

The phase shift is then found by

$$\phi = -2\pi\left(\frac{\Delta P}{P}\right)$$

Using these values found from multiple perturbations in feed flow rate it is possible to construct the following Bode plots:

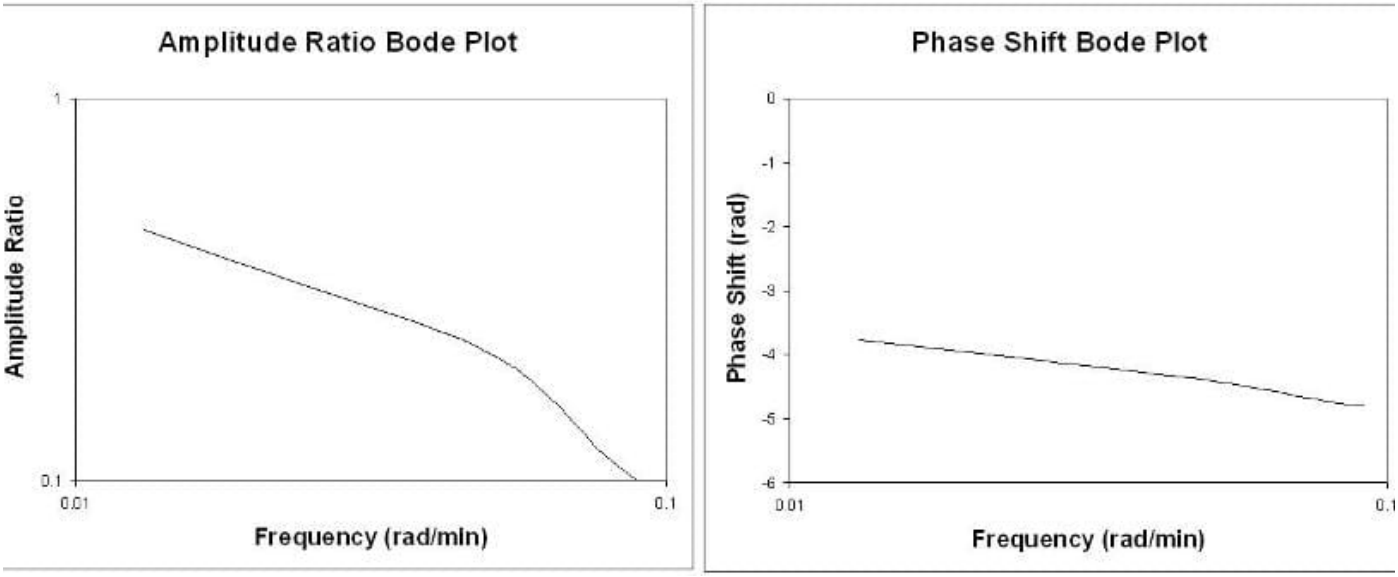


Figure 9.4.5: (left) Bode Plot (Amplitude Ratio) [Log-Log Plot]. (right) Bode Plot (Phase Shift) [Semi-Log]

Rules of Thumb when analyzing Bode plots

Generally speaking, a gain change shifts the amplitude ratio up or down, but does not affect the phase angle. A change in the time delay affects the phase angle, but not the amplitude ratio. For example, an increase in the time delay makes the phase shift more negative for any given frequency. A change in the time constant changes both the amplitude ratio and the phase angle. For example, an increase in the time constant will decrease the amplitude ratio and make the phase lag more negative at any given frequency.

Table 9.4.1: Rules of Thumb

Change	Effect
Gain Change	Amplitude ratio is shifted up or down; no change in phase angle
Time Delay	Phase angle changes, but no change in amplitude ratio

Historical Note

Prior to the advent of powerful computer modeling tools, controls engineers would model systems using transfer functions. Readers interested in learning more about how these were used to construct Bode plots should refer to Bequette's Process Control: Modeling, Design, and Simulation (see References). This wiki assumes that the engineer already has data in Excel, etc, that shows the sinusoidal behavior of input and outputs.

Bode Stability Criterion

A control system is unstable if the open-loop frequency response exhibits an AR exceeding unity at the frequency for which the phase lag is 180 degrees. *This frequency is called the* crossover frequency. *Adjusting the gain so that the AR < 1 for - phi < (180 minus phase margin) results in acceptable stability if the phase margin is greater than 30 degrees.*

As previously mentioned, the controls engineer uses Bode plots to characterize the stability of the system. If you were to find that the amplitude ratio is greater than 1 the system would be unstable at that frequency. An important stability criterion is that AR < 1 when phase shift = -180 degrees (or -pi radians). This is called the Bode stability criterion and if it holds for a closed-loop system, then that system will be stable. If this criterion is not satisfied and you were to put a feedback controller on this process, as done in PID tuning via optimization, the process would explode due to the response lag. The reason for this is that the system might think temperature is too low when it's actually to high, so it would decrease cooling water rather than increase it. Also, with an AR

< 1, an error in input is amplified in the output, so the system will lose control due to amplification of error. This could obviously be a dangerous situation, for example in an exothermic CSTR, which could explode.

Frequency Response in Mathematica

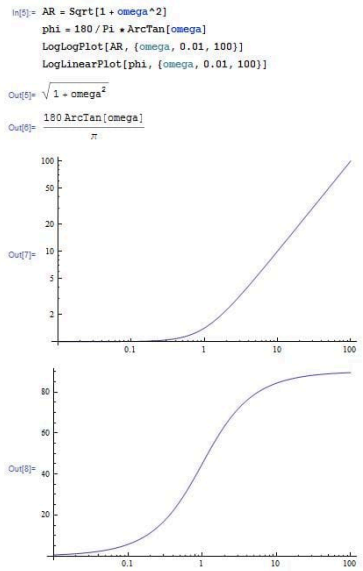
Bode plot for PD controller, $T_D = 1$:

AR = Sqrt[1 + omega^2]

phi = 180/Pi*ArcTan[omega]

LogLogPlot[AR,{omega,0.01,100}]

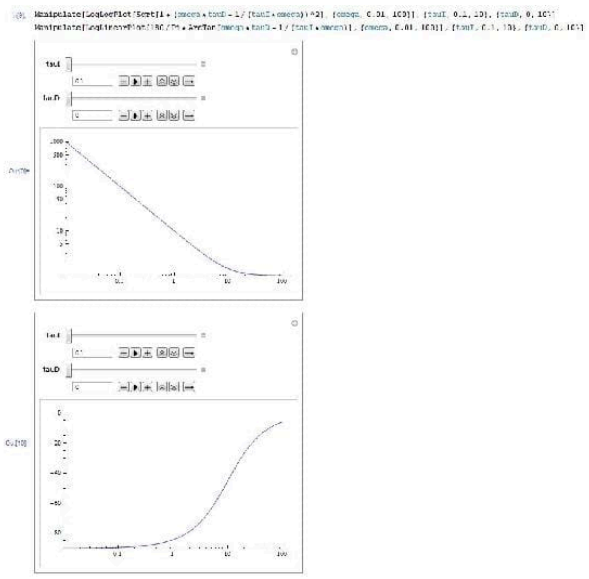
LogLinearPlot[phi,{omega,0.01,100}]



Bode plot for PID controller, using the \Manipulate" function to watch the effect of varying the values of both Tau_I and Tau_D:

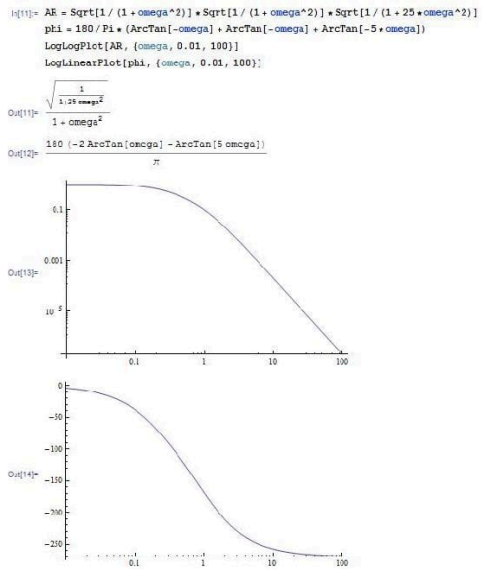
Manipulate[LogLogPlot[Sqrt[1 + (omega*tauD - 1/(tauI*omega))^2],{omega,0.01,100}],{tauI,0.1,10},{tauD,0,10}]

Manipulate[LogLinearPlot[180/Pi*ArcTan[omega*tauD - 1/(tauI*omega)],{omega,0.01,100}],{tauI,0.1,10},{tauD,0,10}]



Bode plot for open-loop operation of third-order system with a P controller:

AR = Sqrt[1/(1+omega^2)]*Sqrt[1/(1+omega^2)]*Sqrt[1/(1+25*omega ^ 2)] phi = 180/Pi*(ArcTan[-omega] + ArcTan[-omega] + ArcTan[-5*omega]) LogLogPlot[AR,{omega,0.01,100}] LogLinearPlot[phi,{omega,0.01,100}]



Find crossover frequency:

FindRoot[phi==-180,{omega,1}]

This yields omega = 1.18322 which is the crossover frequency, and can be used to determine the amplitude ratio (ask mathematica for AR) the gain (Kc=1/AR)

To determine what the gain should be set to at a phase margin of 30 degrees (safe operating conditions) use the following in mathematica:

FindRoot[phi==-150,{omega,1}]

The phase margin of Cohen-Coon and other frequency analysis (such as Zielger-Nicholas) can be calculated by using the gain (Kc) value gotten from the method a typing the following into mathematica:

NSolve[Kc*AR==1,omega]

Derivations

Assume a first-order process with input $X(t)$, output $Y(t)$, and operator

$$\hat{G}^{-1} = \left(\tau\frac{\delta}{\delta t} + 1\right)$$

where

$$X(t) = \hat{G}^{-1}Y(t) = \left(\tau\frac{\delta}{\delta t} + 1\right)Y(t) = \tau Y'(t) + Y(t) \tag{9.4.1}$$

For Bode Plot Analysis, we assume a periodic input to the system:

$$X(t) = \sin(\omega t)$$

Based on this input, we can hypothesize that:

$$Y(t) = AR * \sin(\omega t + \varphi)$$

From this hypothesis, we can replace Equation 9.4.1 with our functions for $X(t)$ and $Y(t)$. This is shown below.

$$X(t) = \tau Y'(t) + Y(t) \tag{9.4.2}$$

$$\sin(\omega t) = \tau\frac{\delta}{\delta t}(AR\sin(\omega t + \phi)) + AR\sin(\omega t + \phi) \tag{9.4.3}$$

$$= AR(\tau\omega\cos(\omega t + \varphi)) + \sin(\omega t + \varphi) \tag{9.4.4}$$

Using trigonometric identities, we can replace $\cos(\omega t + \varphi)$ with $\cos(\omega t)cos(\varphi) - \sin(\omega t)sin(\varphi)$ and $\sin(\omega t + \varphi)$ with $\sin(\omega t)cos(\varphi) + \cos(\omega t)sin(\varphi)$ to get:

$$\sin(\omega t) = AR(\omega T(\cos(\omega t)\cos(\varphi) - \sin(\omega t)\sin(\varphi)) + \sin(\omega t)\cos(\varphi) + \cos(\omega t)\sin(\varphi)) \tag{9.4.5}$$

To solve for AR and φ, we can match the coefficients for the $\sin(\omega t)$ term and $\cos(\omega t)$ term on the left and right side of Equation 9.4.5.

Coefficient of $\cos(\omega t)$:

$$0 = AR(\omega\tau\cos(\varphi) + \sin(\varphi)) \tag{9.4.6}$$

Coefficient of $\sin(\omega t)$:

$$1 = AR(-\omega\tau\sin(\varphi) + \cos(\varphi)) \tag{9.4.7}$$

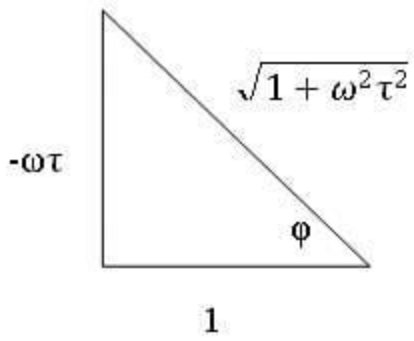
From Equation 9.4.6, we can divide both sides by $AR\cos(\varphi)$ and rearranging to get:

$$\tan(\phi) = -\omega\tau \tag{9.4.8}$$

or

$$\phi = \tan^{-1}(-\omega\tau) \tag{9.4.9}$$

Based on this relationship, we can draw a triangle with the angle φ and the relationship shown in Equation 9.4.9. A picture of the triangle is shown below.



From Equation 9.4.7 and the triangle above, we can solve for AR:

$$AR = \frac{1}{-\omega\tau(\sin(\phi)) + \cos(\phi)} \tag{9.4.10}$$

$$= \frac{1}{-\omega\tau\left(\frac{-\omega\tau}{\sqrt{1+\omega^2\tau^2}}\right) + \frac{1}{\sqrt{1+\omega^2\tau^2}}} \tag{9.4.11}$$

$$= \frac{1}{\sqrt{1+\omega^2\tau^2}} \tag{9.4.12}$$

Determining Stable Controller Gain and Phase Margin

Tuning via phase margin is a more precise method than tuning through **Cohen-Coon Rule** and **Ziegler-Nichols Tuning Rule**. We can find the gain and phase margin by using bode plots or by an analytical method.

Graphical Method

To find a **safe controller gain**, use the following steps:

1. Construct Bode Plot of $\log AR$ versus $\log \omega\tau$ and φ versus $\log \omega\tau$
2. Use the above plot to find $\omega\tau$ such that $\varphi = -180^o + \varphi_{PM}$.
3. Use $\omega\tau$ found from above step to find AR in $\log AR$ versus $\log \omega\tau$ plot.
4. $K_{Csafe} = \frac{1}{AR}$

To find the **phase margin**, use the following steps:

1. Construct Bode Plot of LogAR versus Log ωτ and φ versus Log ωτ
2. Use the above plot to find ωτ such that AR=1.
3. Use ωτ found from above step to find φ in φ versus Log ωτ plot.
4. $\varphi_{PM} = 180 - |\varphi_{AR=1}|$

Analytical Method

Finding the **gain** with a phase margin:

1. Use your expression for Φ, set Φ to the desired value and find the root for ω.
2. Input the value obtained for ω in your expression for AR.
3. $K_C = \frac{1}{AR}$

Finding the phase margin with a gain:

1. Use your expression for AR, set $K_C = \frac{1}{AR}$ and solve for ω.
2. Input the value obtained for ω in your expression for Φ.

Determining the Slope of a Bode Plot

Sometimes it is necessary to determine the relative shape of a Bode Plot without the use of a computer or ODE equation solver. In sketching the relative shape, there are three main parts of a basic Bode Plot that must be considered.

1. Crossing point
2. Starting value
3. Slope

Crossing Point

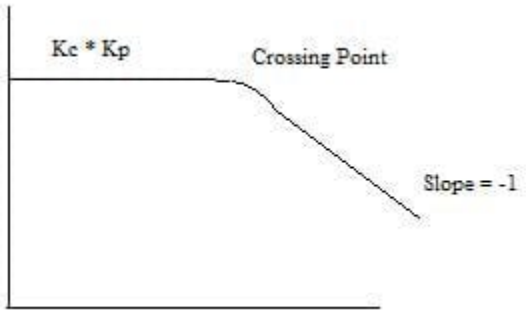
The crossing point of the Bode Plot is when the frequency, ω, is equal to 1 / τ . If the x-axis of the Bode Plot is ωτ, the crossing point will be at x-axis = 1. If there are multiple values of τ, there may be multiple crossing points, each at the frequency where ω is equal to 1 / τ . If there are multiple τ values, and the x-axis is ωτ, it must be specified which τ the x-axis is referring to.

Starting Value

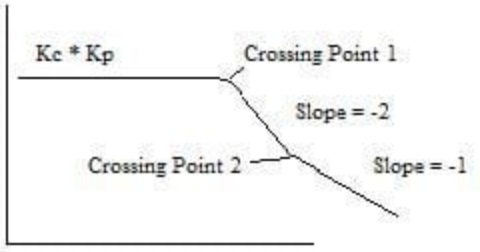
The starting value depends on the K_C and K_P values contributing to the process.

Slope

The slope for an uncontrolled process is equal to the negative of the order of the process. For example, if there is a first-order process with no controller the slope of the Bode Plot would be -1 (-1 for first order), after the crossing point has been reached. A picture of this Bode Plot is shown below.

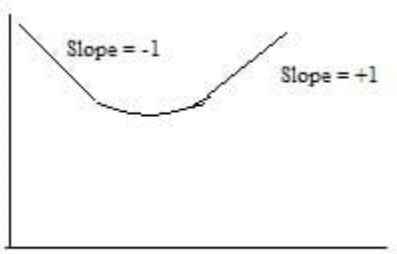


For a process controlled by a PD controller, the AR is essentially the inverse of the first-order system AR, meaning that the slope addition from a PD controller is a +1, instead of a -1 as in a first-order process. For a second-order system with a PD controller the final slope will be -1 (-2 from the second order process, +1 from PD). However, there may be multiple crossing points, whose location depends on the value of τ_P and τ_D. Remember that the crossing point is where ωτ equals one. Since the slope contribution only comes after the crossing point, a Bode Plot for τ_D less than τ_P would look like the picture below.



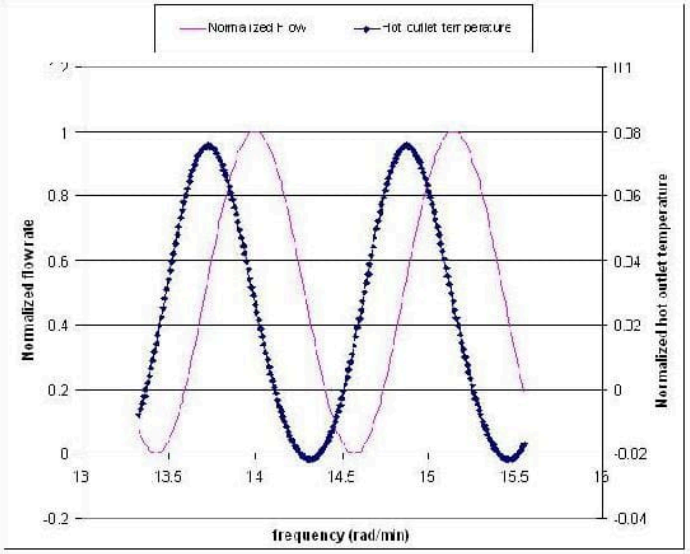
- Crossing point 1 is where $\omega\tau_P$ is 1 and crossing point 2 is where $\omega\tau_D$ is 1.
- Similarly, if τ_D is greater than τ, the slope would first go to +1, then to -1.

For a PI controller, the slope contribution from the controller comes before the crossing point, and then goes to zero after the crossing point. A PID controller would therefore look like the picture shown below, assuming $\tau_D = \tau_I$.



Example 9.4.1: Determine the Amplitude Ratio

You are a controls engineer and wish to characterize a heat exchanger used in a chemical process. One of the many things you are interested in knowing about the system is how the hot outlet temperature responds to fluctuations in the inlet flow rate. Using data for a particular inlet flow rate, you graphed normalized (why?) flow rate and normalized hot outlet temperature vs. frequency (rad/min). Use this graph to determine the amplitude ratio.



Solution

Because flow rate units and temperature do not have the same units, these values needed to be normalized before calculating the amplitude ratio. To normalize flow, use the following equation:

$$\mathrm{F_{norm}} = \frac{F - \mathrm{Fmin}}{\mathrm{Fmax} - \mathrm{Fmin}}$$

where F is the flow rate of a particular data point.

In this problem temperature, T, is in ˚C, so

$$\mathrm{Tnorm} = \frac{T - 273.15}{100}$$

To find the amplitude of both wave functions, first recall that the amplitude of a wave is the maximum height of a wave crest (or depth of a trough). For one steady-state wave produced from a column of values in Excel, you could quickly calculate the amplitude by using the max( ) and min( ) Excel functions. [This can be found using Excel help.] If you opted to find the amplitude this way, then the amplitude for a single wave function would be

$$\mathrm{Amplitude} = \frac{[\max() - \min()]}{2}$$

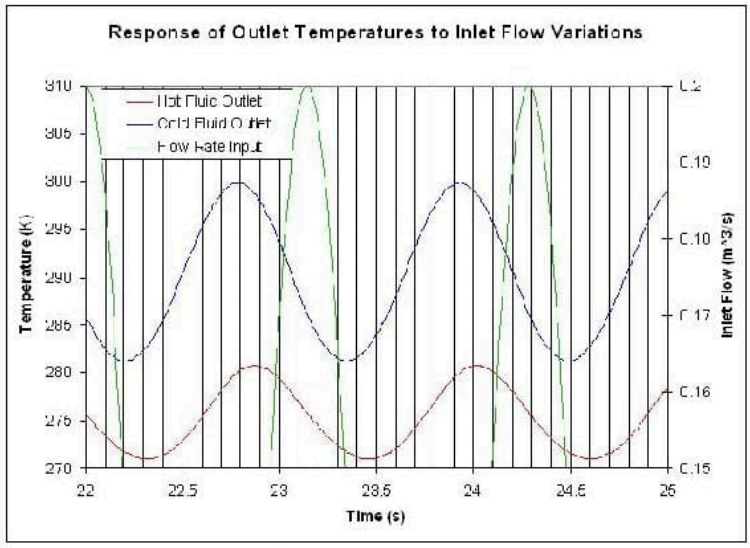
Note that this is just one way to find the amplitude of a wave. It could also be found simply by reading the values off the graph of a wave.

Once the amplitudes of the inlet and outlet waves are found, use the values to find the amplitude ratio, AR, as follows:

$$AR = \frac{\text{outlet streams amplitude}}{\text{inlet streams amplitude}} = \frac{0.0486}{0.499} = 0.0974$$

Example 9.4.1: Determine the Phase Shift

The following graph shows how inlet flow and both the hot and cold outlet temperatures vary as sinusoidal functions of time. This graph was generated using the same data for the heat exchanger of Example 1. Use this graph to find the phase shift between the inlet flow and the hot outlet stream.



Solution

Determine the period (P) – This can be done by finding the time from one peak to the next of a given wave. In this case, we want to know the period of the inlet flow rate, so $P = 1.14\ s$.

Determine the lag (delta P) – This can be done by finding the time difference between the peak of the inlet flow rate and the peak of the hot outlet stream. (Remember that the hot outlet wave lags the wave of the inlet flowrate).

$\delta P = 0.87s$

$$\text{Phaseshift} = \frac{0.87\,s}{1.14\,s} * -2\pi = -4.80$$

Since we're only concerned with time values for finding the phase shift, the data doesn't need to be normalized.

References

- Bequette, B. Wayne. Process Control: Modeling, Design, and Simulation. pp 172-178, 214-235. Upper Saddle River, N.J. Prentice Hall, PTR 2003.
- Liptak, Bela G. Process Control and Optimization. Volume 2. Instrument Engineers Handbook. Fourth Edition. CRC Press, Boca Raton Fl. 2003. *Insertformulahere*

Contributors and Attributions

- written by: Tony Martus, Kegan Lovelace, Daniel Patrick, Merrick Miranda, Jennifer DeHeck, Chris Bauman, Evan Leonard
- Editted by: Alfred Chung (Derivations), Ran Li (Determining Stable Controller Gain), Nirala Singh (Determining Slope of Bode Plot), Katy Kneiser and Ian Sebastian (synthesis with 2006 wiki "Bode Plots"), Robert Appel, Jessica Rilly and Jordan Talia (Formatting

This page titled 9.4: PID tuning via Frequency Responses with Bode Plots is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

9.5: PID tuning via Optimization

Introduction

Tuning a controller is a method used to modify the effect a process change will have on the piece of equipment being controlled. The goal of tuning a system is to construct the most robust process possible. The method chosen for tuning a system varies depending on the parameter being measured, the sensitivity of the materials, the scale of the process, and many other variables unique to each process. This chapter discusses the basics of tuning a controller using predictive methods. To learn more about using an effect based method see the Classical Tuning section.

Optimization

When tuning a PID controller, constants K_c, T_i, and T_d need to be optimized. This can be accomplished using the following equation from the Classical Tuning section:

$$MV = K_c \left(e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau + T_d \frac{de(t)}{dt} \right) \tag{9.5.1}$$

where

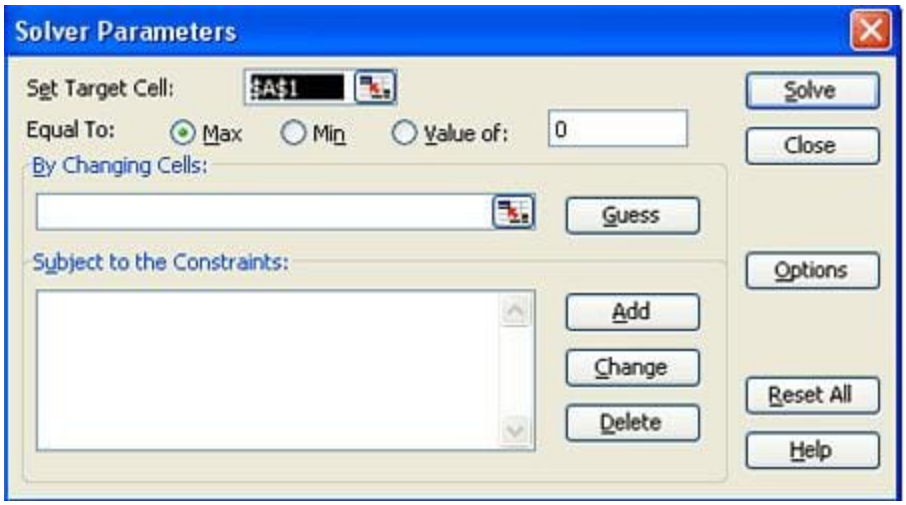
- MV is the manipulated variable (i.e. valve position) that the controller will change
- K_c accounts for the gain of the system
- T_i accounts for integrated error of the system
- T_d accounts for the derivative error of the system

Tuning by optimization uses computer modeling programs, such as Microsoft Excel, to find the optimal values for the coefficients K_c, T_i, and T_d to yield the minimum error (the Solver function in Excel can be used in this situation). For more information about the use of these parameters and their overall effects on the control system see the P, I, D, PI, PD, and PID control section.

Excel Modeling

For instructions on how to install Solver in Excel, see the Adding in the Solver Application in Excel 2007 section. For instruction on how to use Solver in Excel, see the Excel's Solver Tool section. Equation 9.5.1 can be used in Excel to optimize the values of K_c, T_i, and T_d. The following are the steps to optimize these constants:

- Pick a set point. Examples of set points include the temperature at which a reaction is expected to stay and the flow rate of a cooling stream.
- Make a column for error. Create a function in each cell so that the error is equal to the set point minus the actual value.
- Make a column that will calculate Equation 1.
- At the end of the error column, use the sum function to add up all of the errors in the column.
- Make three cells for K_c, T_i, and T_d, and insert the initial values.
- Open the Solver function in Excel.
- In Solver, set the target cell to the summed errors cell. Set the target cell equal to the minimum value possible. Set the changing cells to K_c, T_i, and T_d.
- Click 'Solve' to minimize the sum of the error and therefore optimize the K_c, T_i, and T_d constants.



Tips for using Excel Solver to optimize PID controller parameters:

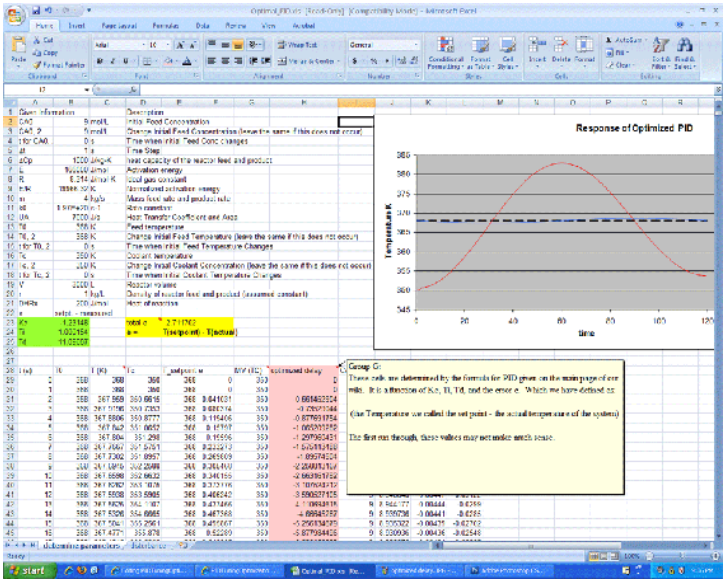
- When choosing initial values for your parameters, try starting with K_c=1, T_i=1000, and T_d=1. Since Solver only does a limited number of iterations (usually set at 100), it may settle on an incorrect answer.
- Change the values for K_c manually and optimize with Solver until you get the lowest value in the target cell. This is the most efficient as K_c usually has the largest effect on a PID controller.
- Make sure to set constraints on your parameters in the 'Subject to the Constraints:' box in Solver (e.g. $K_c > 0$).
- Make sure you incorporate any physical limitations (e.g. maximum heater temperature) into your Excel model. Even though your system may not be close to the limit initially, Solver may try and go past that limit if it is not programmed into your model.

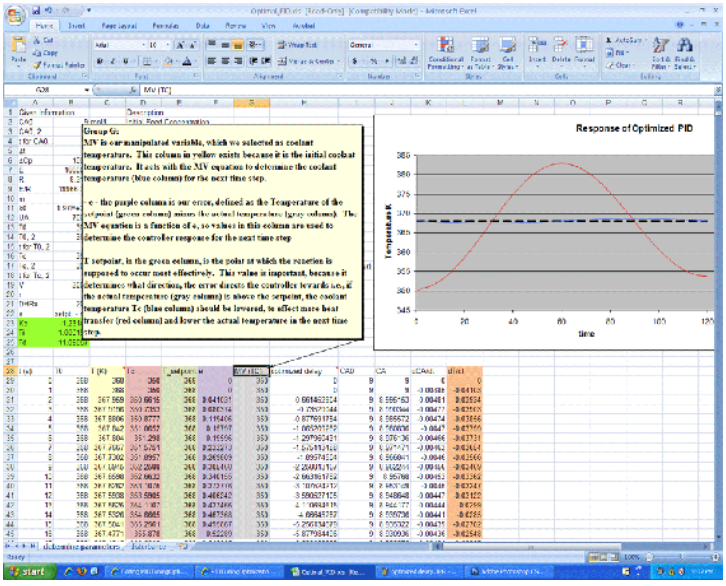
Industrial Application

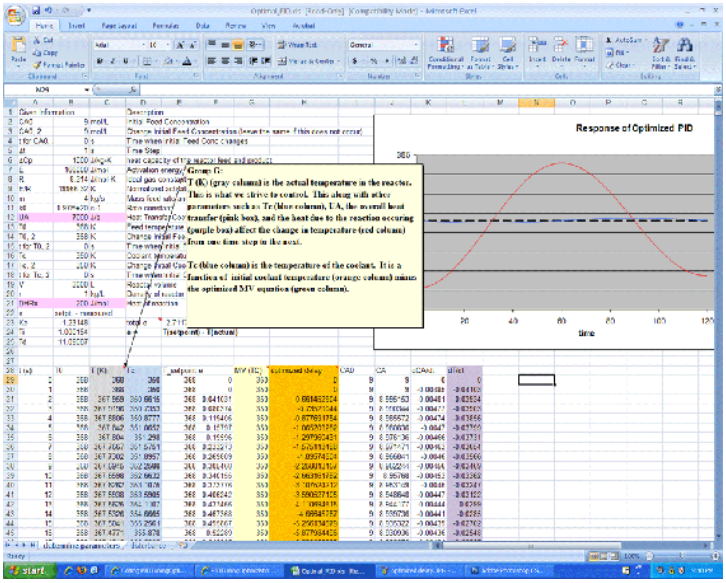
The parameters (K_c, T_i, T_d) found through optimization will not necessarily give the best control. To optimally tune a controller, a technician must adjust equipment to tune the process. So why is it necessary to tune by optimization? The parameters found by optimization give an accurate starting point for the technician.

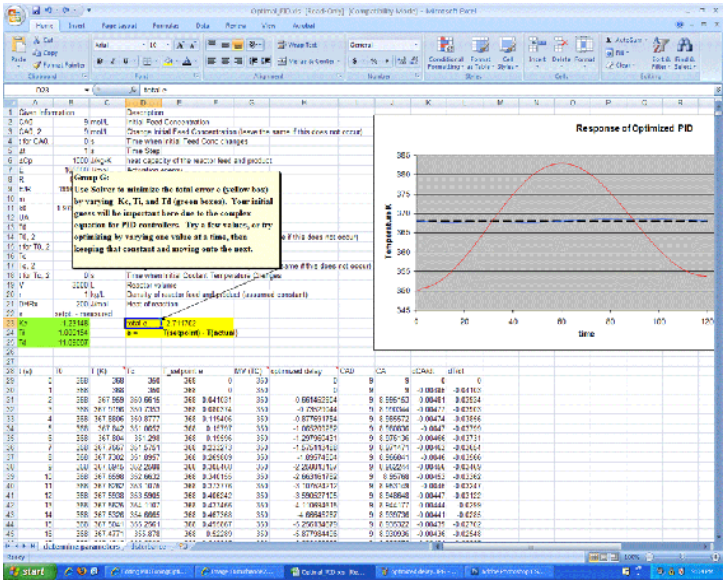
Example of PID Tuning by Optimization

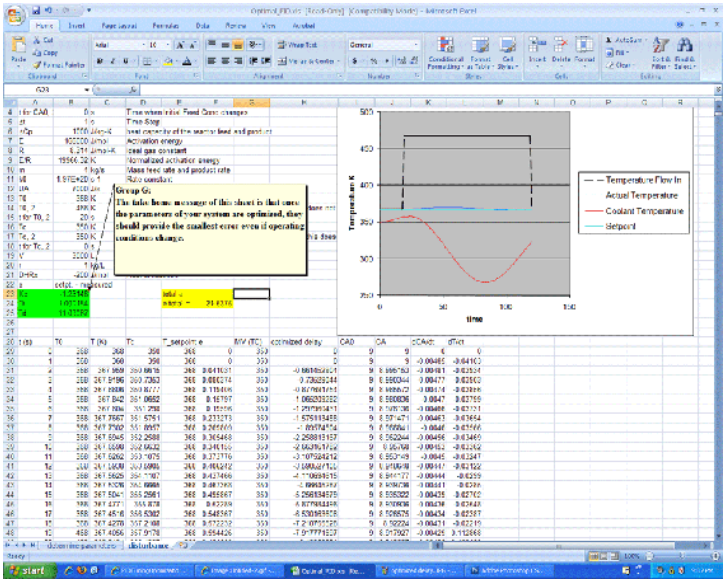
This example is based on the example in the CSTR Heat Exchange Model section. The goal is to optimize the PID controller for the coolant temperature. Follow the fully worked out example at: [Media:Optimal PID.xls Optimal PID]


Response of Optimized PID
Temperature,K
time


Response of Optimized PID
Temperature,K
time


Response of Optimized PID
time


Response of Optimized PID
Temperature,K
time



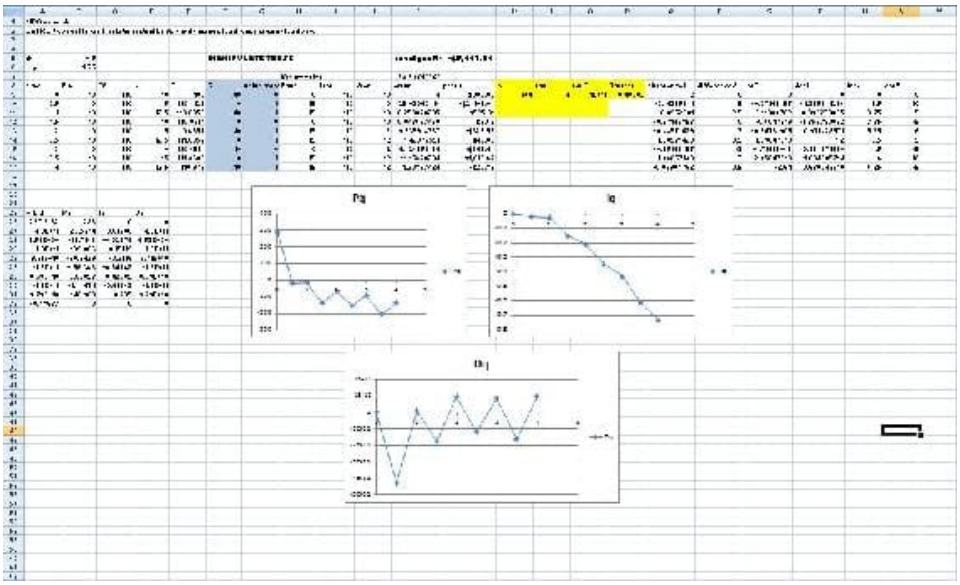
As shown in the Excel file, the coolant temperature changes in response to variations from the set point by a PID type controller. Changing the K_c, T_i, and T_d values using Solver, to minimize the total error, provided the values in the green box. On the second sheet, labeled 'disturbances', it can be seen that once optimized, the parameters should be fit to the system, regardless of changes in operating conditions.

It is important to note that the starting values you choose for the PID parameters will greatly affect your final results from using an excel solver. So you should always use your intuition to judge whether your starting values are reasonable or not. According to ExperTune , starting PID settings for common PID control loops are:

- Loop Type: Flow ; P = 50 to 100 ; I = 0.005 to 0.05 ; D = none ;
- Loop Type: Liquid Pressure ; P = 50 to 100 ; I = 0.005 to 0.05 ; D = none
- Loop Type: Gas Pressure ; P = 1 to 50 ; I = 0.005 to 0.05 ; D = 0.02 to 0.1 ;
- Loop Type: Liquid Level ; P = 1 to 50 ; I = 1 to 100 ; D = 0.01 to 0.05 ;
- Loop Type: Temperature ; P = 2 to 100 ; I = 0.2 to 50 ; D = 0.1 to 20 ;

The above values are rough, assume proper control loop design, ideal or series algorithm and do not apply to all controllers. These information should only be used as a possible consideration and should not be taken as an absolute starting value for all PID controls. For more information about the above values please go to www.expertune.com/tutor.html

As mentioned before, to optimize the parameters in a PID-controlled system the solver function can be used in excel. However the solver function can sometimes run into its own problems because of the process it uses to solve for these “optimal” values. To ensure that solver gives you the correct optimal values, some manual “optimization” is needed. Using your excel sheet, the controller on the system can be separated into its different components P, I, and D. The behavior of each component can be monitored as the parameters are changed manually. Theoretically you would choose values that have every component reach steady-state over time. The excel solver can then be used as a precision tool. If your values are already closer to the optimal values, then solver should have no problem making them more precise. It is recommended that the graphs of the components be checked after using solver with your initial guesses. You may not have to go through this long and most of the time painful process. This is an example of how the P, I, and D plots change with different parameters. Note that the combined PID plot will also reach steady-state if each of the components reach steady-state.



(Parameters are not optimized)

(Optimized parameters)

Example: Optimization of a Heat Exchanger

This example is based on the ODE Modeling in Excel of the Heat Exchange Model.

Find the best values of K_c, T_i, and T_d in the given spreadsheet Media:PID_HeatExchange.xls.

Note the three different tabs:

- the first tab shows the problem to be solved
- the second is a step by step method for creating this sheet on your own
- the third is the answer to the problem

After using Solver:

- Click through the cells in one row to determine which parameters are affected by MV changes
- Change the values of K_c, T_i, and T_d to see how they will effect the overall temperature profile. Which parameter has the

greatest effect? Based on the equation, does this make sense?

Exercise 9.5.1

What do you want to minimize when optimizing the constants K_c, T_i, and T_d?

Answer available at the end of this section.

References

- Svrcek, William Y., Mahoney, Donald P., Young, Brent R. "A Real Time Approach to Process Control", 2nd Edition. John Wiley & Sons, Ltd.

Contributor

- Authors: Andrew MacMillan, David Preston, Jessica Wolfe, Sandy Yu
- Stewards: YooNa Choi, Yuan Ma, Larry Mo, Julie Wesely

Answers

Exercise 9.5.1

time

This page titled 9.5: PID tuning via Optimization is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

9.6: PID Downsides and Solutions

Introduction

A **proportional-integral-derivative (PID) controller** is one of the most common algorithms used for control systems. It is widely used because the algorithm does not involve higher order mathematics, but still contains many variables. The amount of variables that are used allows the user to easily adjust the system to the desired settings. The algorithm for the PID uses a feedback loop to correct the difference between some measured value and the setpoint. It does this by calculating and outputting some action that will correct this error in the system. A PID controller has a proportional, integral and a derivative control which handles the current, past and predicted future of the signal error. For more information about PID, please refer to PID Intro. The PID controller can operate systems that run in a linear or nonlinear fashion. Tuning processes are done to the controller to tackle the possible nonlinear system. Limitations arise within the system because tuning is limited to only three different parameters (proportional, integral, and derivative controls). Additional information on tuning of PID can be found at [1] or [2]. The most common limitations that occur within the PID control specifically involve the integral control. The following article addresses some of the common limitations faced by each control type, with an emphasis on the integral control, and some solutions to overcome each of these limitations.

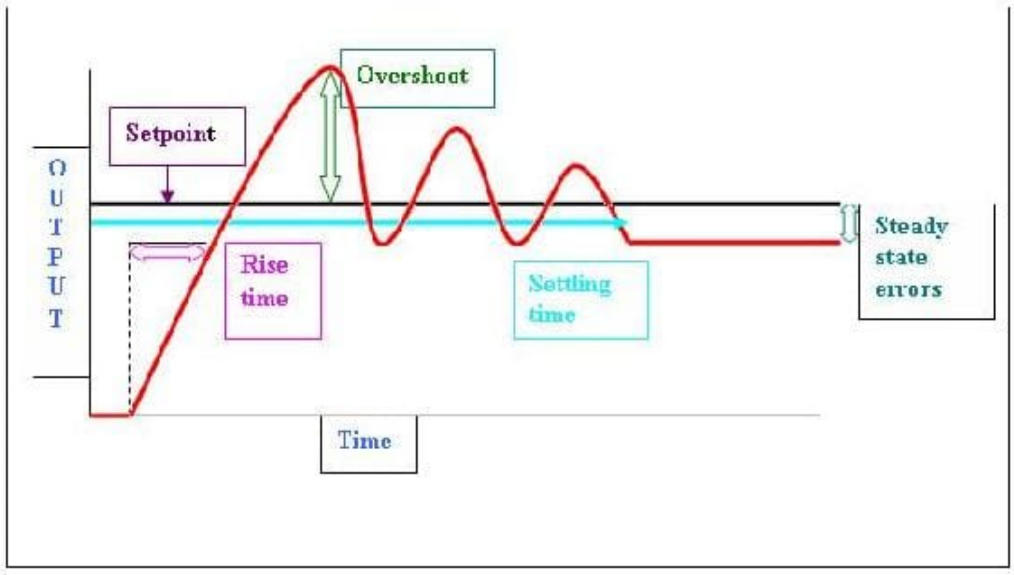
Proportional Control

The main purpose of the proportional control is minimize the fluctuations that occur within the system.

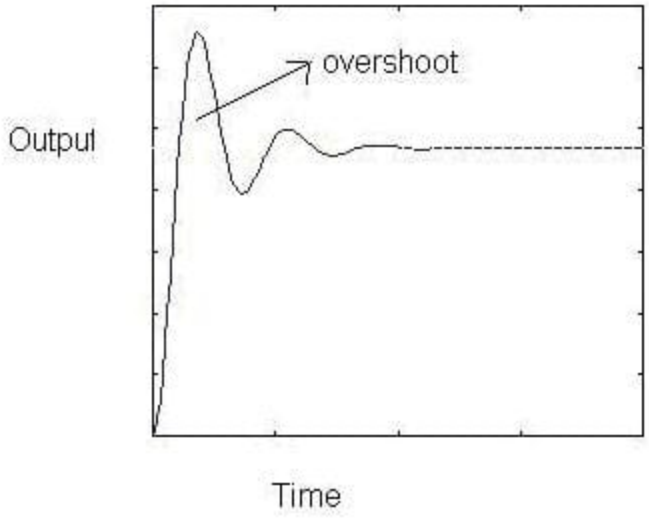
Limitations

The P-controller usually has steady-state errors (the difference in set point and actual outcome) unless the control gain is large. As the control gain becomes larger, issues arise with the stability of the feedback loop. For instance, reducing the rise time implies a high proportional gain, and reducing overshoot and oscillations implies a small proportional gain. This is not possible to achieve in all systems.

Below is a general process outcome diagram showing the terminology used above.



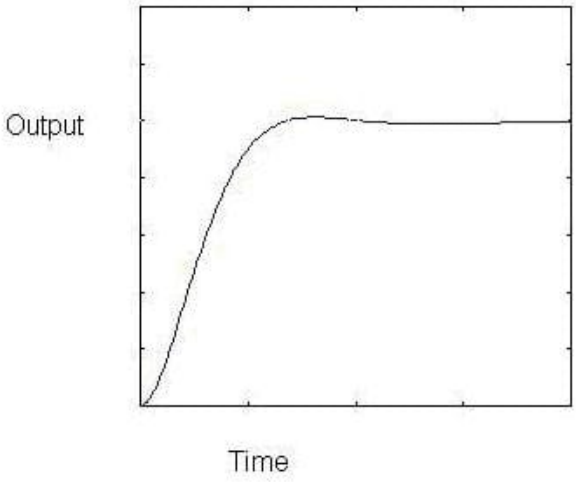
Below is a sample diagram of process output of proportional control.



Solutions

The way to eliminate these steady-state errors is by adding an integral action. The integral term in the equation drives the error to zero. Higher Integral constant (1 / T_I) drives the error to zero sooner but also invites oscillations and instability. Read on the integral control section below to know more about limitations associated with this integral term.

Below is a sample process output diagram when integral control is added.



The above picture shows the reduction of overshoots and oscillations compared to the picture before adding the integral action.

Integral Control

The contribution from the integral term is proportional to both the magnitude of the error and the duration of the error. Summing the instantaneous error over time (integrating the error) gives the accumulated offset that should have been corrected previously. The accumulated error is then multiplied by the integral gain and added to the controller output. The magnitude of the contribution of the integral term to the overall control action is determined by the integral gain, K_i.

The integral term is given by:

$$D_{\text{out}} = K_d \frac{de}{dt}$$

where

- D_{out}: is the **Derivative output**
- K_d: is the **Derivative Gain**, a tuning parameter
- e: is the **Error** = *SP* – *PV*
- t: is the **Time** or instantaneous time (the present)

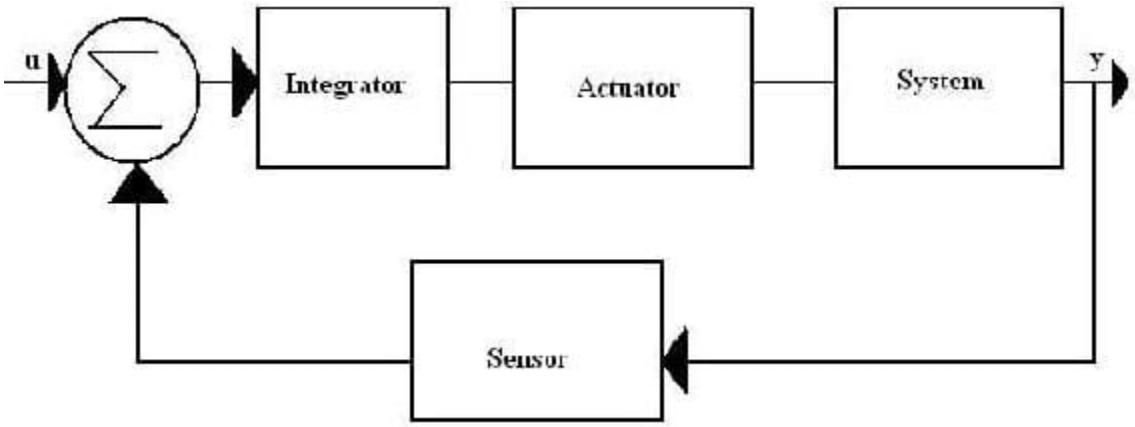
Limitations

Windup

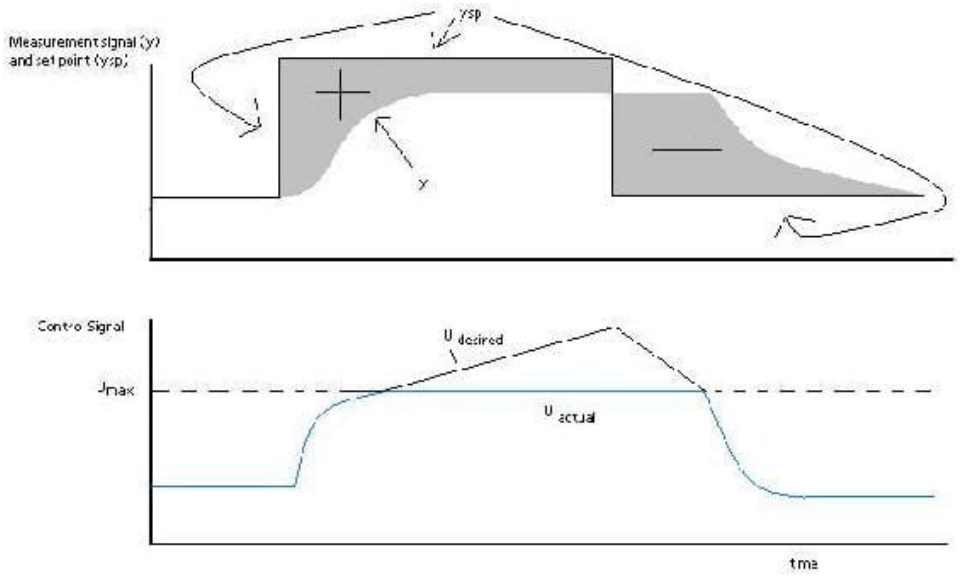
A basic knowledge of the concept of windup is useful before describing a specific type. Windup is defined as the situation when the feedback controller surpasses the saturation (i.e. maximum) limits of the system actuator and is not capable of instantly responding to the changes in the control error. The concept of the control variable reaching the actuator's operation limits is reasonable considering the wide variety of operating conditions that are possible. When windup occurs the actuator constantly runs at its saturation limit despite any output the system might have. This means that the system now runs with an open loop instead of a constant feedback loop.

Integrator Windup

The most common type of windup that occurs is integrator windup. This occurs when the input into the system receives a sudden positive step command and causes a positive error when the system first responds to the actuator. If the rate of integration is larger than the actual speed of the system the integrator's output may exceed the saturation limit of the actuator. The actuator will then operate at its limit no matter what the process outputs. The error will also continue to be integrated and the integrator will grow in size or "wind up". When the system output finally reaches the desired value, the sign of the error reverses (e.g. \ yy_{sp} " src="/@api/deki/files/18198/image-870.png">) and causes the integrator to "wind down" until things return back to normal. Through the wind down the control signal is still maximum for a long period of time and the response becomes delayed. The integrator takes a long period of time to fully recover to the operating range of the actuator. Integrator windup may occur from large set point changes, significant disturbances, or equipment malfunctions.



The illustration is a flowchart showing the specific steps that take place through the integrator controller. It shows the input (u) and output of the system (y) , along with the integrator, the actuator, the system, and sensor involved in the process. The sigma used in each flowchart is used to represent the summation of all variables inputed to it.



The graphs above are illustrations of integrator windup. The error (e) is shown in the top graph as $e = y_{sp} - y$, where y_{sp} is the set point and y is the measured signal. The bottom graph displays the control signal (u). In this specific case of integrator windup, a set point change occurs and the control signal becomes saturated to its maximum amount u_{max}. The system's error is still not eliminated though because the control signal is too small to make the error go to zero. In turn, this causes the integral to increase, which causes the desired control signal to increase. Therefore, there continues to be a difference between the desired control signal and the true control signal. After a period of time, the set point is lowered to a level where the controller is able to eliminate the control error. At this point \ yy_{sp} " src="/@api/deki/files/18198/image-870.png">, which causes a change in the sign of the error. This change in sign causes the control signal to begin decreasing. The true signal (u_{out}) is stuck at u_{max} for a while due to desired control signal u being above the limit u_{max}. The changes in setpoints throughout this specific example occur because the input is changed in order to get a minimal error in the system.

File:Integrator windup output control integral5.JPG

The images above are another display of integrator windup. The top most shows the error, the middle shows the control signal, and the bottom shows the integral portion. The integral term begins to decrease, but remains positive, as soon as the error becomes negative at t_1. The output stays saturated due to the large integral termal that developes. The output signal remains at this level until the error becomes negative for a sufficiently long time (t_2). The control signal then fluctuates several times from its maxmimum to its minimum value and eventually settles down on a specific value. At every place where the control signal reaches its maximum the integral value has a large overshoot, and it follows that every place the control signal reaches a minimum the integral has a dampened oscillation. The integral term accounts for the removal of error in the oscillating system, which in turn causes the system's output to eventually reach a point that is so close to the setpoint that the actuator is no longer saturated. At this point, each graph begins to behave linearly and settles down. This example will be used later to show a solution to integral windup. The following section lists several ideas of how to prevent windup.

Solutions

The solutions involved in integral windup are termed as anti-windup schemes. Anti-windup is a method of counteracting the windup that happens in the integration that occurs in the integral controller of the PID.

Back-Calculation and Tracking

When the output is saturated, the integral term in the controller is recomputed so that it too is at the saturation limit. The integrator changes over time with the constant T_t . The difference between the output controller (v) and the actuator output (u) is found, and this is defined as an error signal. This error signal is then fed back to the input of the integrator with a gain

$\frac{1}{T_t}$ of $\frac{1}{T_t}$. The controller output is proportinal to the amount of time that the error is present. The error signal is only present when the actuator is saturated, therefore it will not have any effect on normal operation. With this method, the normal feedback path around process is broken and a new feedback path around the integrator is formed. The integrator input becomes:

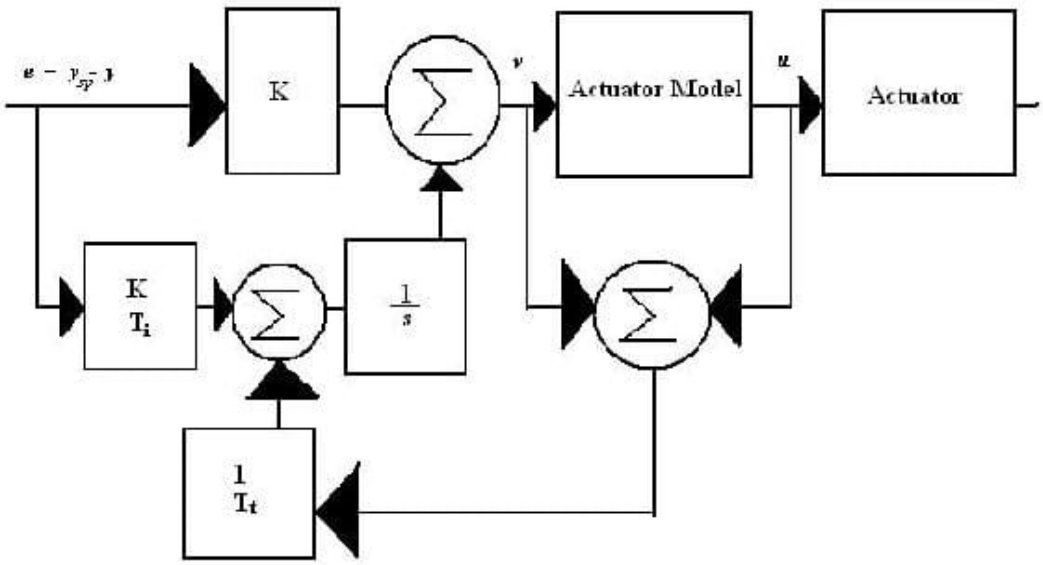
$$frac1T_te_s + \frac{K}{T_i}e$$

$$e_s = \frac{-KT_t}{T_i}e$$

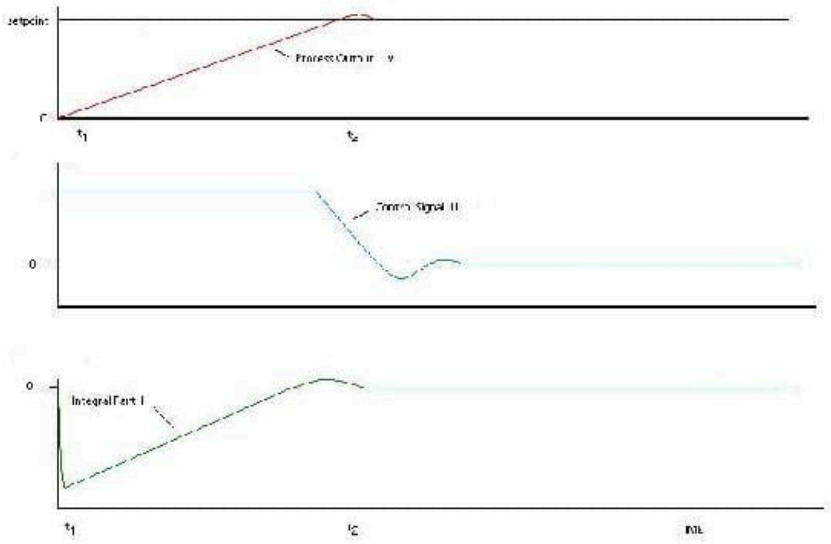
Knowing that $e_s = u - v$, the following equation can be made:

$$v = u_{lim} + \frac{KT_t}{T_i}e$$

u_{lim} is the control variable when saturated. The signs for u_{lim} and e are always the same; therefore v is always bigger than u_{lim}. This in turn prevents the windup in the integrator. The feedback gain is $\frac{1}{T_t}$ and is the rate at which the controller output is reset. T_t is called the tracking time constant and is used to tune the amount of anti-windup. The flowchart shown below shows a PID controller with anti-windup implemented. The error signal in this process is only non-zero when the actuator is saturated. When this occurs the normal feedback path around the process is broken because the process input remains constant. A new feedback path is then formed around the integrator and causes the integrator output to be driven towards a value such that the integrator input becomes zero. Back-calculating never allows the input to the actuator to reach its actual saturation level because it forecasts what will actually go into the actuator model beforehand.



The following plots describe what happens to the same system (second example) described in the introduction to Integrator Windup section, only the controller now has anti-windup. The integrator output is at a value that causes the controller to be at the saturation limit, however, the integral is intially negative when the actuator is saturated (t_1), which contrasts with the original behavior of the integrator. This has a positive effect on the process output, as it converges to the set point much quicker than the normal PI controller. There is only a slight overshoot in the process output (t_2).



Conditional Integration

Conditional integration operates similarly to back-calculating. In this method the integration is turned off when the control is far from steady state. Integral action is then only used when certain conditions are met, otherwise the integral term is held constant. There are two possible conditions that may cause the integral to be set as a constant value. First, if the controller output is saturated (integration is shut off if error is positive but not negative), then the integrator input is set to a constant value. A second approach involves making the integrator constant when there is a large error. There is a disadvantage to the second approach because the controller may get stuck at a nonzero control error if the integral term has a large value at the time of switch off. For this reason, the first approach looks at the saturated controller output, not the saturated actuator output because referring to the actuator output will generate the same disadvantage[1] .

There is very little difference performance wise between integration and tracking, but they move the proportional bands differently. Read on to learn about proportional bands.

To demonstrate what it means to turn the integral term off, here is the equation (with logic) representing a control with conditional integration.

$$u(t) = K\left(e(t) + \frac{1}{T_i}\int e(t)dt\right) + u\left(t_0\right)$$

If $e(t) < \frac{e_{max}}{K}$, then $T_i = T_i$

Else $T_i = 1,000,000$

Series Implementation

To avoid windup in an interacting controller a model of the saturation can be incorporated into the system. The model of saturation will limit the control signal indirectly. Too hard a limitation will cause an unnecessary limitation of the control action. Too weak a limitation will cause windup. More flexibility is provided if the saturation is positioned as follows:

With this setup it is possible to force the integral part to assume other preload values during saturation. The saturation function is replaced by the nonlinearity shown. This anti windup procedure is sometimes called a “batch unit” and is considered a type of conditional integration. It is mainly used for adjusting the overshoot during startup when there is a large set point change.

The Proportional Band

The proportional band is useful to understand when dealing with windup and anti windup schemes. The proportional band is a range such that if the instantaneous value of the process output or predicted value on the interval falls within the range, the actuator does not saturate. The proportional band (defined as the area between y_{low},y_{high}) is given by the following equations:

Control signal for PID control without derivative gain limitation:

$$u = K * (b * y_{sp}) + I - K * T_d * \frac{dy}{dt}$$

The predicted process output is given by

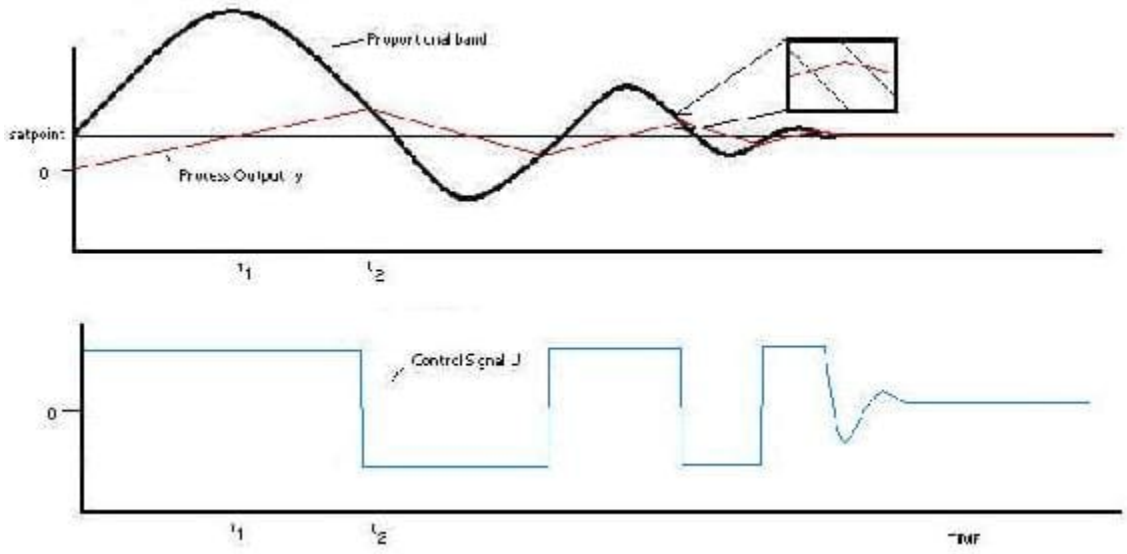
$$y_p = y + T_d * \frac{dy}{dt}$$

Combining the above equations for a maximum and minumum control signal (u_{max} and u_{min})which corresponds to the points for which the actuator saturates,

$$y_{low} = b * y_{sp} + \frac{I - u_{max}}{K}$$

$$y_{high} = b * y_{sp} + \frac{I - u_{min}}{K}$$

If the predicted output is in the proportional band, the controller operates linearly. The control signal saturates when the predicted output is outside of the proportional band. The following plot illustrates how the proportional band is useful in understanding windup for the same example. At time equals zero, the proportional band is increased, indicating that the integral part is increasing. Immediately after t_1, the output is greater that the setpoint, which causes the integral part to start to decrease. As can be seen, the output value, y does not reach the proportional band until it is greater than the setpoint, y_{sp} . When the output finally reaches the proportional band at t_2, it goes through the band very quickly, due to the band changing at a high rate. The control signal u decreases immediately and saturates in the opposite direction, causing the output to decrease and approach the setpoint. This process repeats as the output value converges with the setpoint.



The tracking time constant has a major influence on the proportional band. The proportional band is moved closer to the output because of tracking. The tracking time constant determines the rate that the proportional band moves at.

Real Life Scenario

Imagine a situation where a water storage tank that is needed to store water and supply for use in a plant and it is necessary to keep the water level constant. A level sensor is installed to keep track of the water level in the tank, and it sends this data to the controller. The controller has a set point for instance 75% full. The rate of water entering the tank is increased by opening valve and decreased by closing the valve. In the case of integral windup, the valve is all the way closed or open and the actuator sends the signal to close it more or open it more which is not possible.

Derivative Control

Derivative action can be thought of as making smaller and smaller changes as one gets close to the right value, and then stopping in the correct region, rather than making further changes. Derivative control quantifies the need to apply more change by linking the amount of change applied to the rate of change needed. For example, an accelerator would be applied more as the speed of the car continues to drop. However, the actual speed drop is independent of this process. On its own, derivative control is not sufficient to restore the speed to a specific value. Pairing the match in change with a proportionality constant is enough to properly control the speed.

The derivative control is usually used in conjunction with P and/or I controls because it generally is not effective by itself. The derivative control alone does not know where the setpoint is located and is only used to increase precision within the system. This is only control type that is open loop control (also known as feedforward loop). The derivative control operates in order to determine what will happen to the process in the future by examining the rate of change of error within the system. When the derivative control is implemented the following general equation is used:

$$u(t) = \frac{1}{T_d}\frac{d}{dt}e(t) + u\left(t_0\right)$$

where

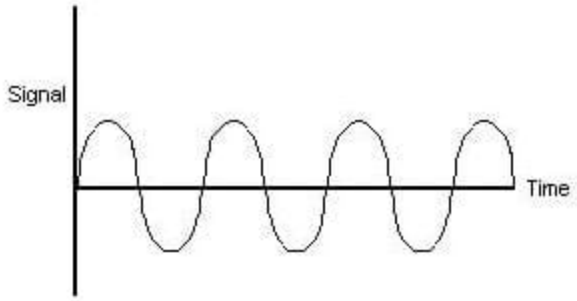
- $u(t)$ = control signal
- T_d = derivative time
- $e(t)$ = error
- $u(t_0)$ = initial control signal

This equation shows the derivative control is proportional to the change in error within the system.

Limitations of Derivative Control

The major problem associated with this control is the noise problem. When the frequency within the system is high (change in error of the system is large), taking the derivative of this signal may amplify the signal greatly. Therefore, small amounts of noise within the system will then cause the output of system to change by a great amount. In these circumstances, it is often sensible use a PI controller or set the derivative action of a PID controller to zero.

The following is a graphical represenation of some sinusoidal output signal function of a system.



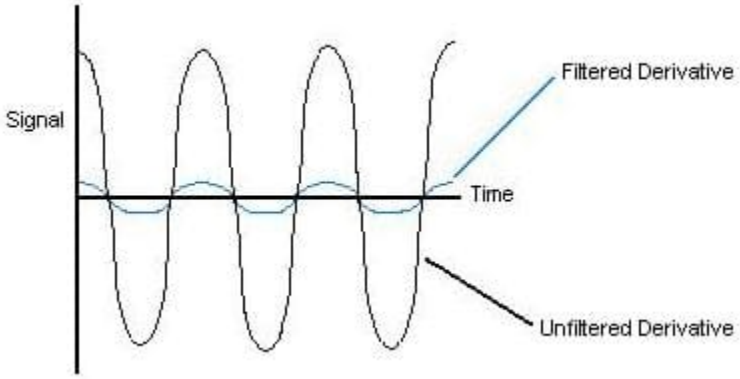
Solutions

To eliminate/minimize this problem, an electronic signal filter can be included in the loop. Electronic signal filters are electronic circuits which perform signal processing functions, specifically intended to remove unwanted signal components and/or enhance wanted ones. Electronic filters can be: passive or active, analog or digital, discrete-time (sampled) or continuous-time, linear or non-linear, etc. The most common types of electronic filters are linear filters, regardless of other aspects of their design.

For more information on electronic signal filters, reference the following website:

en.Wikipedia.org/wiki/Electronic_filter

The derivative term in the equation is usually changed by putting a first order filter on the term so that the derivative does not amplify the high frequency noise that is attenuated. Below is a sample outcome figure of a possible derivative of the output signal shown above along with the filtered signal.



As shown, it is possible for the amplitude to be magnified when the derivative is taken for a sinusoidal function. A filter is usually a set of complicated equations that are added to the derivative that effect the function as shown.

Robustness of PID controllers

When designing a PID controller it is important to not only create a controller that will correct the error in an efficient manor, but also a robust controller. No matter how good the controller is if it cannot handle the magnitude of certain disturbances then it will not be able to perform well and complete the job. The robustness of a controller is the ability of the controller to withstand disturbances and still operate effectively. In order to calculate controller robustness, a weighted sum of the integral of the squared controller effort (ISC) and the integral of the squared controller error (ISE) is calculated. The smaller the weighted sum the more robust the controller.

ISE = 0 ∫ ∞ e(t)2 dt

ISC = 0 ∫ ∞ [u(∞)-u(t)]2 dt

IT = w ISC + ISE

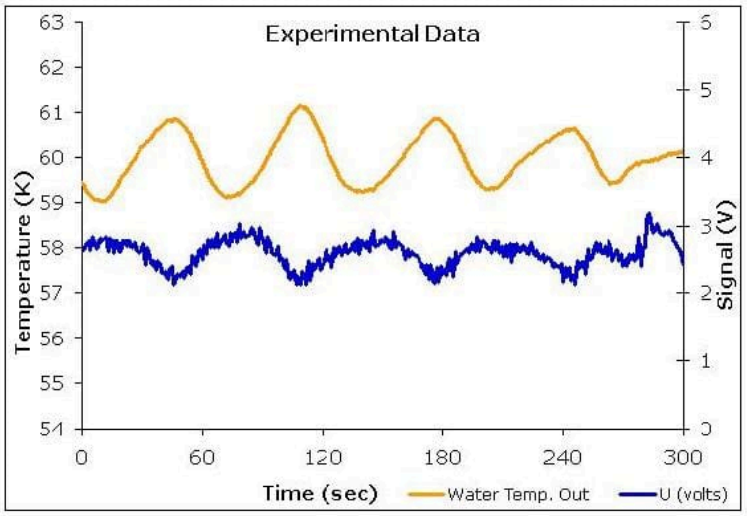
In these equations u is the input variable, w is the weight factor which is based on downstream processes and IT is the total error and controller effort. The weight factor is used to assign a higher significance to the ISC or the ISE. The significance depends on the design and whether it is more important to have minimal error and tight control or less controller effort with more error. Low values of w are used when tight control is desired. Robustness calculations are determined experimentally. The robustness depends a lot on the type of control used, whether it is feedback, feed- forward, cascade, or a combination.

Notes on robustness

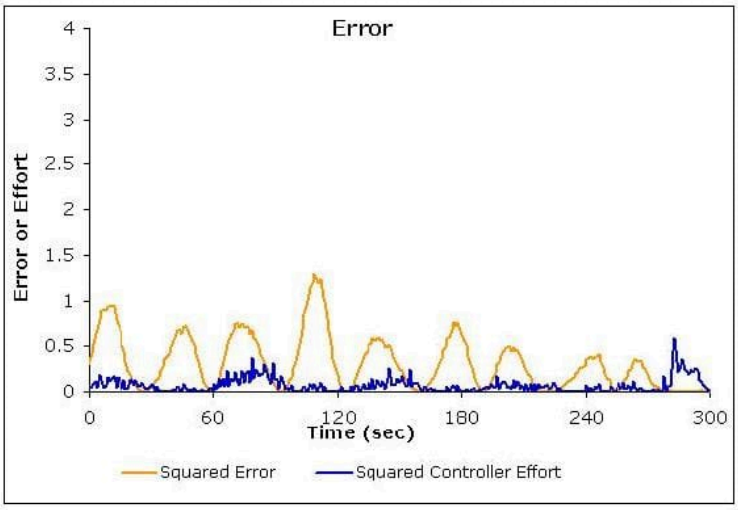
It is important to realize that when calculating the robustness of a feed forward system, the system will reach a steady state, but it is unlikely that the system will reach the original set point. As a result, the feed forward system will most likely be less robust than a feedback system. The reason the feed-forward system doesn't return to the set point but finds a new steady state is because it uses open loop control and there are no sensors to inform it that its adjustment to the disturbance did not work correctly. For this reason, feed-forward controllers are sometimes coupled with a feedback system. When combined the feed-forward controller can now see if its corrections achieved the desired result or not, resulting in a more robust controller.

Example Robustness Calculation

In this system, the controller is regulating the temperature of the feed to a certain downstream process. After collecting experimental data from your controller, the following graph was made.



Using this graph the squared controller effort and the squared error can also be graphed and the ISE and ISC can be calculated. The graph of the corresponding effort and error can be seen below.



Using the squared controller effort and the squared error graph, the robustness can be calculated. To find the integral of the squared error and the squared controller effort the trapezoid rule can be used.

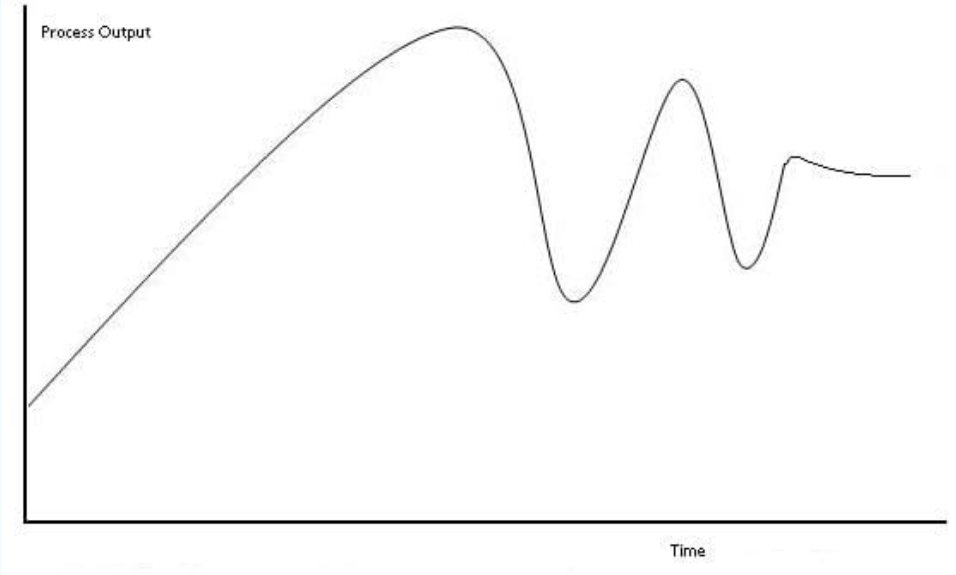
$$IT = wISC + ISE$$

$$ISC = 21ISE = 94$$

Assuming a weight factor of 1, the robustness of the controller is 115.

Example 9.6.1

Consider the below output from using one of the controllers. The general equation of the output signal, U = K(e). What problems do you see in the below picture and how can you correct them?

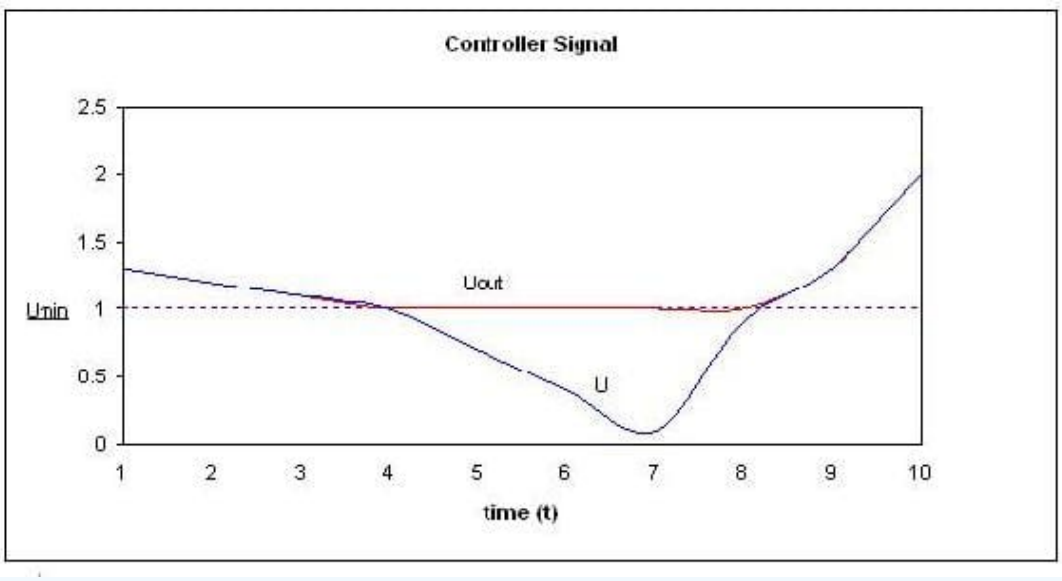


Solution

The response time is very high, and it has a high overshoot and oscillation problems as well as steady-state errors. The equation confirms that it is a P-controller and

Example 9.6.2

A heat exchanger uses a PID control system to control the temperature of the outlet of tube side of the heat exchanger. It's required to keep water at 20 C for the cooling of an effluent from an exothermic reaction. The specific controller has a min output of 1 volt and a max output of 6 volts. The experimental data shows a large lag time in the controller's ability to cool the incoming fluid (the lower portion of the sinusoidal curve). The following data shows the correlation between the controller output and the calculated signal. Describe the type of problem and a possible solution to the problem.



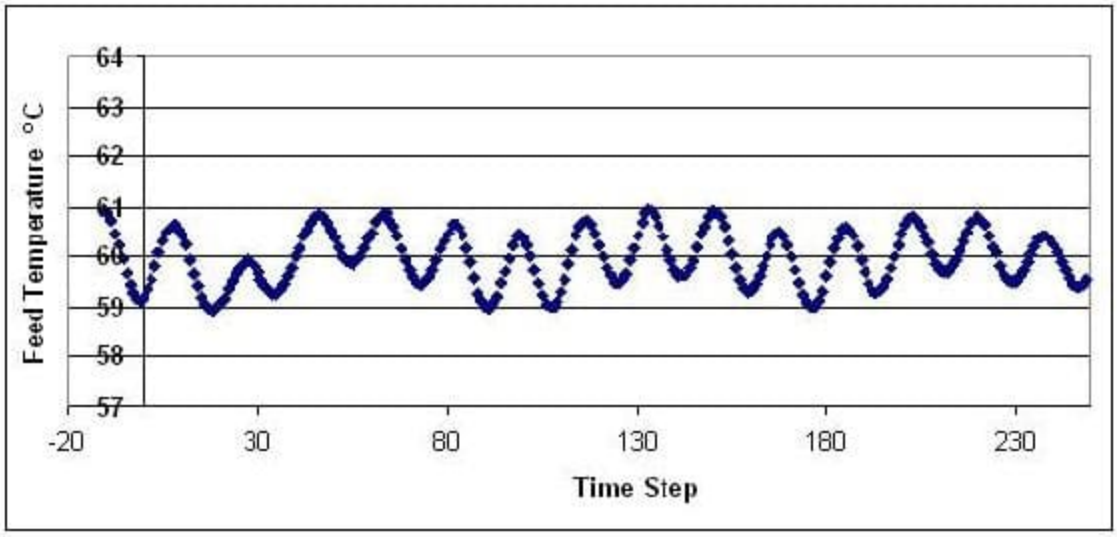
Solution

The Problem shown is an example of integrator-windup. The controllers calculated signal is lower than the min output that it can handle. This explains the lag time seen for the lower portion of the sinusoidal curve. To fix this problem the integral control can be recomputed using back track calculation.

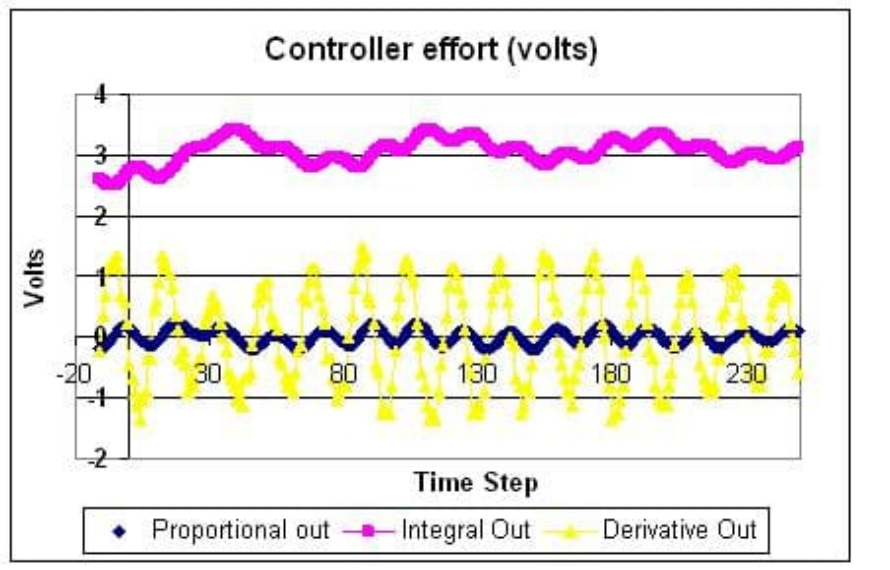
Example 9.6.3

(a) You are using a PID controller to maintain the feed to a reactor at a constant temperature. It is imperative that the feed remain within 1 °C of the set-point. You've tuned your PID control system according to the Ziegler-Nichols open loop system and obtained the following values for Kc, τI, and τD: Kc = 0.218, 1/τI = 0.0536, τD = 4.67

You want to know if these are good control values or if they can be improved. Your control variable output looks like this:



Additionally, you have available the following plot of the voltage output from your control parameters:

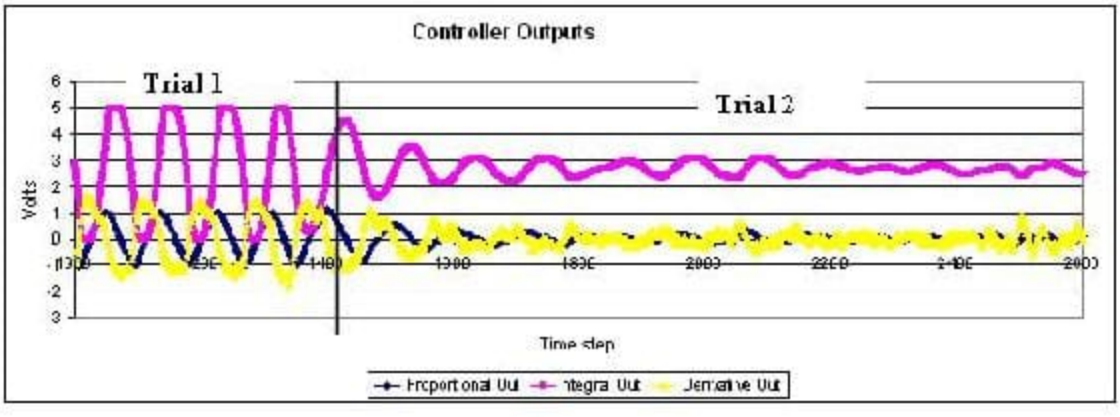


Is this system in control? If so, explain why. If not, explain why and what (qualitative) improvements to the PID gains will bring the system into better control.

Solution

The system is in control in that it appears to roughly stay within the defined acceptable boundaries of ±1°C. However, it does display oscillation, indicating that the system is working a little harder than maybe it has to. Looking at the controller voltage output, we see that the derivative gain is giving high oscillations. Thus maybe decreasing this gain would keep the system in control with a little less wear and tear on the control valves.

(b) At another flow rate, you observe the following behavior in your controller:



Is the system under control in Trial 1? In Trial 2? Why or why not? What physically happened between trials to give the observed change in output?

Solution: In trial 1, without even looking at the direct system output, we can guess that it is probably not in control. The controller output voltage indicates that the integral gain is way too high, and its wild oscillations will doubtlessly be causing oscillations in the system. Physically, this means that the system is opening and closing the valve constantly, creating a lot of wear and tear on the machinery, and probably overcompensating for small changes and further destabilizing the system. In trial 2, we can see that the integral voltage oscillations have damped, indicating that the gain has been reduced. As these oscillations level out, the oscillations in the other control gains also damp down, as the system comes into better control and none of the terms are working as hard as they had been before.

Example 9.6.4

Note this is an open-ended problem meant to give readers an intuition about PID controllers

Imagine the classic differential equation problem of two tanks in a series governed by the differential equations:

$$\frac{dh_1}{dt} = PID - K_1\sqrt{h_1}$$

$$\frac{dh_2}{dt} = K_1\sqrt{h_1} - K_2\sqrt{h_2}$$

\[\P I D=1+K_{p} e(t)+\frac{1}{t_{i}} \int_{0}^{t} e(\tau) d \tau+t_{d} \frac{d e(t)}{d t}]

where

- h_1 = height of tank 1
- h_2 = height of tank 2
- K_i = valve constant associated to tank 1 or 2

The valve constants $K_1 = 5$ and $K_2 = 5$ and the set point is 3. What values of K_p, t_i, and t_d will cause the two tanks to reach steady state the fastest? Also what values will give the most oscillations?

Solution

Click on this link for the worked out Mathematica solution

Other questions that can be explored with slight modifications of the code are:

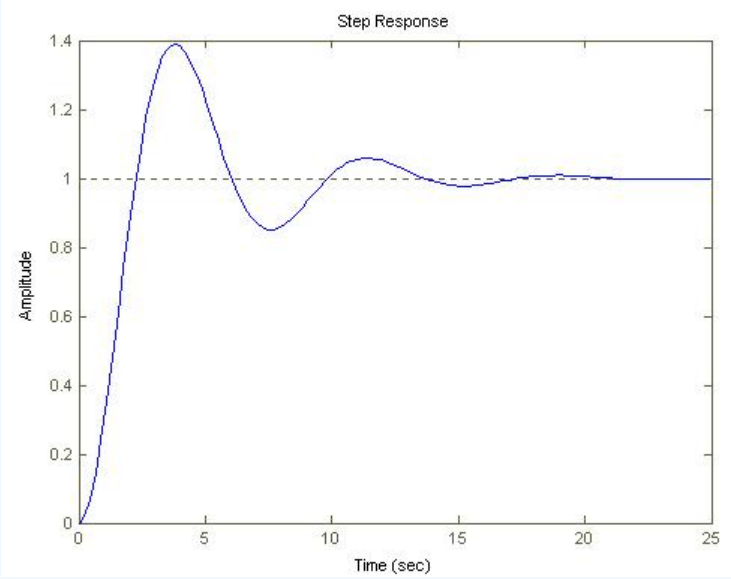
1. Will the values of K_p, t_i, and t_d yield similar results if the PID controller was placed on tank 2? What new values of K_p, t_i, and t_d will cause the two tanks to reach steady state the fastest?

2. For the values of K_p, t_i, and t_d that causes the two tanks o reach steady state the fastest, what happens if you increase K_1? What will happen if you increase K_2?

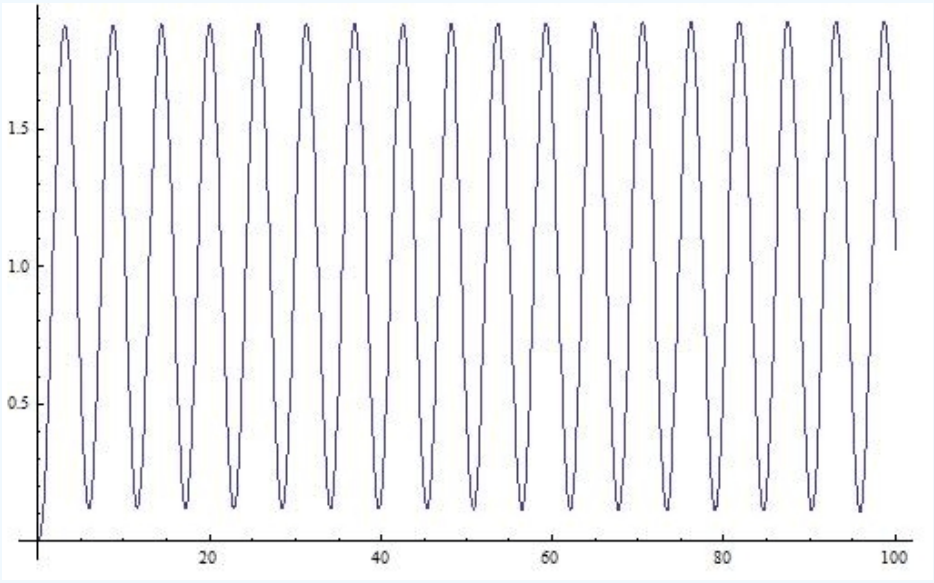
Example 9.6.5

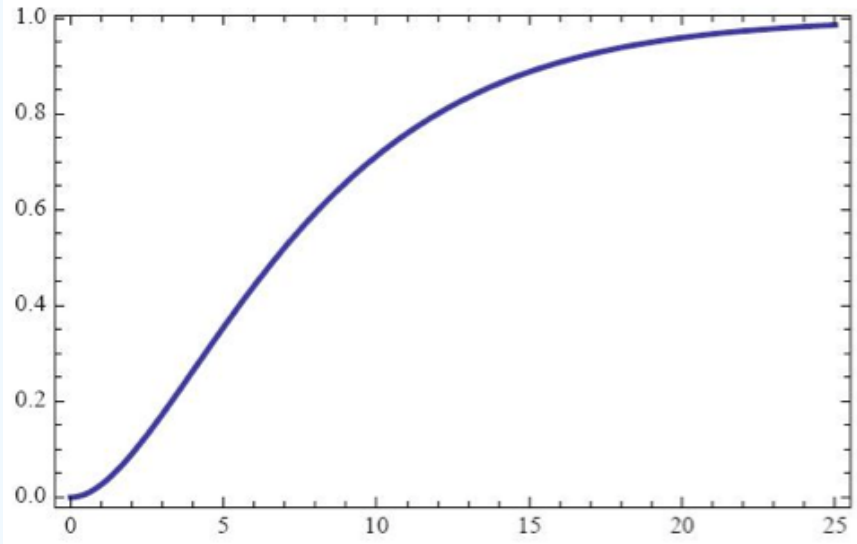
Label the following general response charts as overdamped, critically damped, underdamped, undamped, and unstable.

(a)

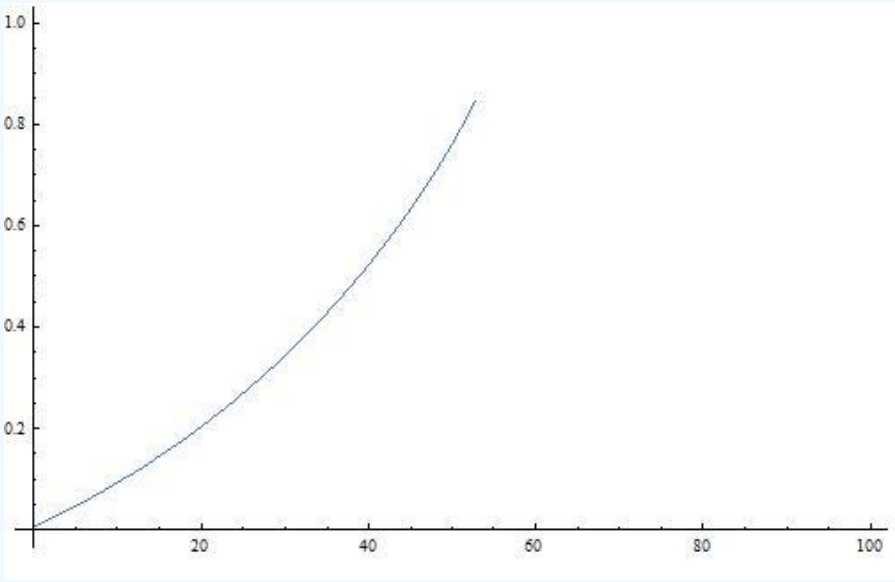


(b)

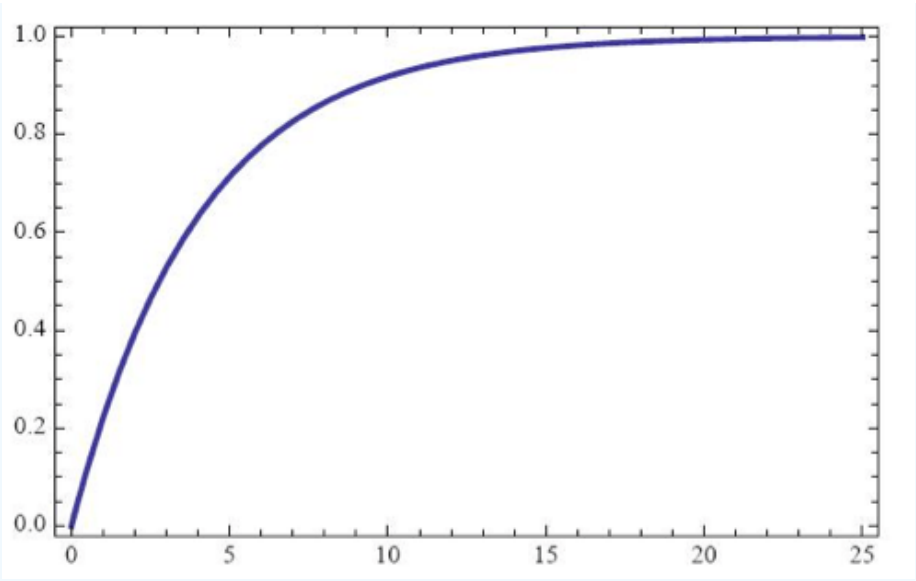




(c)



(d)

(e) 

Answers

a)Underdamped b)Undamped c)Critically Damped d)Unstable e)Overdamped

Exercise 9.6.1

What is a back-calculation used for?

a. To reduce noise problems
b. To reduce steady-state errors
c. To reduce errors in integral term
d. To reduce errors in proportional term

Exercise 9.6.1

What is a proportional band?

a. A way to eliminate errors in integral term
b. A way to eliminate errors in proportional term
c. A way to eliminate errors in derivative term
d. A way to understand anti-windup problem

References

- Bhattacharyya, Shankar P., Datta & Silva, PID Controllers for Time-Delay Systems, 2005 [1]
- Astrom, Karl J., Hagglund, Tore., Advanced PID Control, ISA, The Instrumentation, Systems and Automation Society.
- U of Texas chE
- Wikipedia Article, PID Controller, en.Wikipedia.org/wiki/PID_control

Contributors

- Authors: Ashwini Miryala, Kyle Scarlett, Zachary Zell, Brandon Kountz
- Stewards: Brian Hickner, Lennard Gan, Addison Heather, Monique Hutcherson

This page titled 9.6: PID Downsides and Solutions is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

10: DYNAMICAL SYSTEMS ANALYSIS

10.1: Finding fixed points in ODEs and Boolean models
10.2: Linearizing ODEs
10.3: Eigenvalues and Eigenvectors
10.4: Using Eigenvalues and Eigenvectors to Find Stability and Solve ODEs
10.5: Phase Plane Analysis - Attractors, Spirals, and Limit cycles
10.6: Root Locus Plots - Effect of Tuning
10.7: Routh Stability - Ranges of Parameter Values that are Stable

This page titled 10: Dynamical Systems Analysis is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

10.1: Finding fixed points in ODEs and Boolean models

Authors: Nicole Blan, Jessica Nunn, Pamela Anne Roxas, Cynthia Sequerah
Stewards: Matthew Kerry Braxton-Andrew, Josh Katzenstein, Soo Kim, Karen Staubach

Introduction

Engineers can gain a better understanding of real world scenarios by using various modeling techniques to explain a system's behavior. Two of these techniques are ODE modeling and Boolean modeling. An important feature of an accurate ODE model is its fixed point solutions. A fixed point indicates where a steady state condition or equilibrum is reached. After locating these fixed points in a system, the stability of each fixed point can be determined (see subsequent Wikis). This stability information enables engineers to ascertain how the system is functioning and its responses to future conditions. It also gives information on how the process should be controlled and helps them to choose the type of control that will work best in achieving this.

Concept Behind Finding Fixed Point

A fixed point is a special system condition where the measured variables or outputs do not change with time. In chemical engineering, we call this a steady state. Fixed points can be either stable or unstable. If disturbances are introduced to a system at steady state, two different results may occur:

1. the system goes back to those original conditions (stable point)
2. the system deviates from those conditions rapidly (unstable point)

Subsequent wiki articles will discuss these different types of fixed points in more detail. The focus of this article will be simply finding fixed points, not classifying them. We will discuss several methods of finding fixed points, depending on the type of model employed.

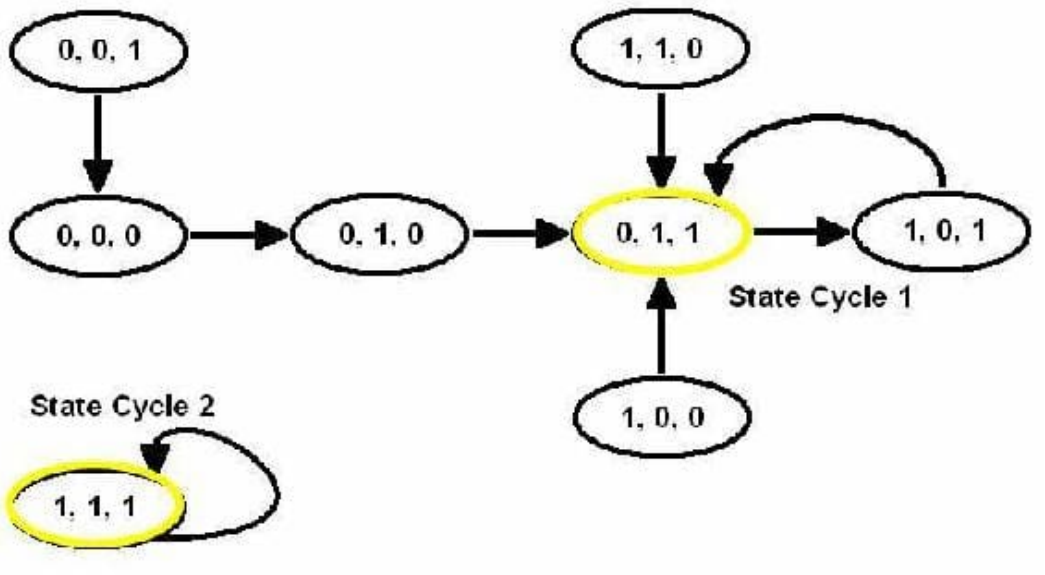
ODE Model

When a process or system is modeled by an ODE or a set of ODEs, the fixed points can be found using various mathematical techniques, from basic hand calcuations to advanced mathematical computer programs. Independent of the method used, the basic principle remains the same: **The ODE or set of ODEs are set to zero and the independent variables are solved for.** At the points where the differential equations equal zero there is no change occurring. Thus, the solutions found by setting the ODEs equal to zero represent the numerical values of independent variables (i.e. temperature, pressure, concentration) at steady state conditions. If a single ODE or set of ODEs becomes too complicated to be solved by hand, a mathematical program such as Mathematica can be used to find fixed points. The latter part of this article focuses on how to use Mathematica to find fixed points of complicated systems of ODEs.

Note that in some cases there may not be an analytical method to find a fixed point. This case commonly occurs when the solution to a fixed point involves a high degree polynomial or another mathematical function that does not have an analytical inverse. In these cases, we can still find fixed points numerically if we have the parameters.

Boolean Model

A Boolean Model, as explained in “Boolean Models,” consists of a series of variables with two states: True (1) or False (0). A fixed point in a Boolean model is a condition or set of conditions to which the modeled system converges. This is more clearly seen by drawing state transition diagrams.



State Transition Diagram from Boolean Models

From the state transition diagram above, we can see that there are two fixed points in this system: 0,1,1 and 1,1,1. Starting in any state on the diagram and following the arrows, one of these two states will be reached eventually, indicating that the system tends to achieve either of these sets of operating conditions. If slight disturbances are introduced to the system while it is operating at one of these sets of conditions, it will return to 0,1,1 or 1,1,1. Also noted in the state transition diagram are state cycles. The difference between a state cycle and a fixed point is that a state cycle refers to the entire set of Boolean functions and transition points leading to the steady-state conditions, whereas a fixed point merely refers to the one point in a state cycle where steady-state conditions are reached (such points are indicated by a yellow circle in the diagram).

Finding Fixed Points: Four Possible Cases

There are four possible scenarios when finding the fixed points of an ODE or system of ODEs:

1. One fixed point
2. Multiple fixed points
3. Infinite fixed points
4. No fixed points

One Fixed Point

The first type of ODE has only one fixed point. An example of such an ODE is found in the Modeling of a Distillation Column. An ODE is used to model the energy balance in the nth stage of the distillation column:

$$\frac{dT_n}{dt} = \frac{1}{M_W}\left[L_{n-1}x_{n-1} - Wx_W\right]\left[T_{n-1} - T_n\right] + \frac{q_r}{M_W c_p}$$

Which can also be written as:

$$\frac{dT_n}{dt} - \frac{1}{M_W}\left[L_{n-1}x_{n-1} - Wx_W\right]\left[T_{n-2}\right] + \frac{q_r}{M_W c_p} + \frac{1}{M_W}\left[L_{n-1}x_{n-1} - Wx_W\right]\left[-T_n\right]$$

If initial conditions i.e. T_{n-1},L_{n-1},x_{n-1} are known, the equation above reduces to:

$$\frac{dT_n}{dt} = a + bT_n$$

where a and b are constants since all the variables are now known.

$$a = \frac{1}{M_W}\left[L_{n-1}x_{n-1} - Wx_W\right]\left[T_{n-1}\right] + \frac{q_r}{M_W c_p}$$

$$b = \frac{-1}{M_W}\left[L_{n-1}x_{n-1} - Wx_W\right]$$

By analyzing the equation

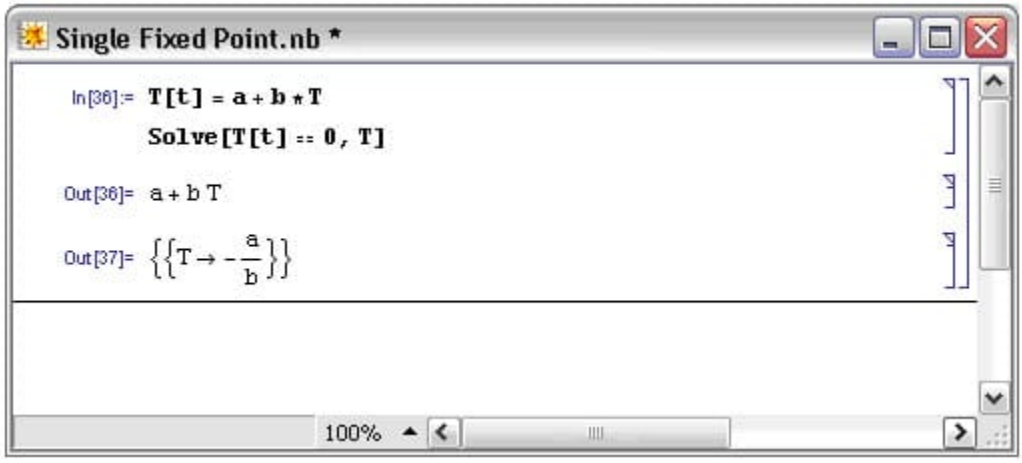
$$\frac{dT_n}{dt} = 0 = a + bT_n$$

we can immediately deduce that at steady state $T_n = -a/b$. Clearly, there is only one fixed point in this system, only one temperature of the distillation column which will be at steady-state conditions. We can use Mathematica to solve for the fixed point of this system and check our results. In Mathematica, the Solve[] function can be used to solve complicated equations and systems of complicated equations. There are some simple formatting rules that should be followed while using Mathematica:

1. Type your equation and let the differential be called an arbitrary variable (e.g. T[t])
2. Type Solve[T[t]==0,T] and hit Shift+Enter
3. This produces an output contained inside curly brackets

Please read the Solving ODEs with Mathematica section for more information on syntax and functions.

A sample of how the format in Mathematica looks like is shown below:

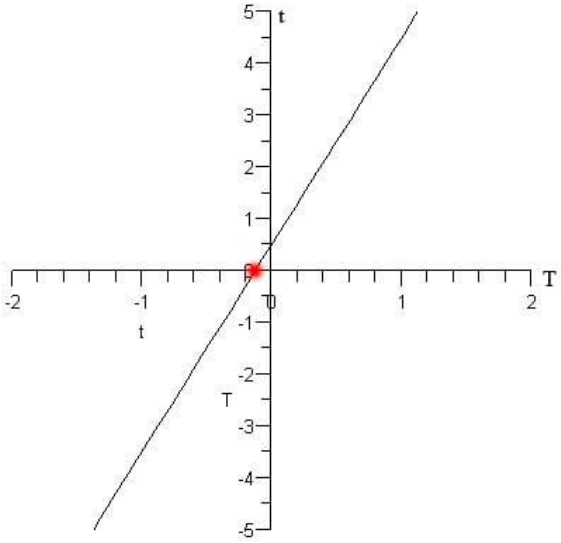


Maple can be used to visualize a single fixed point. Wherever the plot intersects the x-axis represents a fixed point, because the ODE is equal to zero at that point.

The following Maple syntax was used to plot the ODE: plot(0.5+4t, t=-2..2,T=0..5,color=black);

The constant a = 0.5 and the constant b = 4 in the above example.

The resulting graph is below, the red point indicates at what T a fixed point occurs:



Solving a single fixed point for an ODE and a controller in Mathematica

1. Identify what type of controller it is (P, I , PI, or PID etc.)
2. Identify your ODE equations (Is the controller a function of the ODE?)

Example: Solve for the fixed points given the three differential equations and the two controllers (u1 and u2).

Missing \left or extra \right

Extra \left or missing \right

$$\frac{dF_{out})}{dt} = K_{(v2}\Big)\,(u_2)\,H$$

Where H is the level in the tank, F_{in} is the flow in, F_{out} the flow out, and u_1 and u_2 are the signals to the valves v_1 and v_2. $Kv1$ and $Kv2$ are valve gains (assumed to be linear in this case, although this does not have to be). Note that the exit flow also depends on the depth of fluid in the tank.

You next parameterize your model from experimental data to find values for the constants:

A=2.5 meters squared

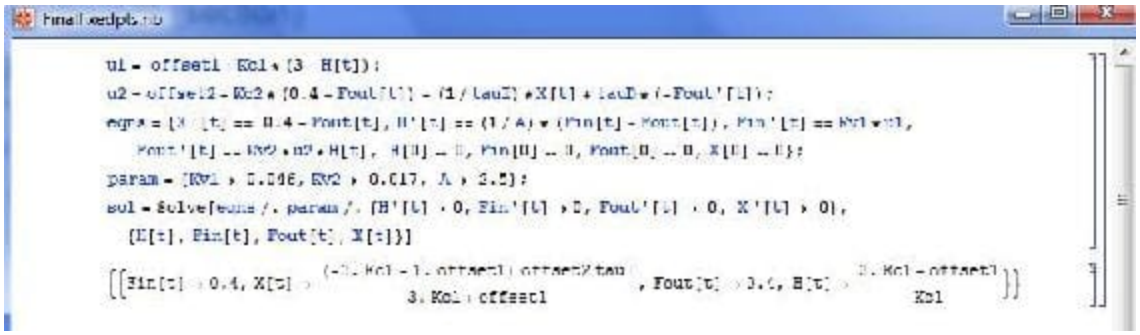
K_(v1)=0.046 meters cubed/(minute mA)

K_(v2)=0.017 meters squared/(minute mA)

Next you want to add:

- A full PID controller to regulate Fout via FC1 connected to v2.
- A P-only controller to regulate H via LC1 connected to v1.

For this system you want to maintain the tank level at 3 meters and the exit flow (Fset) at 0.4 m3 /minute. The following Mathematica code should look as follows:



Multiple Fixed Points

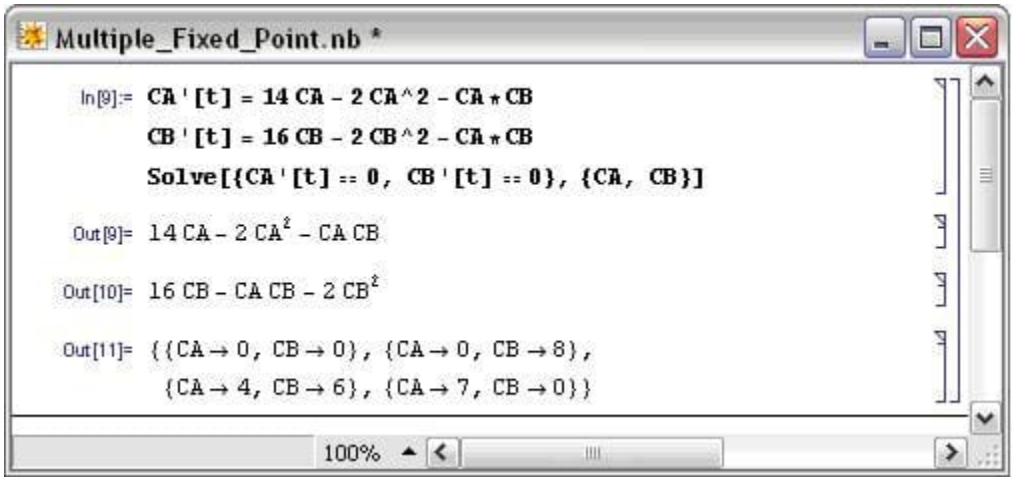
Multiple fixed points for an ODE or system of ODEs indicate that several steady states exist for a process, which is a fairly common situation in reactor kinetics and other applications. When multiple fixed points exist, the optimal steady-state conditions are chosen based on the fixed point's stability and the desired operating conditions of the system.

The following is an example of a system of ODEs with multiple fixed points:

$$\frac{dC_A}{dt} = 14C_A - 2C_A^2 - C_AC_B$$

$$\frac{dC_B}{dt} = 16C_B - 2C_B^2 - C_AC_B$$

The above system of ODEs can be entered into Mathematica with the following syntax:

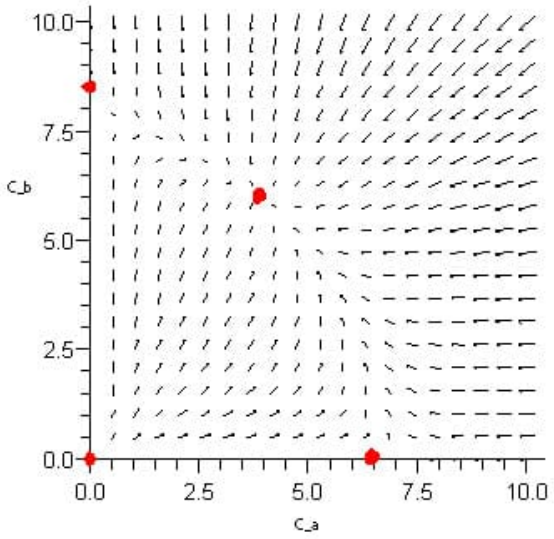


This system in particular has four fixed points. Maple can be used to visualize the fixed points by using the following syntax:

with(plots):

*fieldplot([14*x-2*x^2-x*y,16*y-2*y^2-x*y],x=0..10,y=0..10,fieldstrength=log);*

The first line initializes the plotting package within Maple that allows for plotting vector fields. The second line uses the command “fieldplot” and inputs the two ODEs that make up the system. The scales of the x and y-axis are set to range from 0 to 10. The fieldstrength command is mainly used for visual purposes, so that the direction of the arrows becomes more apparent. Below is the resulting plot:



The red dots indicate the fixed points of the system. On the plot, these points are where all the surrounding arrows converge or diverge. Converging arrows indicate a stable fixed point, in this example the point at (4,6) is a stable fixed point. Diverging arrows indicate an unstable fixed point, in this example (0,0), (0,8) and (7,0) are unstable fixed points.

Infinite Fixed Points

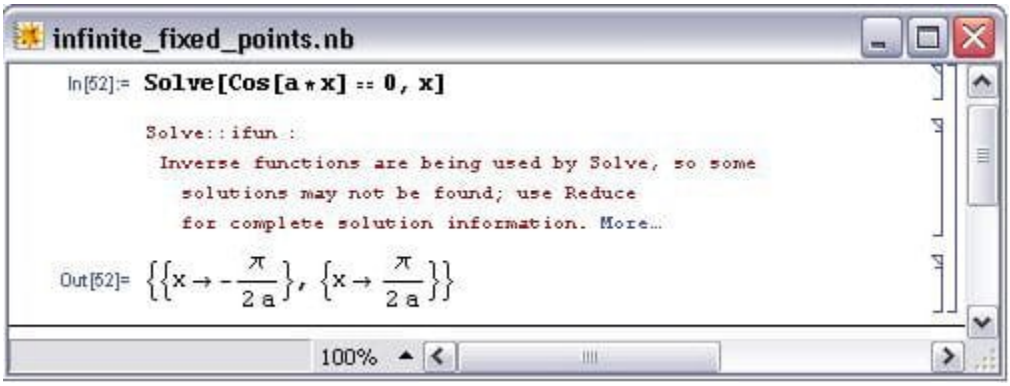
An example of an ODE with infinite fixed points is an oscillating ODE such as:

$$\frac{dy}{dx} = cos(ax)$$

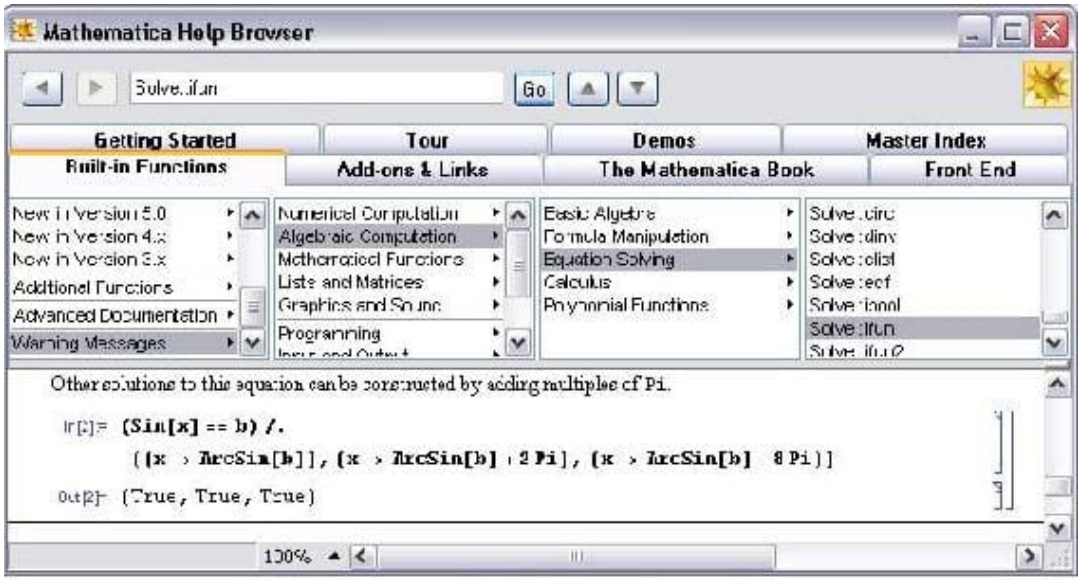
where *a* is a constant.

Using Mathematica to solve for the fixed points by setting

$$\frac{dy}{dx} = 0 = cos(ax)$$



If you click the "More" link on Mathematica it will basically state that there are other solutions possible according to the Help section shown below:

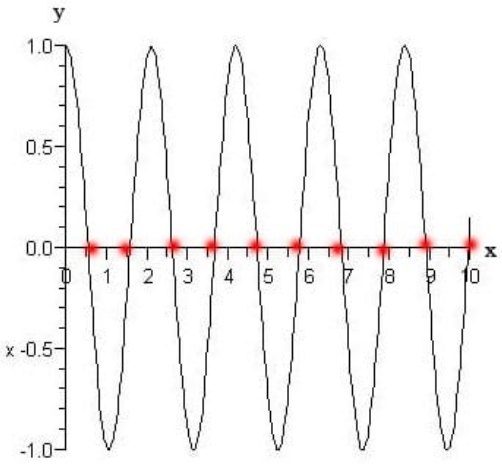


The Maple syntax used to graph the solved differential equation is:

plot(cos(3t),t=0..10,T=-1..1,color=black);

The constant a = 3 in this case.

The infinite fixed points can be seen in the graph below, where anytime the function crosses the x axis, we have a fixed point:



No Fixed Points

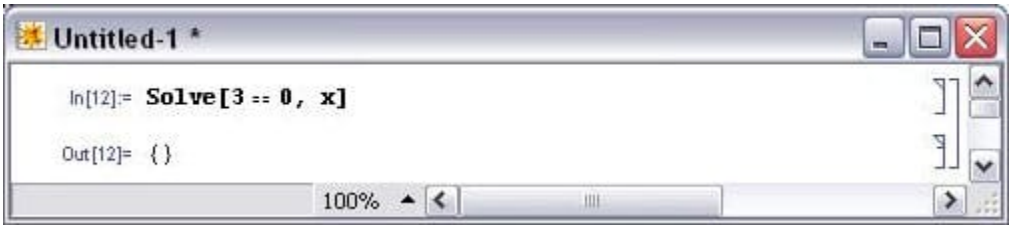
The fourth type of ODE does not contain fixed points. This occurs when a certain variable (such as temperature or pressure) has no effect on a system regardless of how it changes. Generally, systems with this sort of behavior should be avoided because they are difficult to control as they are always changing.

This can be modeled by vertical or horizontal lines due to the fact that no fixed points are found by setting the line equal to zero. An ODE is used to model a line held constant at a:

$$\frac{dT}{dt} = a$$

Where, a can be any constant except 0.

Intuitively, trying to find a fixed point in this system is not possible, because a constant such as 3 can never equal zero. Solving this ODE is not possible even by analyzing the system. Therefore, when inputting this into Mathematica, it yields {}. The notation {} means that there are no fixed points within the system. The image below is how Mathematica solves the ODE.

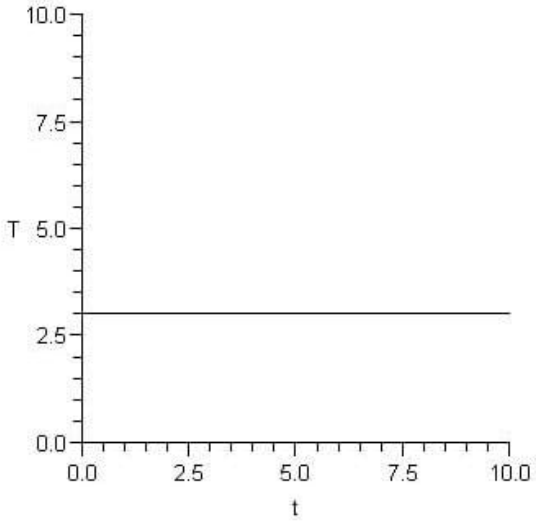


By using Maple (version 10), one can visually see a lack of fixed points by using the following syntax:

plot(3, t = 0..10, T = 0..10, color = black);

The constant a = 3 in the above case.

This image shows that the line is horizontal and never crosses the x axis, indicating a lack of fixed points.



Summary

A fixed point is a system condition where the measured variables or outputs do not change with time. These points can be stable or unstable; refer to Using Eigenvalues to evaluate stability for an introduction to a common method for determining stability of fixed points.

There are four possible cases when determining fixed points for a system described by ODEs:

- One fixed point
- Multiple fixed points
- Infitite fixed points
- No fixed points

There are methods described above for using Mathematica or Maple to solve for the fixed points in each case. Fixed points can also be determined for a Boolean model.

Knowing the fixed points of a system is very important when designing a control architecture for the system. These are the operating conditions that the system will exhibit at steady-state. Controllers can have influence on the fixed points, so a thorough analysis of fixed points using equations describing the system and the controllers should be conducted before implementation of the control scheme.

Worked out Example 1: Manipulating a System of Equations

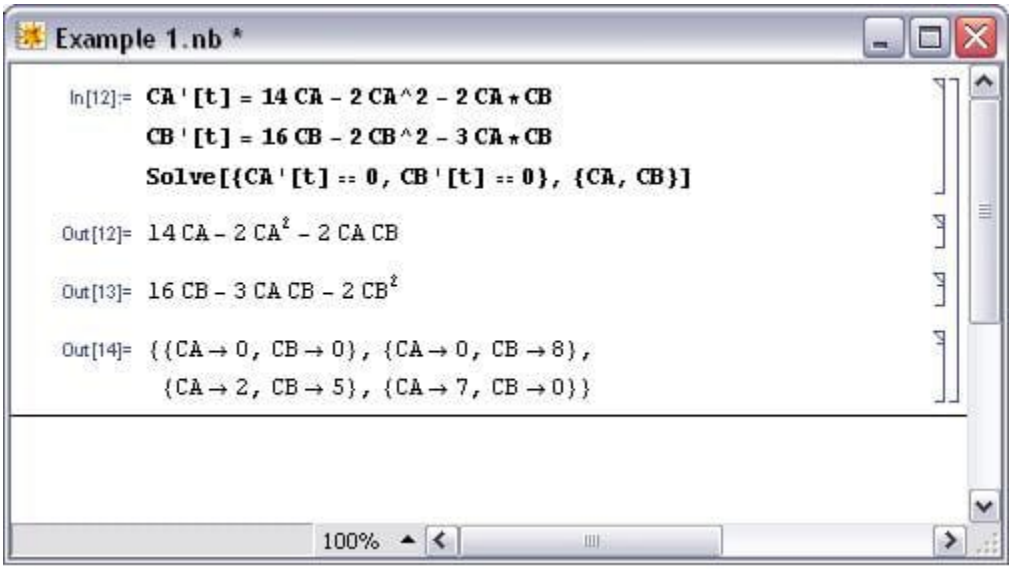
Recall the example system of ODEs used in the Multiple Fixed Points Section:

$$\frac{dC_A}{dt} = 14C_A - 2C_A^2 - mC_AC_B$$

$$\frac{dC_B}{dt} = 16C_B - 2C_B^2 - nC_AC_B$$

Find how the fixed points change when m = 2 and n = 3.

Solution



Worked out Example 2: System of ODEs

You are asked to estimate the fixed point(s)of a heat exchanger in your plant. The fixed points corresponds to the temperatures of your tube-side (hot) and shell-side (cold) fluids respectively, $T_{t,out}$ and $T_{s,out}$. Neglect heat loss to the surrounding area i.e. $hA(T_s - T_\infty) = 0$.

Given the two main ODEs used to model a heat exchanger, use Mathematica to solve for the fixed points of the system in terms of the known variables.

$$mc_p\frac{dT_{t,out}}{dt} = \rho c_p F_{t,in}T_{t,in} - \rho c_p F_{t,out}T_{t,out} - \frac{kA}{\Delta z}(T_{t,out} - T_{s,out}) \quad \text{(equation 1)}$$

$$mc_p\frac{dT_{s,out}}{dt} = \rho c_p F_{s,in}T_{s,in} - \rho c_p F_{s,out}T_{s,out} - \frac{kA}{\Delta z}(T_{s,out} - T_{t,out}) - hA(T_s - T_\infty)$$ (equation 2)

The values for m, c_p, ρ, $F_{t,in}$, $F_{t,out}$, k, A, δz, and $T_{t,in}$, $F_{s,in}$, $F_{s,out}$, $T_{s,in}$ are given and fixed.

Please refer to the Wiki article on HeatExchangeModel for detailed explanation on the meaning of the variables and the derivation of the ODEs above.

Hint: Lump up all known variables under one general variable

Solution

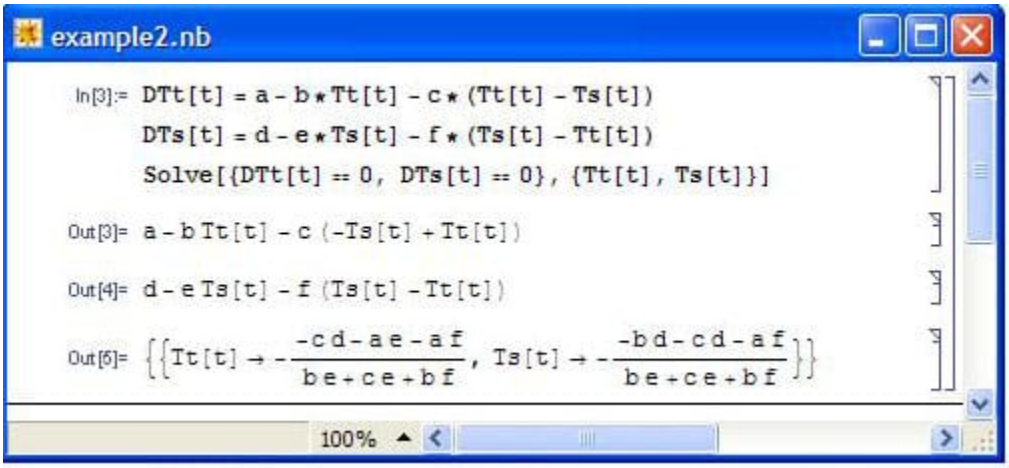
Simplify the 2 equations to the ODEs below:

$$\frac{dT_{t,out}}{dt} = a - bT_{t,out} - c(T_{t,out} - T_{s,out})$$ (equation 1a)

$$\frac{dT_{s,out}}{dt} = d - eT_{s,out} - f(T_{s,out} - T_{t,out})$$ (equation 2a)

where $a = \frac{\rho F_{t,in}T_{t,in}}{m}$, $b = \frac{-\rho F_{t,out}}{m}$, $c = \frac{kA}{\Delta z m c_p}$ and $d = \frac{\rho F_{s,in}T_{s,in}}{m}$, $e = \frac{\rho F_{s,out}}{m}$, $f = \frac{kA}{\Delta z m c_p}$

Type these equation into Mathematica using the appropriate syntax and use the Solve[] function to find the fixed points.



We have found our fixed point. Just plug in the variables as defined earlier for a, b, c, d, e, f and you will obtain the temperatures in terms of the useful parameters.

Multiple Choice Question 1

The solutions found by setting the ODE equal to zero represent:

a) independent variables not at steady state conditions

b) dependent variables not at steady state conditions

c) independent variables at steady state conditions

d) dependent variables at steady state conditions

Answer: C

Multiple Choice Question 2

How many fixed points are there when the following equation is solved by Mathematica?

$$\frac{dT}{dt} = 8t - 4t^2 - t^3$$

a) none

b) 1

c) 2

d) 3

Answer: D

Sage's Corner

Stability of Fixed Points
video.google.com/googleplayer...80917022362654

slides for this talk

References

- Edwards H., Penney D.(2003), *Differential Equations: Computing and Modeling*, Third Edition. Prentice-Hall. ISBN 0130673374
- Strogatz, Steven H.(2001), *Nonlinear Dynamic and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, 1st Edition. Addison-Wesley. ISBN 0738204536

This page titled 10.1: Finding fixed points in ODEs and Boolean models is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

10.2: Linearizing ODEs

Introduction

Chemical engineering processes often operate in nonlinear and unsteady manners (i.e. not always at steady state), and are generally governed by nonlinear ordinary differential equations (ODEs). The ODE is a relation that contains functions of only one independent variable and derivatives with respect to that variable. Many studies have been devoted to developing solutions to these equations, and in cases where the ODE is linear it can be solved easily using an analytical method. However, if the ODE is nonlinear and not all of the operating parameters are available, it is frequently difficult or impossible to solve equations directly. Even when all the parameters are known, powerful computational and mathematical tools are needed to completely solve the ODEs in order to model the process. In order to simplify this modeling procedure and obtain approximate functions to describe the process, engineers often linearize the ODEs and employ matrix math to solve the linearized equations.

A linear equation is an equation in which each term is either a constant or the product of a constant times the first power of a variable. These equations are called "linear" because they represent straight lines in Cartesian coordinates. A common form of a linear equation in the two variables x and y is $y = mx + b$. This is opposed to a nonlinear equation, such as

$$m = e^x + x^2 + 2x + 5.$$

Even though $2x + 5$ is a linear portion of the equation, e^x and x^2 are not. Any nonlinear terms in an equation makes the whole system nonlinear.

Non-linear system of equations:

$$\frac{dA}{dt} = 3A^2 + 2B + C - 7D^3$$
$$\frac{dB}{dt} = A + C^2 + 2D$$
$$\frac{dC}{dt} = A + 4B^2 - C^2$$
$$\frac{dD}{dt} = 2C - D$$

After linearization (around the steady state point {-0.47,-0.35,0.11,0.23}:

$$\begin{pmatrix} A' \\ B' \\ C' \\ D' \end{pmatrix} = \begin{pmatrix} -2.83 & 2 & 1 & 0 \\ 1 & 0 & 0.23 & 0 \\ 1 & -2.78 & -0.23 & 0 \\ 0 & 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} + \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix}$$

Note that each equation is comprised solely of first order variables.

Even though it is unlikely that the chemical engineering process to be modeled operates in a linear manner, all systems can be approximated as linear at a point. This is preferred as linear systems are much easier to work with than nonlinear equations. Although linearization is not an exact solution to ODEs, it does allow engineers to observe the behavior of a process. For example, linearized ODEs are often used to indicate exactly how far from steady state a given process deviates over specified operating ranges. This wiki page discusses how to solve a linearized ODE by hand and by using Mathematica, and proceeds to work out several examples of linearized ODEs commonly seen in chemical engineering practice.

Applications to Chemical Engineering

As mentioned above, linearizing ODEs allows engineers to understand the behavior of their system at a given point. This is very important because many ODEs are impossible to solve analytically. It will also lead to determining the local stability of that point. Most of the time a system will be linearized around steady state, but this is not always the case. You may be interested

in understanding the behavior of your system at its operating point or equilibrium state (not necessarily steady state). The linearization approach can be used for any type of nonlinear system; however, as a chemical engineer, linearizing will usually involve ODEs. Chemical engineers use ODEs in applications such as CSTRs, heat exchangers, or biological cell growth.

It is also important to understand the advantages and disadvantages of linearizing a system of ODEs:

Advantages

- Provides a simpler, more convenient way to solve the ODEs
- The behavior of a process can be observed
- Any type or order of ODE can be used

Disadvantages

- The solution is only an exact solution at the chosen point; otherwise it is an approximation and becomes less accurate away from the point
- Although linearizing is a quicker alternative, it takes time to initially learn the process (ex: using Mathematica)

Another use for linearization of the equations that govern chemical processes is to determine the stability and characteristics of the steady states. Systems of linearized ODEs can be used to do this, and the methods of doing so can be found in Fixed Points, Eigenvectors and Eigenvalues, and Using eigenvalues and eigenvectors to find stability and solve ODEs

General Procedure for Linearization

Linearization is the process in which a nonlinear system is converted into a simpler linear system. This is performed due to the fact that linear systems are typically easier to work with than nonlinear systems. For this course, the linearization process can be performed using Mathematica. The specific instructions on how to do this can be found below.

1. Choose a relevant point for linear approximation, two options available are:
   - Steady state- points where system does not change
   - Current location- given where you are now
2. Calculate the Jacobian matrix at that point. The Jacobian is essentially a Taylor series expansion.
3. Solve to find unknown constants using algebraic methods.

Linearization by Hand

In order to linearize an ordinary differential equation (ODE), the following procedure can be employed. A simple differential equation is used to demonstrate how to implement this procedure, but it should be noted that any type or order of ODE can be linearized using this procedure.

1. Use a Taylor series expansion (truncating after the linear terms) to approximate the right-hand side of the ODE.

Let's say we start with the following ODE:

$$\frac{dx}{dt} = f(x) = 3x^2.$$

This ODE describes the behavior of some variable, x, with respect to time.

A Taylor series is a series expansion of a function about a point. If $x = a$, an expansion of a real function is given by:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^n(a)}{n!}(x-a)^n$$

$$f(x) = f(a) + f'(a)(x-a) + \frac{f^2(a)}{2!}(x-a)^2 + \frac{f^3(a)}{3!}(x-a)^3 - ... + \frac{f^n(a)}{n!}(x-a)^n + ...$$

When $x = 0$, the function is also known as Maclaurin series. Taylor's theorem states that any function satisfying certain conditions can be expressed as a Taylor series.

For simplicity's sake, only the first two terms (the zero- and first-order) terms of this series are used in Taylor approximations for linearizing ODEs. Additionally, this truncation (ie. "chopping" off the $n = 2$ and higher terms from the polynomial shown above) assures that the Taylor Series is a linear polynomial. If more terms are used, the polynomial would have $(x - a)^2$ and higher order terms and become a nonlinear equation. The variable 'a' in the Taylor series is the point chosen to linearize the function around. Because it is desired that most processes run at steady state, this point will be the **steady state point**. So, our differential equation can be approximated as:

$$\frac{dx}{dt} = f(x) \approx f(a) + f'(a)(x - a) = f(a) + 6a(x - a)$$

Since a is our steady state point, $f(a)$ should always be equal to zero, and this simplifies our expression further down to:

$$\frac{dx}{dt} = f(x) \approx f'(a)(x - a) = 6a(x - a)$$

The graph shown above shows the approximation of f(x) at (x,f(x)). As mentioned previously, linearization is only an approximation for any given function near a continuous point. When working with a system of ODEs, the **Jacobian** is written as a matrix. It is the matrix of constants needed to describe a system's linearity. The Jacobian may be thought of as how much a system is distorted to take on a linear identity. A Jacobian matrix will always be a square(#rows = #columns) and it shows how each equation varies with each variable. The Jacobian matrix is defined as:

$$J(x_1, \ldots, x_n) = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_n} \end{bmatrix}$$

And is used as such for this non-linear set of equations

$$\begin{aligned} \tfrac{dx}{dt} &= 2x^2 + \cos(5y) - 12\arctan(z - 3) \\ \tfrac{dy}{dt} &= 7y^3 + \sin(5x) \\ \tfrac{dz}{dt} &= 1.3\log(x + y) \end{aligned}$$

The linearized system is

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} J_1 & J_2 & J_3 \\ J_4 & J_5 & J_6 \\ J_7 & J_8 & J_9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

The square matrix is the Jacobian.

Example 10.2.1

Lets say you have the following set of equations and you want to find its Jacobian matrix around the point \(A=3,\,B=2).

$$\frac{dA}{dt} = 3A - A^2 - AB$$

$$\frac{dB}{dt} = 6B - AB - 2B^2$$

We find the Jacobian by taking the derivative of each equation with respect to each variable.

$$\frac{d(3A - A^2 - AB)}{dA} = 3 - 2A - B$$

$$\frac{d\left(3A - A^2 - AB\right)}{dB} = -A$$

$$\frac{d\left(6B - AB - 2B^2\right)}{dA} = -B$$

$$\frac{d\left(6B - AB - 2B^2\right)}{dB} = 6 - A - 4B$$

These are the equations in the matrix. The values of the variables from whatever point we are linearizing are then put into these equations and calculated out to get the Jacobian.

$$Jac = \begin{bmatrix} 3 - 2A - B & -A \\ -B & 6 - A - 4B \end{bmatrix} A = 4, B = 2$$

$$Jac = \begin{bmatrix} -7 & -4 \\ -2 & -6 \end{bmatrix}$$

2. Change the approximation by linearizing around a steady state point in order to describe how the process deviates from steady state.

The following substitution can be made:

$$\frac{dx}{dt} = \frac{d(x-a)}{dt}$$

$$\frac{d(x-a)}{dt} \approx 6a(x-a)$$

This substitution is allowed because 'a' is a constant, and the derivative of a constant is zero.

Substituting (x-a) for x signifies that our differential equation now shows how our function, x, deviates away from the steady state value, a, with respect to time. This deviation, (x-a), is commonly expressed as x'. It should also be noted that the quantity '6a' is a constant, and thus will be further recognized as 'A'.

Our final linearized equation becomes:

$$\frac{dx'}{dt} \approx Ax'$$

The once nonlinear ODE,

$$\frac{dx}{dt} = f(x) = 3x^2$$

has now been simplified into a linear differential equation.

The procedure of linearization typically occurs around the steady state point or points of a specified process. Engineers anticipate a certain change in output for the particular steady state point, and may proceed to linearize around it to complete their approximation. Please follow the graphic below for further detail. Note that a steady state point occurs when $dx/dt = 0$

$$\backslash \text{rror} PlotColorNew1.jpg$$

As can be seen, moving farther away from the steady state point results in significantly larger deviation, and thus error (actual-linearization).

Occasionally, for very unique operating conditions, plant management may decide to momentarily run a process outside of strict steady state conditions (perhaps a unique start-up procedure, shut-down recovery, etc.). To reinforce the concept of linearization around an unsteady state point (arbitrarily chosen in this example), please consider the following visual representation. Note that a=50 and C=7500 in the first plot.

$$\backslash \text{ew} Image2.jpg$$

The second plot is a magnification of a small section of the first plot. As one can readily notice, both the linear approximation and the actual graph overlap almost exactly over this small range. This illustrates how this particular linearization could be used to approximate this region of the function and can describe its behavior quite accurately. This approximating technique using the linearization of ODE can be performed around different points (different values of a) using the same method in order to model the behavior of the rest of the dx/dt vs x function. Note: frequent recalculation of the integration constant, C, permits for increased accuracy in approximation.

Example of a Simple Linearization Process in Use

Provided a circumstance in which developing a rigorous analytical solution is unfeasible due to time, access to computing resources, or mathematical ability, the linearization process offers a convenient and swift alternative.

An appropriate introduction to linearization application is featured in section 5.3 of the Bequette reading and has been described below.

The following model describes the behavior of tank height as a function of time (t), faced with a steady- state incoming liquid flowrate. dh/dt represents the change in height of the fluid contained within the tank as a function of time, h indicates the height (or level) or the fluid contained within the tank, F represents the magnitude of incoming flow. The following model describes the tank level/height fluctuation as function of both the current height (h) and flow (F). Fluid enters the tank as flow per area of the inlet(F/A), and leaves as a function of a corrective proportional term, β, also per area of the outlet, multiplied by the square root of the current height.

Instead of fluid leaving through this term, we may also consider it a form of feedback control, which subsequently reduces the amount of inlet fluid instead (depends on the application, the final result is identical). Therefore:

$$\frac{dh}{dt} = f(h,F) = \frac{F}{A} - \frac{\beta}{A} * \sqrt{h}$$

Important note about the relevance of the provided variables:

F_S indicates the initial flow rate for the described system and is important in calculating the intercept and subsequently the calculative corrective term for the linearization (if the linearization is not performed at steady-state, as in this example). h_S describes the tank height for the specified flow rate, F_S. Upon plugging in the value of h_S into the linearized formula, you will notice that f(h,F_S) = 0, the point on the plot where both the linearization and the characteristic equation coincide (i.e. are tangent).

h remains a variable for the derivation so that students may notice that it is a dependant variable (dependent upon the flowrate). Bequette must (and does) provide both the values for F_S, h_S so that a linearization may proceed (otherwise we will simply have an equation describing the slope with no characteristic point to pass through).

System Parameters (F_S is an example flow rate at steady state, this information is used to determine the intercept of the linearized formula):

$$A = 1ft^2; h_s = 5ft; \beta = \frac{1}{\sqrt{5}}\frac{ft^{2.5}}{min}; F_s = \frac{1ft^3}{min}$$

Rewriting the system with the input variables:

$$\frac{dh}{dt} = f\left(h,F_s\right) = 1 + \frac{1}{\sqrt{5}} * \sqrt{h}$$

To linearize around a certain point, simply evaluate the derivative of the desired function and add in a corrective constant, C, represented by the value of the function at the initial (specified) condition.

$$\(h,F_s) \approx f(h,F_s) + \frac{\partial f}{\partial h}|_{h,F_s}\left(h - h_s\right)$$

\[\(h,F_s) \approx 0+ \frac{-1}

ParseError: EOF expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

|_h, _{F_s}(h-h_s) \nonumber \]

Selecting, h=5 to linearize around, we present an adapted version of the text example (linearization is in red):

\[\(h,1) \approx 0+ \frac{-1}

ParseError: EOF expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

|_h, _{F_s}(h-5) \nonumber \]

\[\(5,1) \approx 0+ \frac{-1}

ParseError: EOF expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

|_h, _{F_s}(5-5)= 0 \nonumber \]

f**(**h**,1)**approx**. = − 0.1 * (**h **− 5)**

\isual3.*jpg*

The physical significance of this example:

The model describes a situation where a tank maintains the steady state set point, a level (height) of 5 feet, by regulating the input (and/or output) flow rate(s). A height above the preferred level of 5 ft. results in a negative height change per time; conversely, a height below the preferred level of 5 ft. results in a positive height change per time, so as to always approach the steady state level.

Also, please note that the linearized system responds relatively slower than the characteristic equation system (non- linear)when tank level falls below 5 ft. and faster than the characteristic equation system when the tank level is above 5 ft.

Please read on below to learn a step-by-step technique on how to complete this method for most equations.

Linearization using Mathematica

Most ODEs are not as simple as the example worked out above. For those that can be solved analytically, the 'by-hand' method from the previous section can be used, although the math becomes tricky and tedious when one is dealing with systems of ODEs. For those that cannot be solved analytically and can only be approximated numerically, it is sometimes impossible to do the numerical analysis by hand. For these ODEs, powerful computational tools such as Mathematica can be used to perform linearization.

The following Mathematica file contains a worked out example (CSTR w/ heat exchange). This example contains multiple ODEs and multiple types of variables. When dealing with ODEs, there are three main type of variables: state, input, and output. State variables describes the system at any given time (ie. Ca, Cb, T, etc.). Input variables are simply inputs into the system. Output variables are those that are not directly related to the inputs, but are dependent on the state variables (ie. k, the reaction constant). You will discover in the CSTR example that it is necessary to linearize state variables and output variables separately.

Before proceeding to the Mathematica file we will first present some of the Mathematica commands that will be necessary to linearize the CSTR example.

Note that you need to press shift+enter (shift+return on Mac) after a command to evaluate the expression.

Example:

In[1]= ((0.04587*54.32497+4.59058433734947^7)/48700043)^131.94564

Out[1]= 0.000411241

Another thing to note is the use of assignment and equality operators.

= is called **immediate assignment operator** and **:=** is called **delayed assignment operator**. When immediate assignment operator is used, the right hand side of the operator is evaluated every time an assignment is made. However, when delayed assignment operator is used, the right hand side is evaluated only when the value of the expression on the left hand side is requested.

lhs=rhs rhs is intended to be the "final value" of lhs (e.g., f[x_]=1-x^2) lhs:=rhs rhs gives a "command" or "program" to be executed whenever you ask for the value of lhs (e.g., f[x_]:=Expand[1-x^2])

p := x^3 - 6 x^2 + 11 x - 6

This input, which produces no output, assigns the polynomial expression to the variable p. The assignment is done because this polynomial may be used again later, so the assignment saves retyping it. The **equality operator ==**, which is used here to form the equation we are asking Mathematica to solve:

In[2]:= Solve[p == 0]

Out[2]=

ParseError: invalid Primary (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

When solving nonlinear ODEs using Mathematica, it is necessary to form a matrix and there are several commands that can be used to create matrices.

D[X,Y]

This command takes the partial derivative of an expression (X) with respect to a variable (Y) defined by the user. The first term in the square brackets is the experession you wish to differentiate, the second is the variable you wish to differentiate with respect to.

AppendRows[column1,column2]

This command creates a matrix by combining two columns. In order to use this command you must first load the Matrix Manipulation package using the command <

M =

ParseError: EOF expected (click for details)

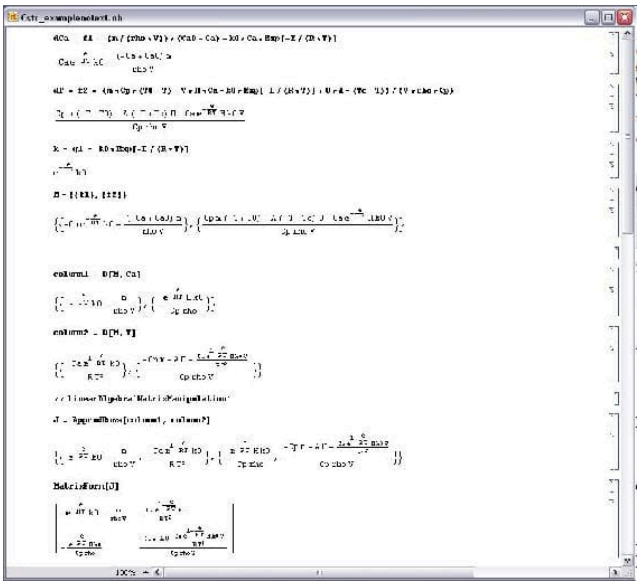
```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

This command creates a matrix, in this case a 3x3 matrix.

MatrixForm[X]

This command displays a specified matrix in matrix form.

These commands will allow us to proceed with the linearization of the CSTR problem shown below.



The first three commands input the equations that govern the behavior of the concentration, Ca, the temperature, T, and the rate of reaction, k. Once the equations have been entered into Mathematica a matrix, M, with one column and two rows containing the equations for Ca and T is created. Next two new matrices are created by taking the derivative of matrix M is with respect to both Ca and T. These matrices, called column1 and column2 respectively, together form the Jacobian of the system and is then displayed in matrix form. The **Jacobian matrix** is the matrix of all first-order partial derivatives of a vector-valued function. Its importance lies in the fact that it represents the best linear approximation to a differentiable function near a given point. For example, given a set y=f(x) of n equations in variables x_1, x_2,...,x_n, the Jacobian matrix is defined as following:

$$\acobianmatrix.gif$$

$$\athematicapg2.JPG$$

Now that we have the Jacobian we need to create a matrix containing the deviation from steady state for each of the variables. This matrix, SS, contains the actual concentration and temperature, C and T, minus the steady state concentration and temperature, Cas and Ts. The matrix is then displayed in matrix form. We now have both the Jacobian and the deviation matrix for the state variables. The next four commands create the Jacobian and deviation matrix for the output variable, k. The first command creates the Jacobian matrix by taking the derivative of the k equation with respect to Ca and T. The Jacobian is then shown in matrix form. Finally the deviation matrix for k is created in the same manner as above and then displayed in matrix form. Note that because k is defined above, this expression is substituted in for k in the deviation matrix. The following Mathematica file contains the code shown above with extra comments explaining why each step is performed Media:cstr_example.nb. It may be useful to downolad this file and run the program in Mathematica yourself to get a feel for the syntax. Downloading the file will also allow you to make any changes and edits to customize this example to another example of interest.

Note: This file needs to be saved to your computer, and then opened using Mathematica to properly run.

To see another example lets linearize the first 4 differential equations given in the introduction section.

Non-linear system of equations:

$$\frac{dA}{dt} = 3A^2 + 2B + C - 7D^3$$

$$\frac{dB}{dt} = A + C^2 + 2D$$

$$\frac{dC}{dt} = A + 4B^2 - C^2$$

$$\frac{dD}{dt} = 2C - D$$

After linearization (around the steady state point {-0.47,-0.35,0.11,0.23}:

$$\begin{pmatrix} A' \\ B' \\ C' \\ D' \end{pmatrix} = \begin{pmatrix} -2.83 & 2 & 1 & -1.10 \\ 1 & 0 & 0.23 & 2 \\ 1 & -2.78 & -0.23 & 0 \\ 0 & 0 & 2 & -1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} + \begin{pmatrix} 0.50 \\ 0.01 \\ 0.47 \\ 0 \end{pmatrix}$$

The 4 differential equations above are added into a Mathematica code as “eqns” and “s1” is the fixed points of the differentials. The steady state values found for “a, b, c, and d” are called "s1doubleBrackets(7)” After the steady state values are found, the Jacobian matrix can be found at those values.

To find “k1, k2, k3, and k4” the constants of the Linearization matrix equation, “m1” must be defined, which is the 2nd matrix on the right-hand side of the Linearization matrix equation.

To determine the k values (in matrix form), execute the dot product of "m1" and the “Jac” matrix, which is done by the "." operator. Therefore it should look like "Jac.m1"

To obtain the k values, determing the "Jac.m1" at the steady state values, which is done by the "/." operator. Therefore it should look like “Jac.m1/.s1doubleBrackets(7)”

Example 10.2.1

Let's say you have a system governed by the following system of equations:

$$\frac{dX_a}{dt} = 3X_a^2 + 2X_b + F_o$$
$$\frac{dX_b}{dt} = 6X_a + 9X_b^2 + X_aX_b + F_o$$
$$Y = X_a^2 + 3 * F_o^2$$

In this case, X_a and X_b are state variables (they describe the state of the system at any given time), F_o is an input variable (it enters the system and is independent of all of the state variables), and Y is the output of the system and is dependent on both the state and input variables. Please linearize this system using Mathematica.

Example 10.2.2

Use the Mathematica file presented in the article to linearize the CSTR example presented in the ODE & Excel CSTR model with heat exchangewiki around the steady state (T = 368.4K and Ca = 6.3mol/L). Use the constants presented in their example. Use the linearization to approximate Ca at 369K.

Solution

The linear approximation gives concentration of a (Ca) of 7.13 mol/L. The Mathematica file used to solve this problem can be downloaded here. Media:LinearizationExample.nb If you are having trouble running the Mathematica file try right clicking the link and using the Save Target As option.

Exercise 10.2.1

In an ODE linearization, what point in the process is the linearization generally centered around?

a. Starting point
b. Ending point
c. Steady-state point
d. Half-way point

Answer available at the end of this section.

Exercise 10.2.2

What does the D[] function in Mathematica do?

a. Find the determinant of a matrix
b. Finds the partial derivative a function or functions
c. Find the dot product of two matrices
d. Integrate a matrix or function

Answer available at the end of this section.

References

- Bequette, B. Wayne. *Process Dynamics: Modeling, Analysis, and Simulation*. Prentice- Hall PTR, Upper Saddle River, NJ 07458 (c) 2006.
- Kravaris, Costas. *Chemical Process Control: A Time Domain Approach, Appendix A*. Department of Chemical Engineering, University of Michigan, Ann Arbor, MI.
- Bhatti, M. Asghar. *Practical Optimization Methods With Mathematica Applications*. Springer Telos.

Answers

Exercise 10.2.1

Answer: c

Exercise 10.2.2

Answer: b

This page titled 10.2: Linearizing ODEs is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Dallas Burkholder, Navin Raj Bora, Nina Mohan, & Sarah Tschirhart via source content that was edited to the style and standards of the LibreTexts platform.

10.3: Eigenvalues and Eigenvectors

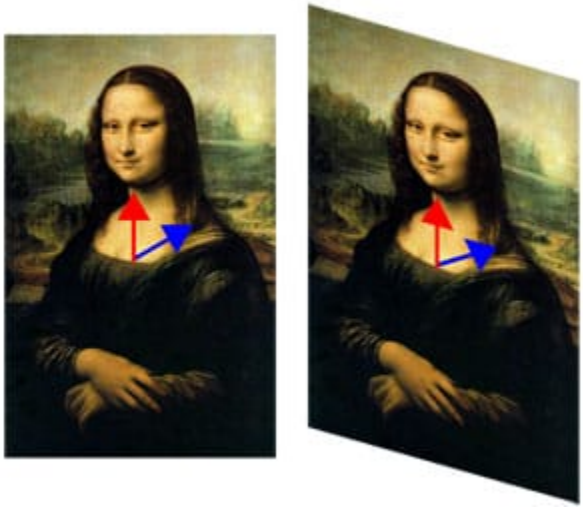
What are Eigenvectors and Eigenvalues?

Eigenvectors ($\mathbf{V}$) and Eigenvalues (λ) are mathematical tools used in a wide-range of applications. They are used to solve differential equations, harmonics problems, population models, etc. In Chemical Engineering they are mostly used to solve differential equations and to analyze the stability of a system.

Definition: Eigenvector and Eigenvalues

- An **Eigenvector** is a vector that maintains its direction after undergoing a linear transformation.
- An **Eigenvalue** is the scalar value that the eigenvector was multiplied by during the linear transformation.

Eigenvectors and Eigenvalues are best explained using an example. Take a look at the picture below.



In the left picture, two vectors were drawn on the Mona Lisa. The picture then under went a linear transformation and is shown on the right. The red vector maintained its direction; therefore, it's an eigenvector for that linear transformation. The blue vector did not maintain its director during the transformation; thus, it is not an eigenvector. The eigenvalue for the red vector in this example is 1 because the arrow was not lengthened or shortened during the transformation. If the red vector, on the right, were twice the size than the original vector then the eigenvalue would be 2. If the red vector were pointing directly down and remained the size in the picture, the eigenvalue would be -1.

Now that you have an idea of what an eigenvector and eigenvalue are we can start talking about the mathematics behind them.

Fundamental Equation

The following equation must hold true for Eigenvectors and Eigenvalues given a square matrix A:

$$\mathrm{A} \cdot \mathrm{v} = \lambda \cdot \mathrm{v} \tag{10.3.1}$$

where:

- A is a square matrix
- v is the Eigenvector
- λ is the Eigenvalue

Let's go through a simple example so you understand the fundamental equation better.

Example 10.3.1

Is $\mathbf{v}$ an eigenvector with the corresponding $\lambda = 0$ for the matrix $\mathbf{A}$?

$$\mathbf{v} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 6 & 3 \\ -2 & -1 \end{bmatrix}$$

Solution

$$A \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$$

$$\begin{bmatrix} 6 & 3 \\ -2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -2 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Therefore, it is true that $\mathbf{v}$ and $\lambda = 0$ are an eigenvector and eigenvalue respectively, for $\mathbf{A}$. (See section on Matrix operations, i.e. matrix multiplication)

Calculating Eigenvalues and Eigenvectors

Calculation of the eigenvalues and the corresponding eigenvectors is completed using several principles of linear algebra. This can be done by hand, or for more complex situations a multitude of software packages (i.e. Mathematica) can be used. The following discussion will work for any *nxn* matrix; however for the sake of simplicity, smaller and more manageable matrices are used. Note also that throughout this article, **boldface** type is used to distinguish matrices from other variables.

Linear Algebra Review

For those who are unfamiliar with linear algebra, this section is designed to give the necessary knowledge used to compute the eigenvalues and eigenvectors. For a more extensive discussion on linear algebra, please consult the references.

Basic Matrix Operations

An *m x n* **matrix A** is a rectangular array of mn numbers (or elements) arranged in horizontal **rows** (*m*) and vertical **columns** (*n*):

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{1j} & a_{1n} \\ a_{i1} & a_{ij} & a_{in} \\ a_{m1} & a_{mj} & a_{mn} \end{bmatrix}$$

To represent a matrix with the element *aij* in the *i*th row and *j*th column, we use the abbreviation A = [*aij*]. Two *m x n* matrices **A** = [*aij*] and **B** = [*bij*] are said to be equal if corresponding elements are equal.

Addition and subtraction

We can add **A** and **B** by adding corresponding elements:

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}$$

This will give the element in row *i* and column *j* of **C** = **A** + **B** to have

$$c_{ij} = a_{ij} + b_{ij}.$$

More detailed addition and subtraction of matrices can be found in the example below.

$$\begin{bmatrix} 1 & 2 & 6 \\ 4 & 5 & 10 \\ 5 & 3 & 11 \end{bmatrix} + \begin{bmatrix} 8 & 3 & 5 \\ 5 & 4 & 4 \\ 3 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 1+8 & 2+3 & 6+5 \\ 4+5 & 5+4 & 10+4 \\ 5+3 & 3+0 & 11+6 \end{bmatrix} = \begin{bmatrix} 9 & 5 & 11 \\ 9 & 9 & 14 \\ 8 & 3 & 17 \end{bmatrix}$$

Multiplication

Multiplication of matrices are NOT done in the same manner as addition and subtraction. Let's look at the following matrix multiplication:

$$A \times B = C$$

A is an $m \times n$ matrix, B is an $n \times p$ matrix, and C is an $m \times p$ matrix. Therefore the resulting matrix, C, has the same number of rows as the first matrix and the same number of columns as the second matrix. Also the number of columns in the first is the same as the number of rows in the second matrix.

The value of an element in **C** (row i, column j) is determined by the general formula:

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

Thus,

$$\begin{bmatrix} 1 & 2 & 6 \\ 4 & 5 & 10 \\ 5 & 3 & 11 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} 1\times 3+2\times 0+6\times 5 & 1\times 0+2\times 1+6\times 1 \\ 4\times 3+5\times 0+10\times 5 & 4\times 0+5\times 1+10\times 1 \\ 5\times 3+3\times 0+11\times 5 & 5\times 0+3\times 1+11\times 1 \end{bmatrix}$$
$$= \begin{bmatrix} 33 & 8 \\ 62 & 15 \\ 70 & 14 \end{bmatrix}$$

It can also be seen that multiplication of matrices is not commutative (**A B** ≠**B A**). Multiplication of a matrix by a scalar is done by multiplying each element by the scalar.
c**A** = **A**c =[*caij*]

$$2\begin{bmatrix} 1 & 2 & 6 \\ 4 & 5 & 10 \\ 5 & 3 & 11 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 12 \\ 8 & 10 & 20 \\ 10 & 6 & 22 \end{bmatrix}$$

Identity Matrix

The identity matrix is a special matrix whose elements are all zeroes except along the primary diagonal, which are occupied by ones. The identity matrix can be any size as long as the number of rows equals the number of columns.

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Determinant

The determinant is a property of any square matrix that describes the degree of coupling between equations. For a 2x2 matrix the determinant is:

$$\det(\mathbf{A}) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Note that the vertical lines around the matrix elements denotes the determinant. For a 3x3 matrix the determinant is:

$$\det(\mathbf{A}) = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} = a(ei - fh) - b(di - fg) + c(dh - eg)$$

Larger matrices are computed in the same way where the element of the top row is multiplied by the determinant of matrix remaining once that element's row and column are removed. Terms where the top elements in odd columns are added and terms where the top elements in even rows are subtracted (assuming the top element is positive). For matrices larger than 3x3 however; it is probably quickest to use math software to do these calculations since they quickly become more complex with increasing size.

Solving for Eigenvalues and Eigenvectors

The eigenvalues (λ) and eigenvectors (**v**), are related to the **square** matrix **A** by the following equation. (Note: In order for the eigenvalues to be computed, the matrix must have the same number of rows as columns.)

$$(\mathbf{A} - \lambda \mathbf{I}) \cdot \mathbf{v} = 0$$

This equation is just a rearrangement of the Equation 10.3.1. To solve this equation, the eigenvalues are calculated first by setting det(**A**-λ**I**) to zero and then solving for λ. The determinant is set to zero in order to ensure non-trivial solutions for **v**, by a fundamental theorem of linear algebra.

$$A = \begin{bmatrix} 4 & 1 & 4 \\ 1 & 7 & 1 \\ 4 & 1 & 4 \end{bmatrix}$$

$$A - \lambda I = \begin{bmatrix} 4 & 1 & 4 \\ 1 & 7 & 1 \\ 4 & 1 & 4 \end{bmatrix} + \begin{bmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{bmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} 4-\lambda & 1 & 4 \\ 1 & 7-\lambda & 1 \\ 4 & 1 & 4-\lambda \end{vmatrix} = 0$$

$$-54\lambda + 15\lambda^2 - \lambda^3 = 0$$
$$-\lambda(\lambda - 6)(\lambda - 9) = 0$$
$$\lambda = 0, 6, 9$$

For each of these eigenvalues, an eigenvector is calculated which will satisfy the equation (**A**-λ**I**)**v**=**0** for that eigenvalue. To do this, an eigenvalue is substituted into **A**-λ**I**, and then the system of equations is used to calculate the eigenvector. For $\lambda = 6$

$$(\mathbf{A} - 6\mathbf{I})\mathbf{v} = \begin{bmatrix} 4-6 & 1 & 4 \\ 1 & 7-6 & 1 \\ 4 & 1 & 4-6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 & 1 & 4 \\ 1 & 1 & 1 \\ 4 & 1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

Using multiplication we get a system of equations that can be solved.

$$-2x + y + 4z = 0 \longrightarrow \; y = 2x - 4z \qquad (1)$$
$$x + y + z = 0 \longrightarrow \; y = -x - z \qquad (2)$$
$$4x + y - 2z = 0 \longrightarrow \; y = -4x + 2z \qquad (3)$$

Equating (1) and (3) ...

$$2x - 4z = -4x + 2z$$

$$6x = 6z$$
$$x = z$$

Plugging this into (2)...

$$y = -x - x = -2x$$

There is one degree of freedom in the system of equations, so we have to choose a value for one variable. By convention we choose x = 1 then

$$x = 1$$
$$y = -2$$
$$z = 1$$
$$\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

A degree of freedom always occurs because in these systems not all equations turn out to be independent, meaning two different equations can be simplified to the same equation. In this case a small number was chosen (x = 1) to keep the solution simple. However, it is okay to pick any number for x, meaning that each eigenvalue potentially has an infinite number of possible eigenvectors that are scaled based on the initial value of x chosen. Said another way, the eigenvector only points in a direction, but the magnitude of this pointer does not matter. For this example, getting an eigenvector that is $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ is identical to getting an eigenvector that is $\mathbf{v} = \begin{bmatrix} 2 \\ -4 \\ 2 \end{bmatrix}$ or an eigenvector that is scaled by some constant, in this case 2.

Finishing the calcualtions, the same method is repeated for λ = 0 and λ = 9 to get their corresponding eigenvectors.

For λ = 0, $\mathbf{v} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

For λ = 9, $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

In order to check your answers you can plug your eigenvalues and eigenvectors back into the governing equation $\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$. For this example, λ = 6 and $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ was double checked.

$$\begin{bmatrix} 4 & 1 & 4 \\ 1 & 7 & 1 \\ 4 & 1 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = 6 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 4-2+4 \\ 1-14+1 \\ 4-2+4 \end{bmatrix} = \begin{bmatrix} 6 \\ -12 \\ 6 \end{bmatrix}$$

Therefore, λ = 6 and $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ are both an eigenvalue-eigenvector pair for the matrix $\mathbf{A} = \begin{bmatrix} 4 & 1 & 4 \\ 1 & 7 & 1 \\ 4 & 1 & 4 \end{bmatrix}$.

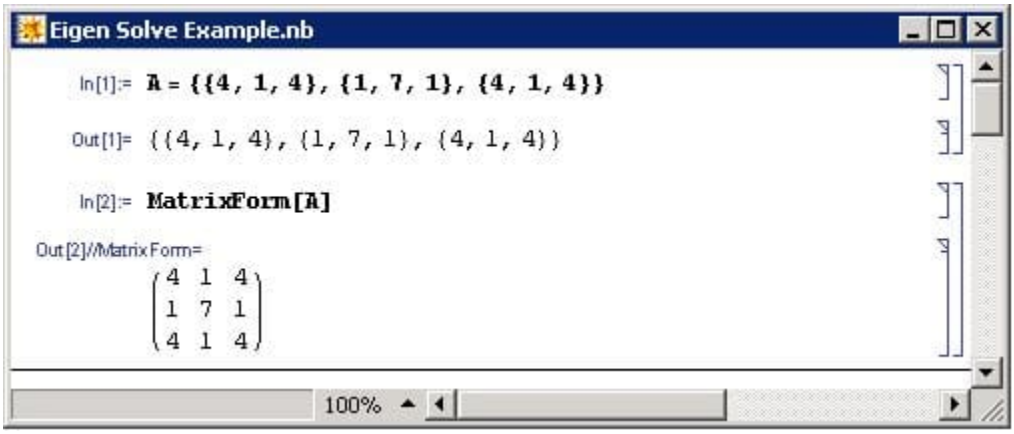
Calculating Eigenvalues and Eigenvectors using Numerical Software

Eigenvalues in Mathematica

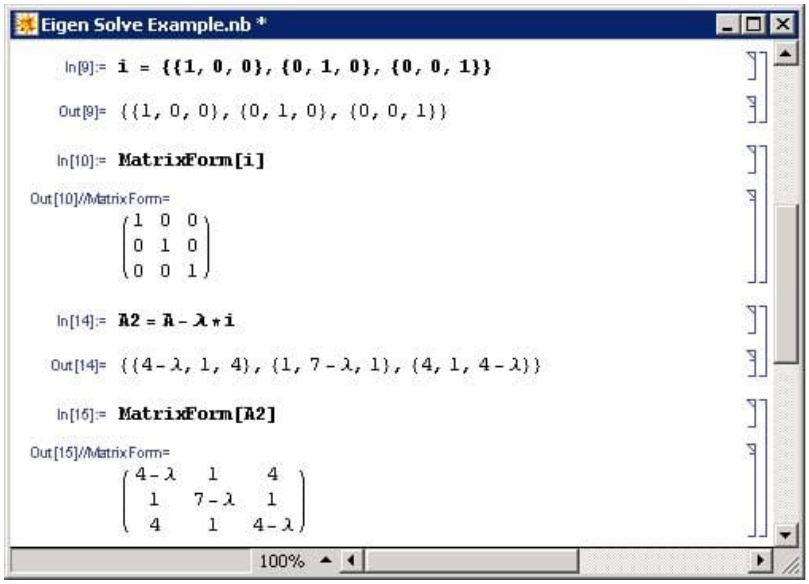
For larger matrices (4x4 and larger), solving for the eigenvalues and eigenvectors becomes very lengthy. Therefore software programs like Mathematica are used. The example from the last section will be used to demonstrate how to use Mathematica. First we can generate the matrix **A**. This is done using the following syntax:

- $A = \{\{4,1,4\},\{1,7,1\},\{4,1,4\}\}$

It can be seen that the matrix is treated as a list of rows. Elements in the same row are contained in a single set of brackets and separated by commas. The set of rows are also contained in a set of brackets and are separated by commas. A screenshot of this is seen below. (Note: The "MatrixForm[]" command is used to display the matrix in its standard form. Also in Mathematica you must hit Shift + Enter to get an output.)



Next we find the determinant of matrix **A**-λ**I**, by first subtracting the matrix λ**I** from A (Note: This new matrix, **A**-λ**I**, has been called A2).

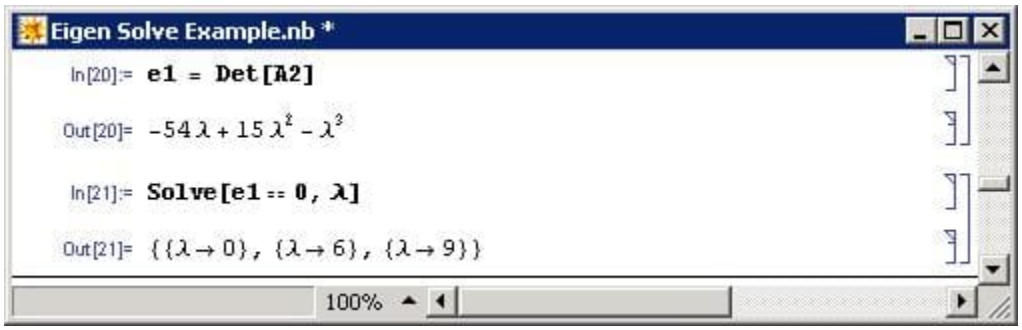


The command to find the determinant of a matrix **A** is:

- Det[A]

For our example the result is seen below. By setting this equation to 0 and solving for λ, the eigenvalues are found. The Solve[] function is used to do this. Notice in the syntax that the use of two equal signs (==) is used to show equivalence whereas a single equal sign is used for defining a variable.

- Solve[{set of equations},{variables being solved}]



Alternatively the eigenvalues of a matrix **A** can be solved with the Mathematica Eigenvalue[] function:

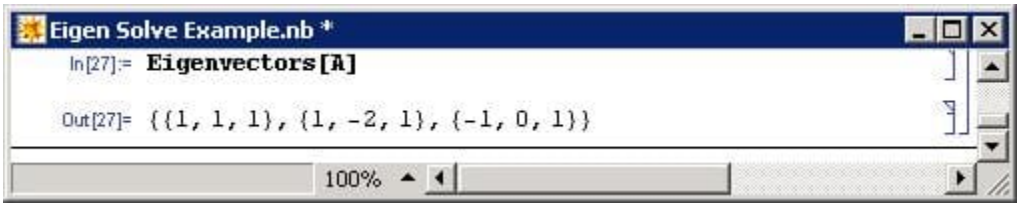
- Eigenvalues[A]

Note that the same results are obtained for both methods.

To find the eigenvectors of a matrix **A**, the Eigenvector[] function can be used with the syntax below.

- Eigenvectors[A]

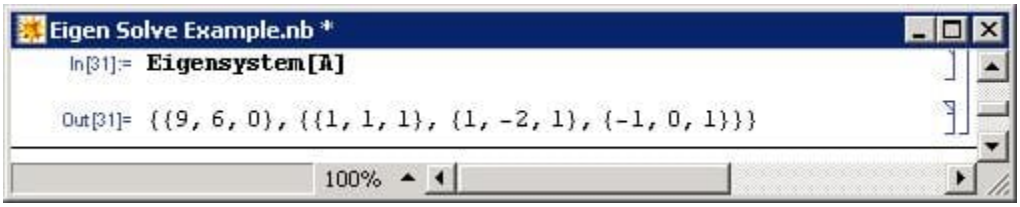
The eigenvectors are given in order of descending eigenvalues.



One more function that is useful for finding eigenvalues and eigenvectors is Eigensystem[]. This function is called with the following syntax.

- Eigensystem[A]

In this function, the first set of numbers are the eigenvalues, followed by the sets of eigenvectors in the same order as their corresponding eigenvalues.

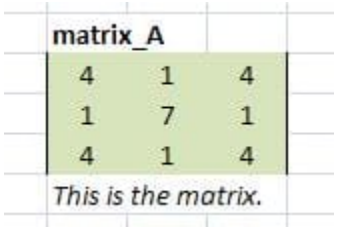


The Mathematica file used to solve the example can be found at this link.Media:Eigen Solve Example.nb

Microsoft Excel

Microsoft Excel is capable of solving for Eigenvalues of symmetric matrices using its Goal Seek function. A symmetric matrix is a square matrix that is equal to its transpose and always has real, not complex, numbers for Eigenvalues. In many cases, complex Eigenvalues cannot be found using Excel. Goal Seek can be used because finding the Eigenvalue of a symmetric matrix is analogous to finding the root of a polynomial equation. The following procedure describes how to calculate the Eigenvalue of a symmetric matrix in the Mathematica tutorial using MS Excel.

(1) Input the values displayed below for matrix A then click menu INSERT-NAME-DEFINE “matrix_A” to name the matrix.



(2) Similarly, define identity matrix I by entering the values displayed below then naming it “matrix_I.”

(3) Enter an initial guess for the Eigenvalue then name it “lambda.”

lambda	6

(4) In an empty cell, type the formula =matrix_A-lambda*matrix_I. Highlight three cells to the right and down, press F2, then press CRTL+SHIFT+ENTER. Name this matrix “matrix_A_lambda_I.”

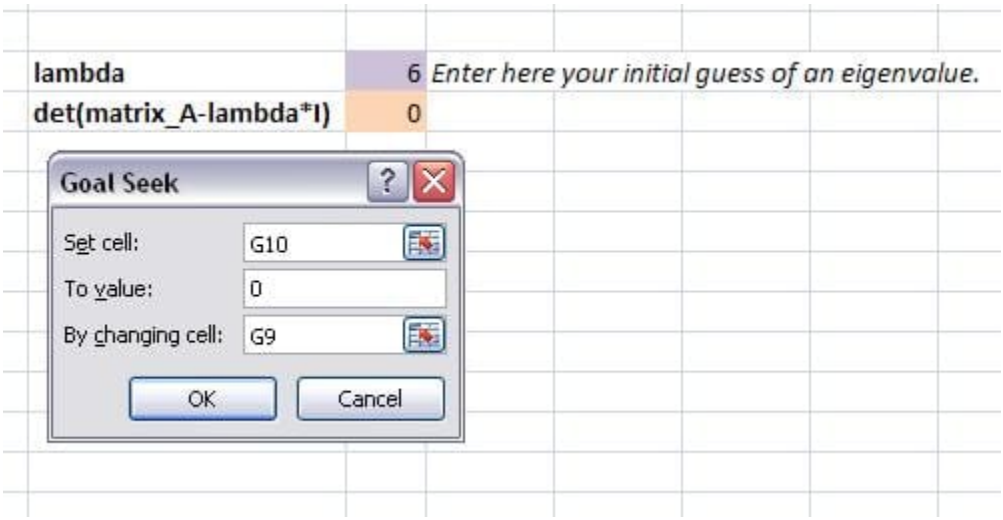
matrix_A_lambda_I		
-2	1	4
1	1	1
4	1	-2

This is matrix_A-lambda*I.

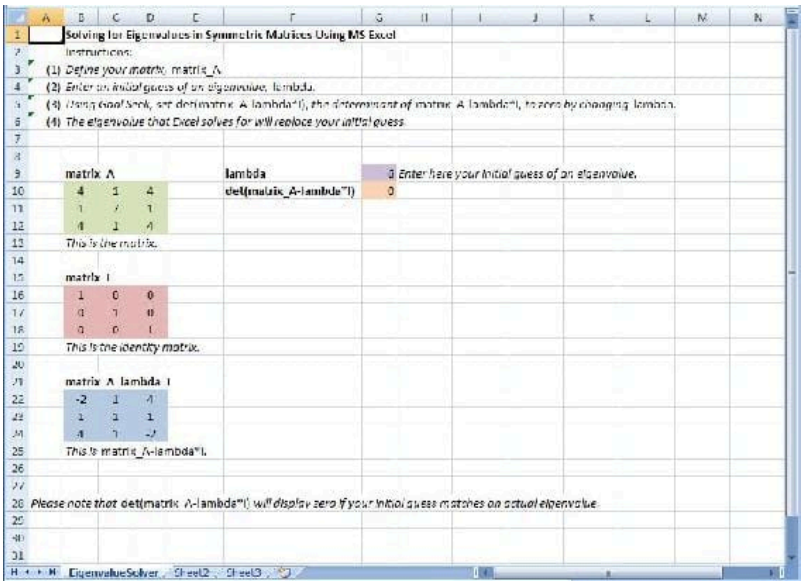
(5) In another cell, enter the formula =MDETERM(matrix_A_lambda_I). This is the determinant formula for matrix_A_lambda_I.

det(matrix_A-lambda*I)	0

(6) Click menu Tools-Goal Seek... and set the cell containing the determinant formula to zero by changing the cell containing lambda.



(7) To obtain all three Eigenvalues for matrix A, re-enter different initial guesses. Excel calculates the Eigenvalue nearest to the value of the initial guess. The Eigenvalues for matrix A were determined to be 0, 6, and 9. For instance, initial guesses of 1, 5, and 13 will lead to Eigenvalues of 0, 6, and 9, respectively.



The MS Excel spreadsheet used to solve this problem, seen above, can be downloaded from this link: Media:ExcelSolveEigenvalue.xls.

Chemical Engineering Applications

The eigenvalue and eigenvector method of mathematical analysis is useful in many fields because it can be used to solve homogeneous linear systems of differential equations with constant coefficients. Furthermore, in chemical engineering many models are formed on the basis of systems of differential equations that are either linear or can be linearized and solved using the eigenvalue eigenvector method. In general, most ODEs can be linearized and therefore solved by this method. Linearizing ODEs For example, a PID control device can be modeled with ODEs that may be linearized where the eigenvalue eigenvector method can then be implemented. If we have a system that can be modeled with linear differential equations involving temperature, pressure, and concentration as they change with time, then the system can be solved using eigenvalues and eigenvectors:

$$\frac{dP}{dt} = 4P - 4T + C$$

$$\frac{dT}{dt} = 4P - T + 3C$$

$$\frac{dC}{dt} = P + 5T - C$$

Note: This is not a real model and simply serves to introduce the eigenvalue and eigenvector method.

A is just the matrix that represents the coefficients in the above linear differential equations. However, when setting up the matrix, **A**, the order of coefficients matters and must remain consistent. Namely, in the following representative matrix, the first column corresponds to the coefficients of **P**, the second column to the coefficients of **T**, and the third column corresponds to the coefficients of **C**. The same goes for the rows. The first row corresponds to $\frac{dP}{dt}$, the second row corresponds to $\frac{dT}{dt}$, and the third row corresponds to $\frac{dC}{dt}$:

$$\mathbf{A} = \begin{bmatrix} 4 & -4 & 1 \\ 4 & -1 & 3 \\ 1 & 5 & -1 \end{bmatrix}$$

It is noteworthy that matrix **A** is only filled with constants for a linear system of differential equations. This turns out to be the

case because each matrix component is the partial differential of a variable (in this case P, T, or C). It is this partial differential that yields a constant for linear systems. Therefore, matrix **A** is really the Jacobian matrix for a linear differential system.

Now, we can rewrite the system of ODE's above in matrix form.

$$\mathbf{x}' = \mathbf{A}\mathbf{x}$$

where

$$\mathbf{x}(t) = \begin{bmatrix} P(t) \\ T(t) \\ C(t) \end{bmatrix}$$

We guess trial solutions of the form

$$\mathbf{x} = \mathbf{v}e^{\lambda t}$$

since when we substitute this solution into the matrix equation, we obtain

$$\lambda \mathbf{v}e^{\lambda t} = \mathbf{A}\mathbf{v}e^{\lambda t}$$

After cancelling the nonzero scalar factor $e^{\lambda t}$, we obtain the desired eigenvalue problem.

$$\mathbf{A}\mathbf{v} = \lambda \mathbf{v}$$

Thus, we have shown that

$$\mathbf{x} = \mathbf{v}e^{\lambda t}$$

will be a nontrivial solution for the matrix equation as long as **v** is a nonzero vector and λ is a constant associated with **v** that satisfies the eigenvalue problem.

In order to solve for the eigenvalues and eigenvectors, we rearrange the Equation 10.3.1 to obtain the following:

$$\left(\mathbf{\Lambda} \quad \lambda \mathbf{I})\mathbf{v} = 0 \right. \qquad \begin{bmatrix} 4-\lambda & -4 & 1 & \\ 4 & 1 & \lambda & 3 \\ 1 & 5 & -1-\lambda & \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

For nontrivial solutions for **v**, the determinant of the eigenvalue matrix must equal zero, $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$. This allows us to solve for the eigenvalues, λ. You should get, after simplification, a third order polynomial, and therefore three eigenvalues. (see section on Solving for Eigenvalues and Eigenvectors for more details) Using the calculated eignvalues, one can determine the stability of the system when disturbed (see following section).

Once you have calculated the three eigenvalues, you are ready to find the corresponding eigenvectors. Plug the eigenvalues back into the equation $(\mathbf{A} - \lambda \mathbf{I})\,\mathbf{v} = 0$ and solve for the corresponding eigenvectors. There should be three eigenvectors, since there were three eigenvalues. (see section on Calculating Eigenvalues and Eigenvectors for more details)

The solution will look like the following:

$$\begin{bmatrix} P(t) \\ T(t) \\ C(t) \end{bmatrix} = c_1 \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} e^{\lambda_1 t} + c_2 \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} e^{\lambda_2 t} + c_3 \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} e^{\lambda_3 t}$$

where

x_1, x_2, x_3, y_1, y_2, y_3, z_1, z_2, z_3 are all constants from the three eigenvectors. The general solution is a linear combination of these three solution vectors because the original system of ODE's is homogeneous and linear. It is homogeneous because the derivative expressions have no cross terms, such as *PC* or *TC*, and no dependence on *t*. It is linear because the derivative operator is linear.

To solve for c_1, c_2, c_3 there must be some given initial conditions (see Worked out Example 1).

This Wiki does not deal with solving ODEs. It only deals with solving for the eigenvalues and eigenvectors. In Mathematica the **Dsolve[]** function can be used to bypass the calculations of eigenvalues and eigenvectors to give the solutions for the differentials directly. See Using eigenvalues and eigenvectors to find stability and solve ODEs for solving ODEs using the eigenvalues and eigenvectors method as well as with Mathematica.

This section was only meant to introduce the topic of eigenvalues and eigenvectors and does not deal with the mathematical details presented later in the article.

Using Eigenvalues to Determine Effects of Disturbing a System

Eigenvalues can help determine trends and solutions with a system of differential equations. Once the eigenvalues for a system are determined, the eigenvalues can be used to describe the system's ability to return to steady-state if disturbed.

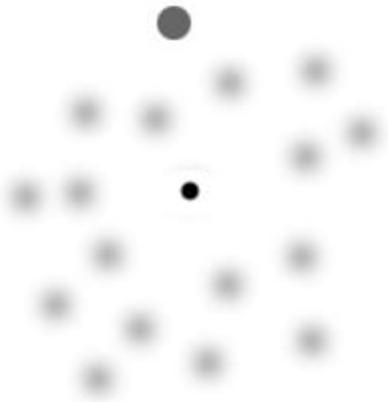
The simplest way to predict the behavior of a system if disturbed is to examine the signs of its eigenvalues. Negative eigenvalues will drive the system back to its steady-state value, while positive eigenvalues will drive it away. What happens if there are two eigenvalues present with opposite signs? How will the system respond to a disturbance in that case? In many situations, there will be one eigenvalue which has a much higher absolute value than the other corresponding eigenvalues for that system of differential equations. This is known as the "dominant eigenvalue", and it will have the greatest effect on the system when it is disturbed. However, in the case that the eigenvalues are equal and opposite sign there is no dominant eigenvalue. In this case the constants from the initial conditions are used to determine the stability.

Another possible case within a system is when the eigenvalue is 0. When this occurs, the system will remain at the position to which it is disturbed, and will not be driven towards or away from its steady-state value. It is also possible for a system to have two identical eigenvalues. In this case the two identical eigenvalues produce only one eigenvector. Because of this, a situation can arise in which the eigenvalues don't give the complete story of the system, and another method must be used to analyze it, such as the Routh Stability Analysis Method.

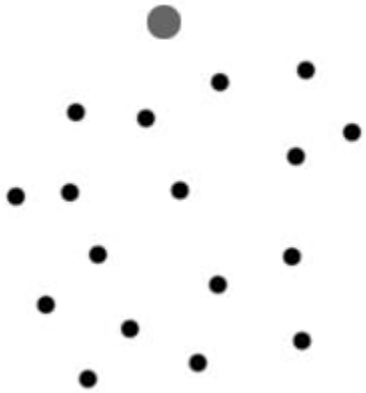
Eigenvalues can also be complex or pure imaginary numbers. If the system is disturbed and the eigenvalues are non-real number, oscillation will occur around the steady state value. If the eigenvalue is imaginary with no real part present, then the system will oscillate with constant amplitude around the steady-state value. If it is complex with a positive real part, then the system will oscillate with increasing amplitude around the function, driving the system further and further away from its steady-state value. Lastly, if the eigenvalue is a complex number with a negative real part, then the system will oscillate with decreasing amplitude until it eventually reaches its steady state value again.

Below is a table of eigenvalues and their effects on a differential system when disturbed. It should be noted that the eigenvalues developed for a system should be reviewed as a system rather than as individual values. That is to say, the effects listed in the table below do not fully represent how the system will respond. If you were to pretend that eigenvalues were nails on a Plinko board, knowing the location and angle of one of those nails would not allow you to predict or know how the Plinko disk would fall down the wall, because you wouldn't know the location or angle of the other nails. If you have information about all of the nails on the Plinko board, you could develop a prediction based on that information. More information on using eigenvalues for stability analysis can be seen here, Using eigenvalues and eigenvectors to find stability and solve ODEs_Wiki.

Eigenvalue	Effect on system when disturbed
Positive real number	Driven away from steady-state value
Negative real number	Driven back to steady-state value
0	Remains at position to which it was disturbed
Identical to another eigenvalue	Effects can not be determined
Complex, positive real number	Oscillates around steady-state value with increasing amplitude
Complex, negative real number	Oscillates around steady-state value with decreasing amplitude
Imaginary	Oscillates around steady-state value with constant amplitude



The above picture is of a plinko board with only one nail position known. Without knowing the position of the other nails, the Plinko disk's fall down the wall is unpredictable.



Knowing the placement of all of the nails on this Plinko board allows the player to know general patterns the disk might follow.

Repeated Eigenvalues

A final case of interest is repeated eigenvalues. While a system of N differential equations must also have N eigenvalues, these values may not always be distinct. For example, the system of equations:

$$\frac{dC_A}{dt} = f_{A,in}\rho C_A = f_{out}\rho C_A\sqrt{V_1} - V_1 k_1 C_A C_B$$
$$\frac{dC_B}{dt} = f_{B,in}\rho C_{Bin} - f_{out}\rho C_B\sqrt{V_1} - V_1 k_1 C_A C_B$$
$$\frac{dC_C}{dt} = -f_{out}\rho C_c\sqrt{V_1} + V_1 k_1 C_A C_B$$
$$\frac{dV_1}{dt} = f_{A,in} + f_{B,in} - f_{out}\sqrt{V_1}$$
$$\frac{dV_2}{dt} = f_{out}\sqrt{V_1} - f_{customer}\sqrt{V_2}$$
$$\frac{dC_{C2}}{dt} = f_{\text{out}}, \rho C_C\sqrt{V_1} - f_{\text{customer}}, \rho C_{\text{C2}}\sqrt{V_2}$$

May yield the eigenvalues: {-82, -75, -75, -75, -0.66, -0.66}, in which the roots '-75' and '-0.66' appear multiple times. Repeat eigenvalues bear further scrutiny in any analysis because they might represent an edge case, where the system is operating at some extreme. In mathematical terms, this means that linearly independent eigenvectors cannot be generated to complete the matrix basis without further analysis. In "real-world" engineering terms, this means that a system at an edge case could distort

or fail unexpectedly.

However, for the general solution:

$$Y(t) = k_1 \exp(\lambda t) V_1 + k_2 \exp(\lambda t)\,(tV_1 + V_2)$$

If $\lambda < 0$, as t approaches infinity, the solution approaches 0, indicating a stable sink, whereas if λ > 0, the solution approaches infinity in the limit, indicating an unstable source. Thus the rules above can be roughly applied to repeat eigenvalues, that the system is still likely stable if they are real and less than zero and likely unstable if they are real and positive. Nonetheless, one should be aware that unusual behavior is possible. This course will not concern itself with resultant behavior of repeat eigenvalues, but for further information, see:

- http://math.rwinters.com/S21b/supplements/newbasis.pdf
- http://www.sosmath.com/diffeq/system/linear/eigenvalue/repeated/repeated.html

Example 10.3.1

Your immediate supervisor, senior engineer Captain Johnny Goonewadd, has brought you in on a project dealing with a new silcone-based sealant that is on the ground level of research. Your job is to characterize the thermal expansion of the sealant with time given a constant power supply. Luckily, you were given a series of differential equations that relate temperature and volume in terms of one another with respect to time (Note: T and V are both dimensionless numbers with respect to their corresponding values at t=0). Solve the system of differentials and determine the equations for both Temperature and Volume in terms of time.

Solution

You are given the initial condition at time t=0, T=1 and V=1

$$\frac{dT}{dt} = 4T - 3V$$

$$\frac{dV}{dt} = 3T + 4V$$

By defining a matrix for both the coefficients and dependant variables we are able to rewrite the above series of differentials in matrix form

$$A = \begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix}$$

$$X = \begin{bmatrix} T \\ V \end{bmatrix}$$

$$A * X = \begin{bmatrix} \frac{dT}{dV} \\ \frac{dV}{dt} \end{bmatrix}$$

Lambda is inserted into the A matrix to determine the Eigenvalues

$$det(A - \lambda I) = 0$$

$$det \begin{bmatrix} 4-\lambda & -3 \\ 3 & 4-\lambda \end{bmatrix} = 0$$

$$(4-\lambda)^2 + 9 = 0$$

$$\lambda_1 = 4 - 3i$$

$$\lambda_2 = 4 + 3i$$

For each eigenvalue, we must find the eigenvector. Let us start with $\lambda_1 = 4 - 3i$

$$(A - \lambda I)v = 0$$

$$\begin{bmatrix} 4-(4-3i) & -3 \\ 3 & 4-(4-3i) \end{bmatrix}$$

$$3iT - 3V = 0$$

$$3T + 3iV = 0$$

$$v_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

Now we find the eigenvector for the eigenvalue $\lambda_2 = 4 + 3i$

$$(A - \lambda I)v = 0$$

$$\begin{bmatrix} 4-(4+3i) & -3 \\ 3 & 4-(4+3i) \end{bmatrix}$$

$$-3iT - 3V = 0$$

$$3T - 3iV = 0$$

$$v_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The general solution is in the form

$$x(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2$$

$$\begin{bmatrix} T(t) \\ V(t) \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ i \end{bmatrix} e^{(4-3i)t} + c_2 \begin{bmatrix} 0 \\ 0 \end{bmatrix} e^{(4+3i)t}$$

$$\begin{bmatrix} T(t) \\ V(t) \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ i \end{bmatrix} e^{(4-3i)t}$$

A mathematical proof, Euler's formula, exists for transforming complex exponentials into functions of sin(t) and cos(t)

$$e^{(a+bi)t} = e^{at}(cos(bt) - isin(bt))$$

Thus

$$\begin{bmatrix} T(t) \\ V(t) \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ i \end{bmatrix} e^{4t}(cos(3t) - isin(3t))$$

Simplifying

$$\begin{bmatrix} T(t) \\ V(t) \end{bmatrix} = e^{4t} \begin{bmatrix} c_1 cos(3t) - c_1 isin(3t) \\ c_1 icos(3t) + c_1 sin(3t) \end{bmatrix}$$

Since we already don't know the value of c_1, let us make this equation simpler by making the following substitution

$$c_3 = c_1$$

$$c_4 = c_1 i$$

Thus, we get have our solution in terms of real numbers

$$\begin{bmatrix} T(t) \\ V(t) \end{bmatrix} = e^{4t} \begin{bmatrix} c_3 cos(3t) - c_4 sin(3t) \\ c_4 cos(3t) + c_3 sin(3t) \end{bmatrix}$$

Or, rewriting the solution in scalar form

$$T(t) = e^{4t}(c_3 cos(3t) - c_4 sin(3t))$$

$$V(t) = e^{4t}(c_3 sin(3t) + c_4 cos(3t))$$

Now that we have our solutions, we can use our initial conditions to find the constants c_3 and c_4

First initial condition: t=0, T=1

$$T(t) = e^{4t}(c_3 cos(3t) - c_4 sin(3t))$$

$$1 = e^{4*0}(c_3 cos(3*0) - c_4 sin(3*0))$$

$$1 = 1*(c_3*1 - c_4*0)$$

$$c_3 = 1$$

First initial condition: t=0, V=1

$$V(t) = e^{4t}(c_3 sin(3t) + c_4 cos(3t))$$

$$1 = e^{4*0}(c_3 sin(3*0) + c_4 cos(3*0))$$

$$1 = 1 * (c_3 * 0 + c_4 * 1)$$

$$c_4 = 1$$

We have now arrived at our solution

$$T(t) = e^{4t}(cos(3t) - sin(3t))$$

$$V(t) = e^{4t}(sin(3t) + cos(3t))$$

See Using eigenvalues and eigenvectors to find stability and solve ODEs_Wiki for solving ODEs using the eigenvalues and eigenvectors.

Example 10.3.2

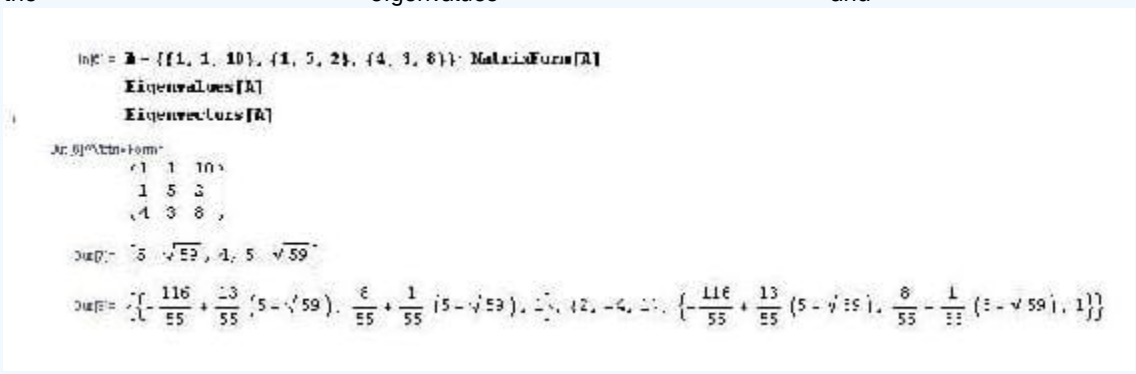
Process Engineer, Dilbert Pickel, has started his first day for the Helman's Pickel Brine Factory. His first assignment is with a pre-startup team formulated to start up a new plant designed to make grousley sour pickle brine. Financial constraints have demanded that the process begin to produce good product as soon as possible. However, you are forced to reflux the process until you reach the set level of sourness. You have equations that relate all of the process variable in terms of one another with respect to time. Therefore, it is Dill Pickles job to characterize all of the process variables in terms of time (dimensionless Sourness, Acidity, and Water content; S, A, & W respectively). Below is the set of differentials that will be used to solve the equation.

\[\begin{array}{l}
\frac{d S}{d t}=S+A+10 W \\
\frac{d A}{d t}=S+5 A+2 W \\
\frac{d W}{d t}=4 S+3 A+8 W
\end{array}]

Thus the coefficient matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 10 \\ 1 & 5 & 2 \\ 4 & 3 & 8 \end{bmatrix}$$

Using mathematica it is easy to input the coefficients of the system of equations into a matrix and determine both the eigenvalues and eigenvectors.



The eigenvectors can then be used to determine the final solution to the system of differentials. Some data points will be necessary in order to determine the constants.

$$\begin{bmatrix} S \\ A \\ W \end{bmatrix} = C_1 \begin{bmatrix} 0.88 \\ 0.38 \\ 1 \end{bmatrix} e^{(5+\sqrt{89}k} + C_2 \begin{bmatrix} 2 \\ -4 \\ 1 \end{bmatrix} e^{4t} + C_3 \begin{bmatrix} -2.74 \\ 0.10 \\ 1 \end{bmatrix}$$

See Using eigenvalues and eigenvectors to find stability and solve ODEs_Wiki for solving ODEs using the eigenvalues and eigenvectors.

Example 10.3.3

It is possible to find the Eigenvalues of more complex systems than the ones shown above. Doing so, however, requires the use of advanced math manipulation software tools such as Mathematica. Using Mathematica, it is possible to solve the system of ODEs shown below.

$$\begin{aligned} \frac{dX}{dt} &= 8X + \frac{10XYF}{X+Z} \\ \frac{dY}{dt} &= 4F - Y - Z - \frac{3XY}{X+Y} \\ \frac{dZ}{dt} &= 9X - 2Z + F \end{aligned}$$

Obviously, this system of ODEs has 4 variables and only 3 equations.

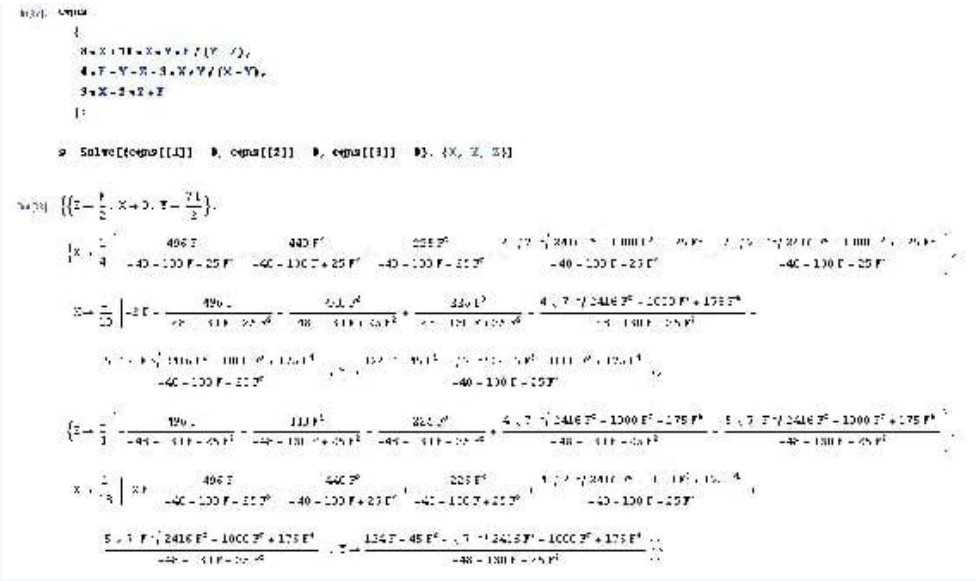
Obviously, this is a more complex set of ODEs than the ones shown above. And even though they will create a more complex set of Eigenvalues, they are solved for in the same way when using Mathematica.

Using the code shown below:

```
In[39]:= eqns =
          {
           8 * X + 10 * X * Y * F / (Y + Z),
           4 * F - Y - Z - 3 * X * Y / (X + Y),
           9 * X - 2 * Z + F
          }

Out[39]= {8 X + 10 F X Y / (Y + Z), 4 F - Y - 3 X Y / (X + Y) - Z, F + 9 X - 2 Z}
```

The equations can be entered into Mathematica. The equations are shown again in the output

Then, using the next bit of code:



The it is possible to find where the equations are equal to 0 (i.e. the fixed points). The results of this is also shown in the image above. It's notable that 3 solutions are found. This makes sense as the system is 3 ODEs.

The Jacobian can then be found by simply using the code shown below.

```
eqns =
 {
  10*X + 10*X*Y*F/(Y + Z),
  4*F - Y - Z - 3*X*Y/(X + Y),
  9*X - 2*Z*F
 };

s = Solve[{eqns[[1]] == 0, eqns[[2]] == 0, eqns[[3]] == 0}, {X, Y, Z}];

Jac = {
  {D[eqns[[1]], X], D[eqns[[1]], Y], D[eqns[[1]], Z]},
  {D[eqns[[2]], X], D[eqns[[2]], Y], D[eqns[[2]], Z]},
  {D[eqns[[3]], X], D[eqns[[3]], Y], D[eqns[[3]], Z]}
 }
```

$$\left\{\left\{0+\frac{10FY}{Y+Z},-\frac{10FXY}{(Y+Z)^2}-\frac{10FX}{Y+Z},-\frac{10FXY}{(Y+Z)^2}\right\},\left\{\frac{3XY}{(X+Y)^2}-\frac{3Y}{X+Y},-1+\frac{3XY}{(X+Y)^2}-\frac{3X}{X+Y},-1\right\},\{9,0,-2\}\right\}$$

The ersults of finding the Jacobian are shown in the equation above.

Finally, to find one of the Eigenvalues, one can simply use the code shown below.

```
In[57]:= eqns =
          {
           8 * X + 10 * X * Y * F / (Y + Z),
           4 * F - Y - Z - 3 * X * Y / (X + Y),
           9 * X - 2 * Z + F
          };

         s = Solve[{eqns[[1]] == 0, eqns[[2]] == 0, eqns[[3]] == 0}, {X, Y, Z}];

         Jac = {
            {D[eqns[[1]], X], D[eqns[[1]], Y], D[eqns[[1]], Z]},
            {D[eqns[[2]], X], D[eqns[[2]], Y], D[eqns[[2]], Z]},
            {D[eqns[[3]], X], D[eqns[[3]], Y], D[eqns[[3]], Z]}
           };

         t1 = Simplify[Eigenvalues[Jac /. s[[1]]]]

Out[60]= {-2, -1, 8 + 35 F/4}
```

This gives the Eigenvalue when the first fixed point (the first solution found for "s") is applied. The other two solutions could be found by simply changing the fixed blade that is referred to when finding t1. The other Eigenvalues are not shown because of their large size.

Exercise 10.3.1: Calculating Eigenvalues and Eigenvectors using Numerical Software

What are the eigenvalues for the matrix **A**?

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 3 & -1 \end{bmatrix}$$

a. $\lambda_1 = -2$ and $\lambda_1 = 5$
b. $\lambda_1 = 2$ and $\lambda_1 = -5$
c. $\lambda_1 = 2$ and $\lambda_1 = 5$
d. $\lambda_1 = -2$ and $\lambda_1 = -5$

Answer available at the end of this section.

Exercise 10.3.2: Using Eigenvalues to Determine Effects of Disturbing a System

When a differential system with a real negative eigenvalue is disturbed, the system is...

a. Driven away from the steady state value
b. Unchanged and remains at the disturbed value
c. Driven back to the steady state value
d. Unpredictable and the effects can not be determined

Answer available at the end of this section.

References

- Kravaris, Costas: Chemical Process Control: A Time-Domain Approach. Ann Arbor: The University of Michigan, pp 1-23, A.1-A.7.
- Bhatti, M. Asghar: Practical Optimization Methods with Mathematica Applications. Springer, pp 75-85, 677-691.

- Strang, Prof. Gilbert: “Eigenvalues and Eigenvectors.” Math 18.06. Lord Foundation of Massachusetts. Fall 1999.
- Edwards, C. Henry and David E. Penney: Differential Equations: Computing and Modeling. Upper Saddle River: Pearson Education, Inc, pp 299-365.
- Teknomo, Kardi. Finding Eigen Value of Symmetric matrix Using Microsoft Excel. http:\people.revoledu.com\kardi\ tutorial\ Excel\EigenValue.html

Contributors and Attributions

- Authors: (October 19, 2006) Tommy DiRaimondo, Rob Carr, Marc Palmer, Matt Pickvet
- Stewards: (October 22, 2007) Shoko Asei, Brian Byers, Alexander Eng, Nicholas James, Jeffrey Leto

Answers

Exercise 10.3.1**: Calculating Eigenvalues and Eigenvectors using Numerical Software**

```
In[1]:= A = {{4, 2}, {3, -1}};
        MatrixForm[A]

Out[2]//MatrixForm=
        ( 4  2 )
        ( 3 -1 )

In[3]:= Eigenvalues[A]

Out[3]= {5, -2}
```

a. $\lambda_1 = -2 \qquad \lambda_2 = 5$

Exercise 10.3.2**: Using Eigenvalues to Determine Effects of Disturbing a System**

c. A real negative eigenvalue is indicative of a stable system that will return to the steady state value after it is disturbed.

This page titled 10.3: Eigenvalues and Eigenvectors is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

10.4: Using Eigenvalues and Eigenvectors to Find Stability and Solve ODEs

Introduction

Eigenvalues and eigenvectors are very useful in the modeling of chemical processes. When designing the controls for a process it is necessary to create a program to operate these controls. Differential equations are used in these programs to operate the controls based on variables in the system. These equations can either be solved by hand or by using a computer program. The solutions for these differential equations will determine the stability of the system. After finding this stability, you can show whether the system will be stable and damped, unstable and undamped (so that there is constant fluctuation in the system), or as an unstable system in which the amplitude of the fluctuation is always increasing. For the first case, a stable and damped system, if there is a change, the system will adjust itself properly to return to steady state. For the other two cases, the system will not be able to return to steady state. For the undamped situation, the constant fluctuation will be hard on the system and can lead to equipment failure. The final situation, with the ever increasing amplitude of the fluctuations will lead to a catastrophic failure.

There are a couple ways to develop the differential equation used to determine stability. First, you can create a differential equation to guide the system where the variables are the readings from the sensors in the system. A second method would be using actual data found from running the system. You could fit a differential equation to this data and use that equation for stability determination.

In this section on Eigenvalue Stability, we will first show how to use eigenvalues to solve a system of linear ODEs. Next, we will use the eigenvalues to show us the stability of the system. After that, another method of determining stability, the Routh stability test, will be introduced. For the Routh stability test, calculating the eigenvalues is unnecessary which is a benefit since sometimes that is difficult. Finally, the advantages and disadvantages of using eigenvalues to evaluate a system's stability will be discussed.

Solving ODEs

Eigenvalues and eigenvectors can be used as a method for solving linear systems of ordinary differential equations (ODEs). The method is rather straight-forward and not too tedious for smaller systems. See The Eigenvector Eigenvalue Method for solving systems by hand and Linearizing ODEs for a linear algebra/Jacobian matrix review. When trying to solve large systems of ODEs however, it is usually best to use some sort of mathematical computer program. Mathematica is a program that can be used to solve systems of ordinary differential equations when doing them by hand is simply too tedious. Once one overcomes the syntax of Mathematica, solving enormous systems of ordinary linear differential equations becomes a piece of cake!

Using Eigenvalues to Solve a System

A linear system will be solve by hand and using Eigenvalues[ ] expression in Mathematica simultaneously. Note that, in the Mathematica inputs below, "In[]:=" is not literally typed into the program, only what is after it. The syntax needed to be typed is the line following "In[]=" . The term is used here to more accurately demonstrate coding in Mathematica. To find a general solution of the linear system of ordinary differential equation:

$$\frac{dx}{dt} = 4x + 8y$$

$$\frac{dy}{dt} = 10x + 2y$$

We first put the system in matrix form:

$$A = \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} 4 & 8 \\ 10 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Where we can see that

$$A = \begin{bmatrix} 4 & 8 \\ 10 & 2 \end{bmatrix}$$

In mathematica, we can use the following code to represent A:

```
In[1]:= MatrixForm [{{4,8},{10,2}}]
Out[1]:= \(\begin{bmatrix}4 & 8\\ 10 & 2\end{bmatrix}\)
```

The eigenvalues λ_1 and λ_2, are found using the characteristic equation of the matrix A, det(A- λI)=0.

$$det(A - \lambda I) = 0$$

$$det(\begin{bmatrix} 4 & 8 \\ 10 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}) = 0$$

$$det \begin{bmatrix} 4-\lambda & 8 \\ 10 & 2-\lambda \end{bmatrix} = 0$$

$$(4-\lambda)(2-\lambda) - 80 = 0$$

$$(\lambda - 12)(\lambda + 6) = 0$$

Therefore, $\lambda_1 = 12$ and $\lambda_2 = -6$

We can use Mathematica to find the eigenvalues using the following code:

```
In[2]:= Eigenvalues[{{4,8},{10,2}}]

Out[2]:={12,-6}
```

Now, for each eigenvalue (λ_1=12 and λ_2=-6), an eigenvector associated with it can be found using $A - \lambda I)\vec{v} = 0$, where $\vec{v}$is an eigenvector such that $vecv = \lambda\vec{v}$

i) For λ_1=12

$$A - \lambda_1 I)\vec{v} = 0$$

$$\begin{bmatrix} 4-12 & 8-0 \\ 10-0 & 2-12 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$$\begin{bmatrix} -8 & 8 \\ 10 & -10 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

This will lead to the Equations 10.4.1 & 10.4.2:

$$-8x + 8y = 0 \tag{10.4.1}$$

$$+10x - 10y = 0 \tag{10.4.2}$$

The Mathematica input is:

```
In[3]:= eqn1= -8x+8y==0
In[4]:= eqn2= 10x-10y==0

In[5]:= Solve[{eqn1,eqn2},{x,y}]

Out[5]:= { { y → x } }
```

Equations 10.4.1 & 10.4.2 lead to the solution $y = x$

ii) For λ_2=-6

$$A - \lambda_2 I)\vec{v} = 0$$

$$\begin{bmatrix} 4+6 & 8-0 \\ 10-0 & 2+6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$$\begin{bmatrix} 10 & 8 \\ 10 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

This will lead to the Equations 10.4.3 & 10.4.4:

$$10x + 8y = 0 \tag{10.4.3}$$

$$10x + 8y = 0 \tag{10.4.4}$$

The Mathematica input is:

In[6]:= **eqn3= 10x+8y==0**

In[7]:= **eqn4= 10x+8y==0**

In[8]:= **Solve[{eqn3,eqn4},{x,y}]**

Out[8]:= { { $y \to -\frac{5}{4}x$ } }

Equations 10.4.3 & 10.4.4 lead to the solution

$$y = -\frac{5}{4}x$$

Recall that the direction of a vector such as $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is the same as the vector $\begin{bmatrix} 4 \\ 8 \end{bmatrix}$ or any other scalar multiple. Therefore, to get the eigenvector, we are free to choose for either the value x or y.

i) For λ_1 = 12

We have arrived at *y* = *x*. As mentioned earlier, we have a degree of freedom to choose for either x or y. Let's assume that x=1. Then, y=1 and the eigenvector $\vec{v_1}$ associated with the eigenvalue λ_1 is $\vec{v_1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

ii) For λ_2 = – 6

We have arrived at $y = -\frac{5}{4}x$. Let's assume that x = 4. Then, y = -5 and the eigenvector associated with the eigenvalue λ_2 is $\vec{v_2} = \begin{bmatrix} 4 \\ -5 \end{bmatrix}$.

These two eigenvalues and associated eigenvectors yield the solution:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{12t} + c_2 \begin{bmatrix} 4 \\ -5 \end{bmatrix} e^{-6t}$$

Hence a general solution of the linear system in scalar form is:

$$lx(t) = c_1 e^{12t} + c_2 4e^{-6t}$$
$$y(t) = c_1 e^{12t} - c_2 5e^{-6t}$$

Solving a System Using DSolve

Using the same linear system of ordinary differential equations:

$$\frac{dx}{dt} = 4x + 8y$$

$$\frac{dy}{dt} = 10x + 2y$$

We input the differential equations to Mathematica with the following command:

In:= **ODEs={x'[t]==4x[t]+8y[t],y'[t]==10x[t]+2y[t]}**

where x'[t] = $\frac{dx}{dt}$ and y'[t] = $\frac{dy}{dt}$

After entering the equations, we use the DSolve function:

In:= **DSolve[ODEs, {x[t], y[t]}, t]**

Mathematica will output the solution:

Out:= { { $x[t] \to \frac{1}{9}e^{-6t}\left(4+5e^{18t}\right)C[1] + \frac{4}{9}e^{-6t}\left(-1+e^{18t}\right)C[2]$, $y[t] \to \frac{5}{9}e^{-6t}\left(-1+e^{18t}\right)C[1] + \frac{1}{9}e^{-6t}\left(5+4e^{18t}\right)C[2]$ } }

This set of equations, although looks more complicated than the first one, is actually the same.

Stability

Eigenvalues can be used to determine whether a fixed point (also known as an equilibrium point) is stable or unstable. A stable fixed point is such that a system can be initially disturbed around its fixed point yet eventually return to its original location and remain there. A fixed point is unstable if it is not stable. To illustrate this concept, imagine a round ball in between two hills. If left alone, the ball will not move, and thus its position is considered a fixed point. If we were to disturb the ball by pushing it a little bit up the hill, the ball will roll back to its original position in between the two hills. This is a stable fixed point. Now image that the ball is at the peak of one of the hills. If left undisturbed, the ball will still remain at the peak, so this is also considered a fixed point. However, a disturbance in any direction will cause the ball to roll away from the top of the hill. The top of the hill is considered an unstable fixed point.

The eigenvalues of a system linearized around a fixed point can determine the stability behavior of a system around the fixed point. The particular stability behavior depends upon the existence of real and imaginary components of the eigenvalues, along with the signs of the real components and the distinctness of their values. We will examine each of the possible cases below.

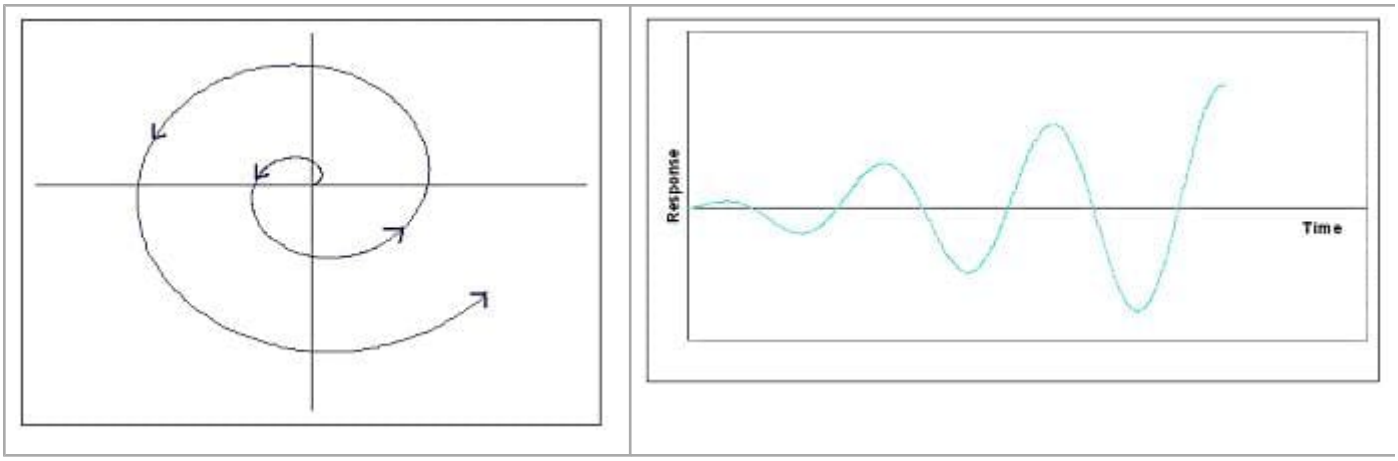
Imaginary (or Complex) Eigenvalues

When eigenvalues are of the form $a + bi$, where $a$and b are real scalars and i is the imaginary number $\sqrt{-1}$, there are three important cases. These three cases are when the real part a is positive, negative, and zero. In all cases, when the complex part of an eigenvalue is non-zero, the system will be oscillatory.

Positive Real Part

When the real part is positive, the system is unstable and behaves as an unstable oscillator. This can be visualized as a vector tracing a spiral away from the fixed point. The plot of response with time of this situation would look sinusoidal with ever-increasing amplitude, as shown below.

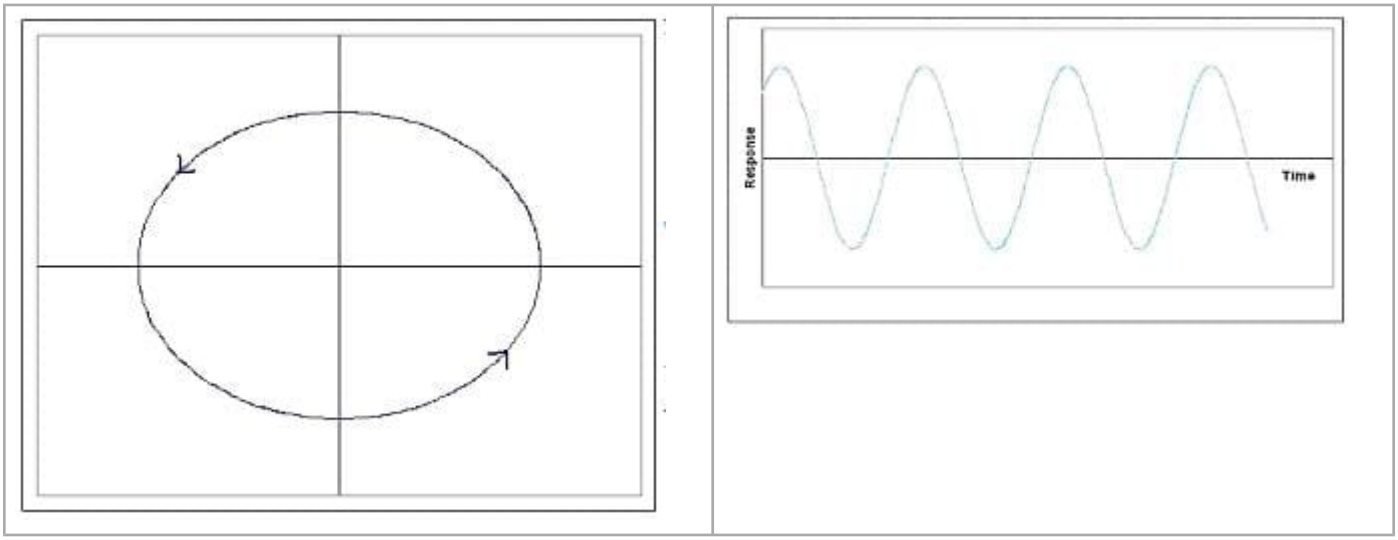
This situation is usually undesirable when attempting to control a process or unit. If there is a change in the process, arising from the process itself or from an external disturbance, the system itself will not go back to steady state.



Zero Real Part

When the real part is zero, the system behaves as an undamped oscillator. This can be visualized in two dimensions as a vector tracing a circle around a point. The plot of response with time would look sinusoidal. The figures below should help in understanding.

Undamped oscillation is common in many control schemes arising out of competing controllers and other factors. Even so, this is usually undesirable and is considered an unstable process since the system will not go back to steady state following a disturbance.

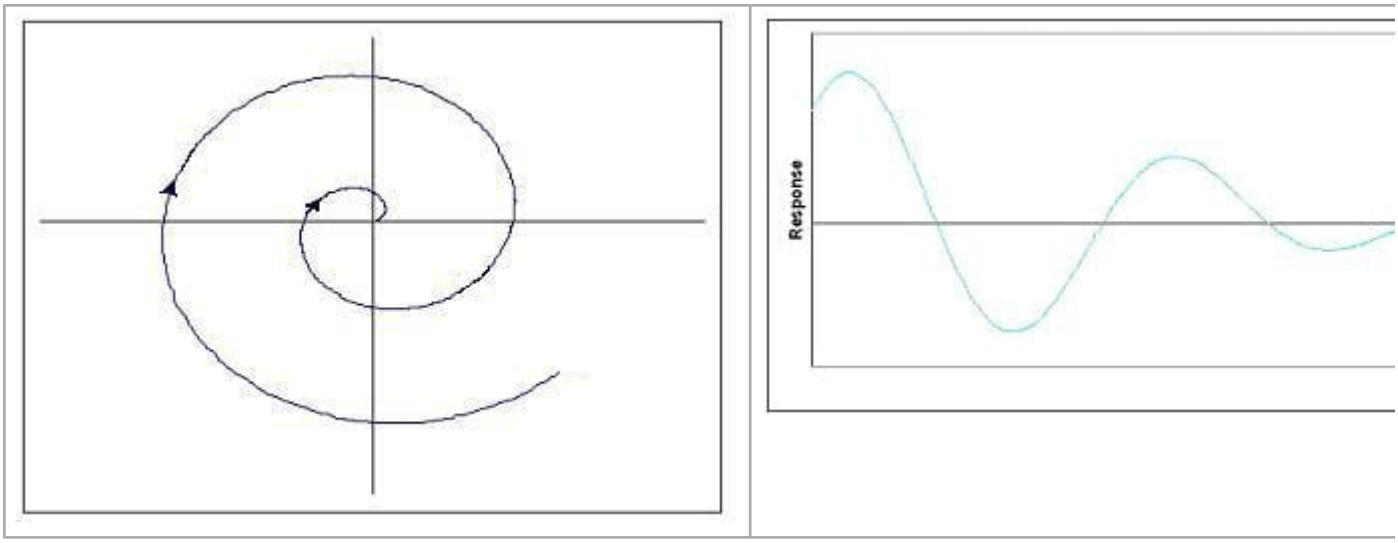


Negative Real Part

When the real part is negative, then the system is stable and behaves as a damped oscillator. This can be visualized as a vector tracing a spiral toward the fixed point. The plot of response with time of this situation would look sinusoidal with ever-decreasing amplitude, as shown below.

This situation is what is generally desired when attempting to control a process or unit. This system is stable since steady state will be reached even after a disturbance to the system. The oscillation will quickly bring the system back to the setpoint, but will over shoot, so if overshooting is a large concern, increased damping would be needed.

While discussing complex eigenvalues with negative real parts, it is important to point out that having all negative real parts of eigenvalues is a necessary and sufficient condition of a stable system.



Complex Part of Eigenvalues

As previously noted, the stability of oscillating systems (i.e. systems with complex eigenvalues) can be determined entirely by examination of the real part. Although the sign of the complex part of the eigenvalue may cause a phase shift of the oscillation, the stability is unaffected.

Real Eigenvalues

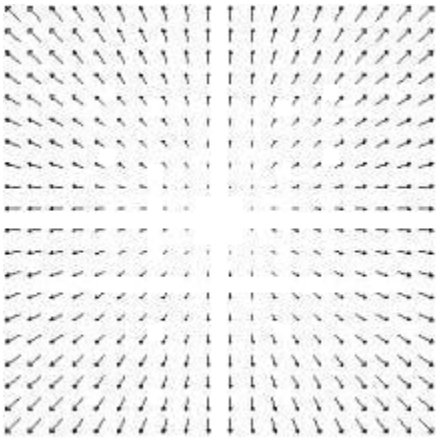
We've seen how to analyze eigenvalues that are complex in form, now we will look at eigenvalues with only real parts.

Zero Eigenvalues

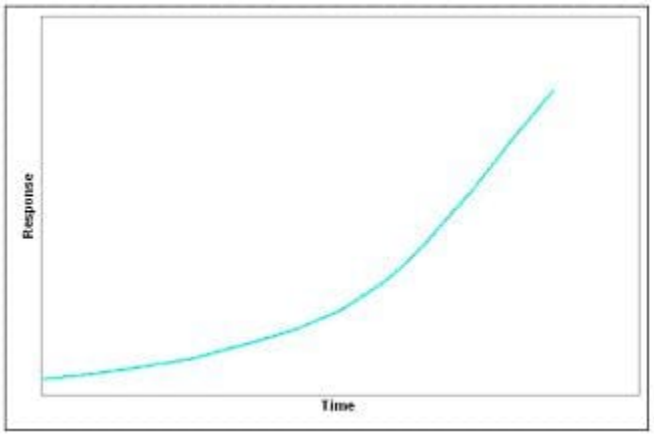
If an eigenvalue has no imaginary part and is equal to zero, the system will be unstable, since, as mentioned earlier, a system will not be stable if its eigenvalues have any non-negative real parts. This is just a trivial case of the complex eigenvalue that has a zero part.

Positive Eigenvalues

When all eigenvalues are real, positive, and distinct, the system is unstable. On a gradient field, a spot on the field with multiple vectors circularly surrounding and pointing out of the same spot (a node) signifies all positive eigenvalues. This is called a source node.

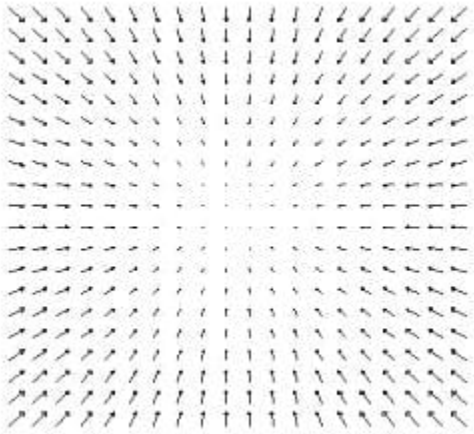


Graphically, real and positive eigenvalues will show a typical exponential plot when graphed against time.

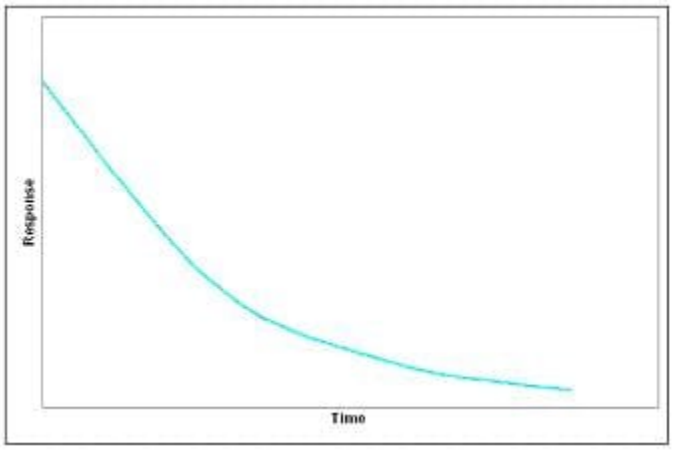


Negative Eigenvalues

When all eigenvalues are real, negative, and distinct, the system is stable. Graphically on a gradient field, there will be a node with vectors pointing toward the fixed point. This is called a sink node.

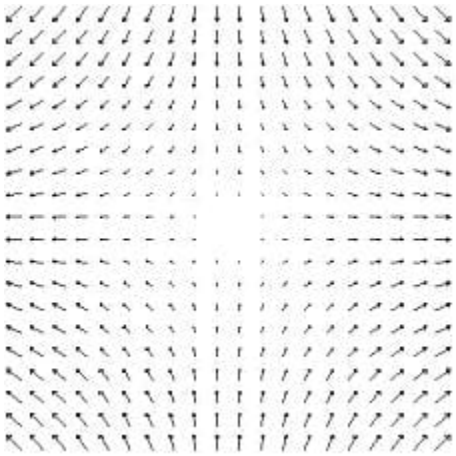


Graphically, real and negative eigenvalues will output an inverse exponential plot.



Positive and Negative Eigenvalues

If the set of eigenvalues for the system has both positive and negative eigenvalues, the fixed point is an unstable saddle point. A saddle point is a point where a series of minimum and maximum points converge at one area in a gradient field, without hitting the point. It is called a saddle point because in 3 dimensional surface plot the function looks like a saddle.

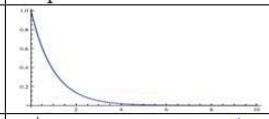
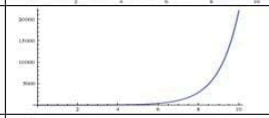
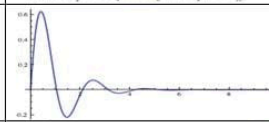
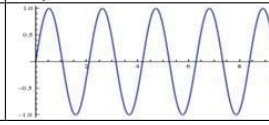


Repeated Eigenvalues

If the set of eigenvalues for the system has repeated real eigenvalues, then the stability of the critical point depends on whether the eigenvectors associated with the eigenvalues are linearly independent, or orthogonal. This is the case of degeneracy, where more than one eigenvector is associated with an eigenvalue. In general, the determination of the system's behavior requires further analysis. For the case of a fixed point having only *two* eigenvalues, however, we can provide the following two possible cases. If the two repeated eigenvalues are positive, then the fixed point is an unstable source. If the two repeated eigenvalues are negative, then the fixed point is a stable sink.

Summary of Eigenvalue Graphs

Below is a table summarizing the visual representations of stability that the eigenvalues represent.

Eigenvalues	Graph
All real and negative	
All real and one or more are positive	
All real eigenvalues are negative and there are imaginary parts	
One or more eigenvlaues have a positive real part and there are imaginary parts	
Real parts of the eigenvlaues are zero and there are imaginary parts	

Note that the graphs from Peter Woolf's lecture from Fall'08 titled Dynamic Systems Analysis II: Evaluation Stability, Eigenvalues were used in this table.

Another method of determining stability

The process of finding eigenvalues for a system of linear equations can become rather tedious at times and to remedy this, a British mathematician named Edward Routh came up with a handy little short-cut.

First, recall that an unstable eigenvalue will have a positive or zero real part and that a stable eigenvalue will have a negative real part.

The first test is to take an n-th degree polynomial of interest:

$$P(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

and look to see if any of the coefficients are negative or zero. If so, there is at least one value with a positive or zero real part which refers to an unstable node.

The way to test exactly how many roots will have positive or zero real parts is by performing the complete Routh array. Referring to the previous polynomial, it works as follows:

Row	
1	$a_o \quad a_2 \quad a_4 \quad a_6 \quad \ldots$
2	$a_1 \quad a_3 \quad a_5 \quad a_7 \quad \ldots$
3	$a_1 \quad b_2 \quad b_3 \quad b_4 \quad \ldots$
4	$a_1 \quad c_2 \quad c_3 \quad c_4 \quad \ldots$
$\vdots$	$\vdots \quad \vdots \quad \vdots \quad \vdots$
n-1	$p_1 \quad p_2$
n	q_1
n+1	v_1

An array of n+1 rows and the coefficients placed as above. After the first two rows, the values are obtained as below:

$$b_1 = \frac{a_1a_2 - a_0a_3}{a_1}, b_2 = \frac{a_1a_4 - a_0a_5}{a_1}, b_3 = \frac{a_1a_6 - a_0a_7}{a_1}, \cdots c_1 = \frac{b_1a_3 - a_1b_2}{b_1}, c_2 = \frac{b_1a_5 - a_1b_3}{b_1}, c_3 = \frac{b_1a_7 - a_1b_4}{b_1}, \cdots$$

Routh's theorem says:

Routh's Theorem

1. For all of the roots of the polynomial to be stable, all the values in the first column of the Routh array must be positive.
2. If any of the values in the first column are negative, then the number of roots with a positive real part equals the number of sign changes in the first column.

So considering the following example,

$$f(x) = 9x^4 + 14x^3 + 7x + 10$$

Preliminary test: All of the coefficients are positive, however, there is a zero coefficient for x^2 so there should be at least one point with a negative or zero real part.

Routh array:

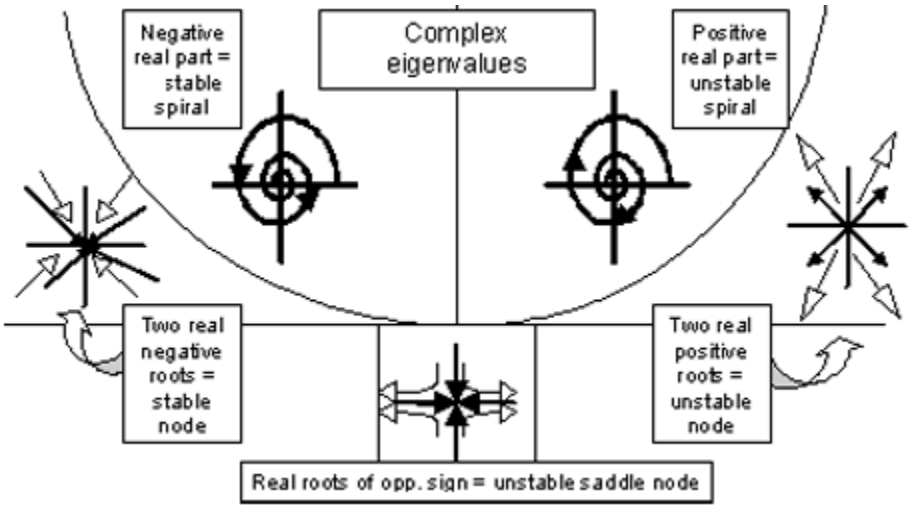
Row	
1	$9 \quad 0 \quad 10$

Row	
2	14 7
3	−4.5 10
4	38.1
5	10

Since Row 3 has a negative value, there is a sign change from Row 2 to Row 3 and again from Row 3 to Row 4. Thus, there are 2 roots with positive or zero real part.

Stability Summary

The following image can work as a quick reference to remind yourself of what vector field will result depending on the eigenvalue calculated.



The table below gives a complete overview of the stability corresponding to each type of eigenvalue.

Eigenvalue Type	Stability	Oscillatory Behavior	Notation
All Real and +	Unstable	None	Unstable Node
All Real and -	Stable	None	Stable Node
Mixed + & - Real	Unstable	None	Unstable saddle point
+a + b*i*	Unstable	Undamped	Unstable spiral
-a + b*i*	Stable	Damped	Stable spriral
0 + b*i*	Unstable	Undamped	Circle
Repeated values	Depends on orthogonality of eigenvectors		

Advantages and Disadvantages of Eigenvalue Stability

There are several advantages of using eigenvalues to establish the stability of a process compared to trying to simulate the system and observe the results. However, there are situations where eigenvalue stability can break down for some models.

Advantages

1. High accuracy for linear systems.
2. General method that can be applied to a variety of processes.
3. Can be used even if all variables are not defined, such as control parameters.

Disadvantages

1. Only applicable for linear models.
2. Linear approximations of nonlinear models break down away from the fixed point of approximation.

Example 10.4.1

For the system described by:

$$\frac{dx}{dt} = y$$
$$\frac{dy}{dt} = 2x + y$$

Find the fixed points and determine their stability.

Solution

Step 1. Determining the fixed points

At the fixed points, nothing is changing with respect to time. Therefore, set the derivatives to zero to find the fixed points.

$$0 = y$$
$$0 = 2x + y$$

Solving these two equations simultaneously, we see that we have one fixed point at {0,0}

Step 2. Determine the eigenvalue of this fixed point

First, let us rewrite the system of differentials in matrix form.

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Next, find the eigenvalues by setting $\det(A - \lambda I) = 0$

$$det \begin{bmatrix} 0-\lambda & 1 \\ 2 & 1-\lambda \end{bmatrix} = 0$$

$$-\lambda(1-\lambda) - 2 = 0$$

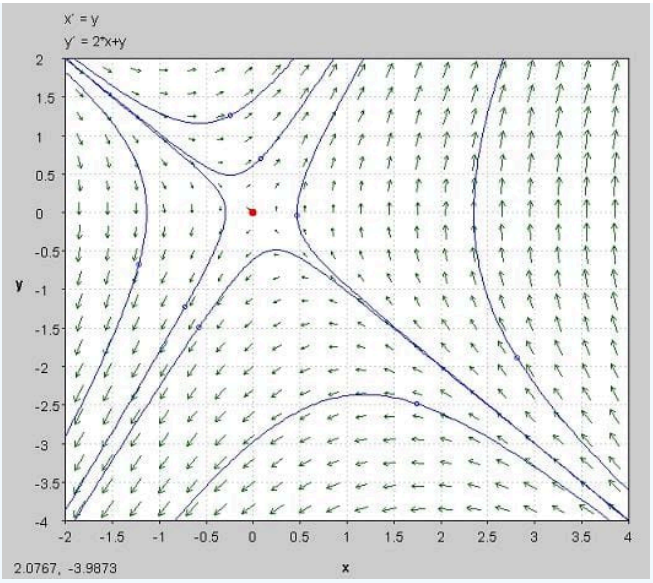
$$\lambda^2 - \lambda - 2 = 0$$

Using the quadratic formula, we find that $\lambda_1 = 2$and $\lambda_2 = -\frac{1}{2}$

Step 3. Determine the stability based on the sign of the eigenvalue

The eigenvalues we found were both real numbers. One has a positive value, and one has a negative value. Therefore, the point {0, 0} is an unstable saddle node.

The stability can be observed in the image below. The fixed point is seen at (0,0). All solutions that do not start at (0,0) will travel away from this unstable saddle point. The solution was found by using the two-dimensional system in PPlane 2005.10.



Example 10.4.2

Determine the Routh array and the number of positive or zero roots of the following equation.

$$f(x) = 6x^5 + 12x^4 + 5x^3 + 3x^2 + 17x$$

Solution

Routh Array

Row	
1	$6 \quad 5 \quad 17$
2	$12 \quad 3 \quad 0$
3	$3.5 \quad 17$
4	-26.1
5	17
6	0

Since you go from a positive value in row three, to a negative value in row four, and back to a positive value in row five, you will have a positive or zero real part for two of your roots.

Example 10.4.1

Use Mathematica to find the eigenvalues of the system defined by:

$$\frac{dx_1}{dt} = 8x_1 + 15x_2 - 3x_3 + 7x_4 + 2x_5$$
$$\frac{dx_2}{dt} = -22x_1 - 21x_2 + 3x_3 - 12x_4 + 11x_5$$
$$\frac{dx_3}{dt} = 10x_1 + 6x_2 + 24x_3 + 3x_4 - 6x_5$$
$$\frac{dx_4}{dt} = 0x_1 - 2x_2 + 21x_3 + 0x_4 + 4x_5$$
$$\frac{dx_5}{dt} = 4x_1 + 9x_2 + x_3 - 22x_4 - 7x_5$$

And comment on the stability of this system

Solution

The matrix that corresponds with this system is the square matrix:

$$\begin{bmatrix} 8 & 15 & -3 & 7 & 2 \\ -22 & -21 & 3 & -12 & 11 \\ 10 & 6 & 24 & 3 & -6 \\ 0 & -2 & 21 & 0 & 4 \\ 4 & 9 & 1 & -22 & -7 \end{bmatrix}$$

Using the Eigenvalues[ ] function in Mathematica the input is:

In[1]:= Eigenvalues[{ {8, 15, -3, 7, 2}, {-22, -21, 3, -12, 11}, {10, 6, 24, 3, -6}, {0, -2, 21, 0, 4}, {4, 9, 1, -22, -7} }]}

In[2]:= N[%] This step produces numerical results

The output contains the 5 eigenvalues:

out[2]:= {27.0612, -10.7653 + 10.0084i, -10.7653 - 10.0084i, -0.765272 + 7.71127i, -0.765272 - 7.71127i}

So the five eigenvalues are:

- 27.0612
- -10.7653 + 10.0084i
- -10.7653 - 10.0084i
- -0.765272 + 7.71127i
- -0.765272 - 7.71127i

Looking at these eigenvalues it is clear that the system as a whole is unstable. This is because one of the eigenvalues has a positive real part.

Exercise 10.4.1

A system is stable if and only if all of the system's eigenvalues:

a. are real
b. are complex
c. have negative real parts
d. have negative imaginary parts

Answer available at the end of this section.

Exercise 10.4.2

What would the following set of eigenvalues predict for the system's behavior?

$$-4, -2+3i, -2-3i, 3$$

a. An unstable oscillation
b. A damped oscillation
c. An undamped oscillation
d. A source
e. A saddle point

Answer available at the end of this section.

References

- Kravaris, Costas. *Chemical Process Control: A Time Domain Approach*.
- Liptak, Bela G., *Process Control and Optimization. Vol. II*. New York: Taylor & Francis.

Answers

Exercise 10.4.1

C

Exercise 10.4.2

d

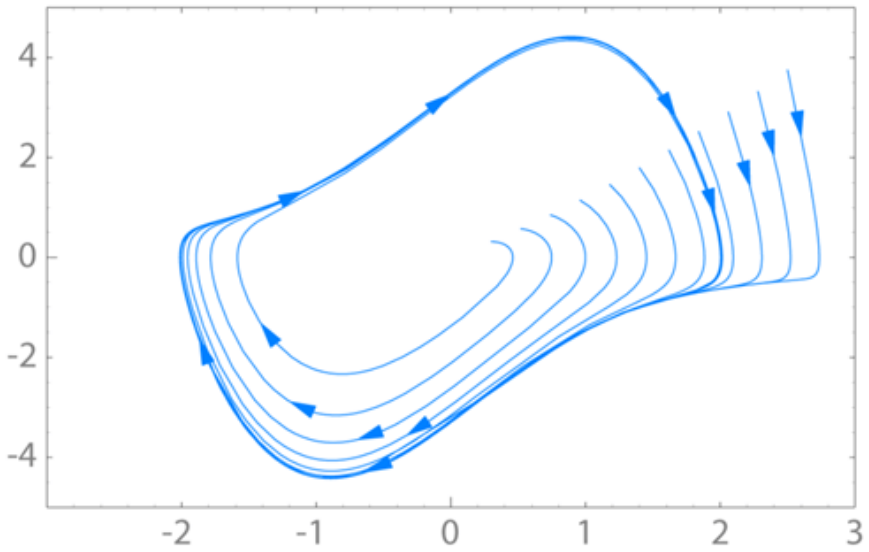
This page titled 10.4: Using Eigenvalues and Eigenvectors to Find Stability and Solve ODEs is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Daniel Katzman, Jason Noelanders, Jessica Moreno, & Mark Winston-Galant via source content that was edited to the style and standards of the LibreTexts platform.

10.5: Phase Plane Analysis - Attractors, Spirals, and Limit cycles

Introduction to Attractors, Spirals and Limit Cycles

We often use differential equations to model a dynamic system such as a valve opening or tank filling. Without a driving force, dynamic systems would stop moving. At the same time dissipative forces such as internal friction and thermodynamic losses are taking away from the driving force. Together the opposing forces cancel any interruptions or initial conditions and cause the system to settle into typical behavior. **Attractors** are the location that the dynamic system is drawn to in its typical behavior. Attractors can be fixed points, limit cycles, spirals or other geometrical sets.

Limit cycles are much like sources or sinks, except they are closed trajectories rather than points. Once a trajectory is caught in a limit cycle, it will continue to follow that cycle. By definition, at least one trajectory spirals into the limit cycle as time approaches either positive or negative infinity. Like a sink, attractive (stable) limit cycles have the neighboring trajectories approaching the limit cycle as time approaches positive infinity. Like a source, non-attractive (unstable) limit cycles have the neighboring trajectories approaching the limit cycle as time approaches negative infinity. Below is an illustration of a limit cycle [1].



Spirals are a similar concept. The attractor is a spiral if it has complex eigenvalues. If the real portion of the complex eigenvalue is positive (i.e. 3 + 2i), the attractor is unstable and the system will move away from steady-state operation given a disturbance. If the real portion of the eigenvalue is negative (i.e. -2 + 5i), the attractor is stable and will return to steady-state operation given a disturbance.

Given the following set of linear equations we will walk through an example that produces a spiral:

$$\frac{dx}{dt} = 2x + 5y$$
$$\frac{dy}{dt} = -5x + 2y$$

The Jacobian matrix would be the coefficients:

$$A = \begin{vmatrix} 2 & 5 \\ -5 & 2 \end{vmatrix}$$

Next we found the eigenvalues:

$$(A-\lambda I)=\begin{vmatrix} 2 & 5 \\ -5 & 2 \end{vmatrix}-\lambda\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}=\begin{vmatrix} (2-\lambda) & 5 \\ -5 & (2-\lambda) \end{vmatrix}$$

where I is the identity matrix

$$I=\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$$

$$\det(A-\lambda I)=(2-\lambda)^2+25=0$$

Eigenvalues:

$$\lambda=2\pm 5i$$

The system is unstable because the real portion of the complex eigenvalues is positive.

To find the first eigenvector we continue by plugging in $2-5i$:

$$\begin{vmatrix} (2-\lambda) & 5 \\ -5 & (2-\lambda) \end{vmatrix}=\begin{vmatrix} 2-(2-5i) & 5 \\ -5 & 2-(2-5i) \end{vmatrix}=\begin{vmatrix} 5i & 5 \\ -5 & 5i \end{vmatrix}$$

$$(A-\lambda I)v=\begin{vmatrix} 5i & 5 \\ -5 & 5i \end{vmatrix}v=0$$

let

$$v=\begin{vmatrix} x \\ y \end{vmatrix}$$

$$\begin{vmatrix} 5i & 5 \\ -5 & 5i \end{vmatrix}\begin{vmatrix} x \\ y \end{vmatrix}=\begin{vmatrix} 0 \\ 0 \end{vmatrix}$$

We now have a system of equations which we can solve for x, y:

$$\begin{aligned} 5ix+5y&=0 \\ -5x+5iy&=0 \end{aligned}$$

Dividing both equations by 5:

$$\begin{aligned} ix+y&=0 \\ -x+iy&=0 \end{aligned}$$

Solution

$$v_1=\begin{vmatrix} -1 \\ i \end{vmatrix}$$

Following the same procedure using the second eigenvalue of 2 + 5*i*, we find the second eigenvector to be:

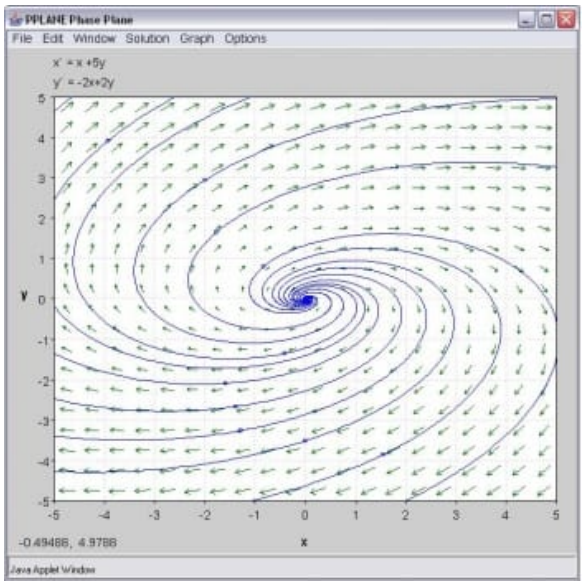
$$v_2=\begin{vmatrix} i \\ -1 \end{vmatrix}$$

Now plugging both eigenvalues and eigenvectors into the characteristic equation:

$$\begin{aligned} x(t)&=e^{2t}\left(C_1\cos 5t+C_2\sin 5t\right) \\ y(t)&=e^{2t}\left(C_3\cos 5t+C_4\sin 5t\right) \end{aligned}$$

For more on this procedure, see: Eigenvalues and Eigenvectors

The phase-plane plot is shown below:



Introduction to Pplane

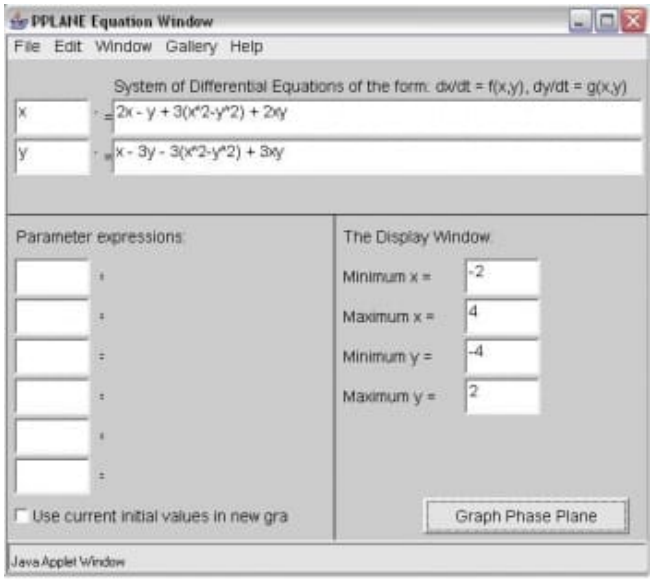
Phase-plane analysis is an important tool in studying the behavior of nonlinear systems since there is often no analytical solution for a nonlinear system model.

PPlane is a JAVA applet for phase plane analysis of two-dimensional systems. It starts in your web browser and you can directly input your equations and parameter values. PPLANE plots vector fields for systems of differential equations. At each point, (x,y), of a grid, PPLANE draws an arrow indicating the direction and magnitude of the vector (x',y'). This vector equals dy/dt / dx/dt = dy/dx, and is independent of t; therefore, it must be tangent to any solution curve through (x,y).

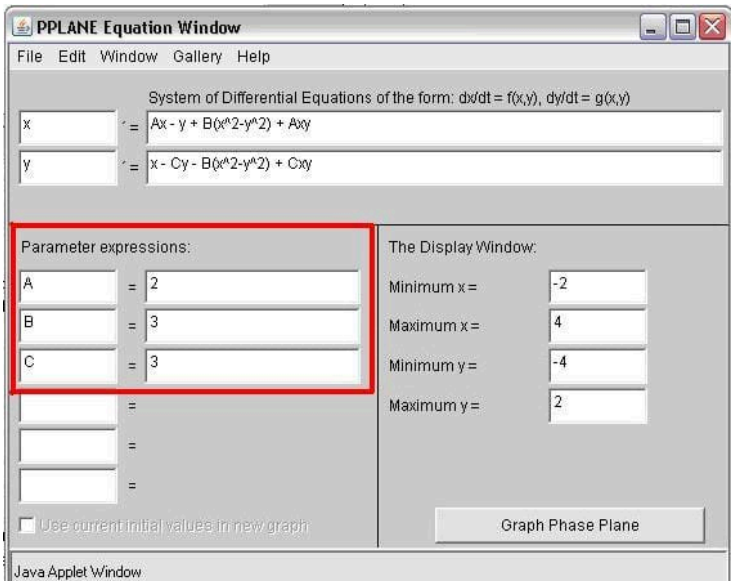
It allows the user to plot solution curves in the phase plane by simple clicking on them. It also enables the user to plot these solutions in a variety of plots. There are a number of advanced features, including finding equilibrium points, eigenvalues and nullclines, that you will find useful later.

How to use Pplane

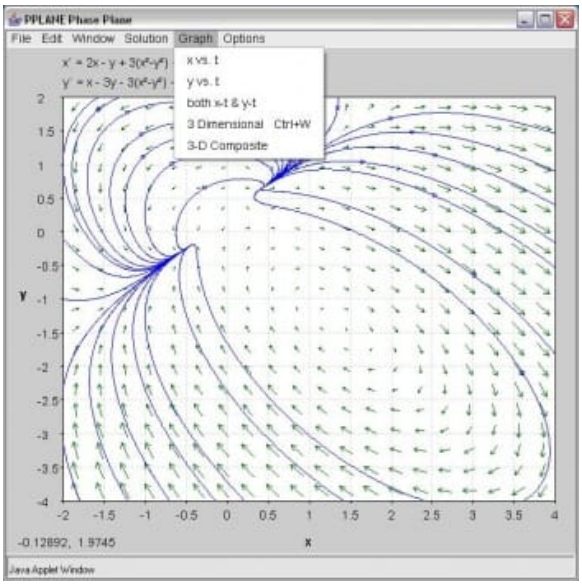
In the PPlane equation window you can enter a system of differential equations of the form $dx/dt = f(x, y)$ and $dy/dt = g(x, y)$, define parameters and resize the display window. Under the Gallery pull down from the menu, you can switch to a linear system.



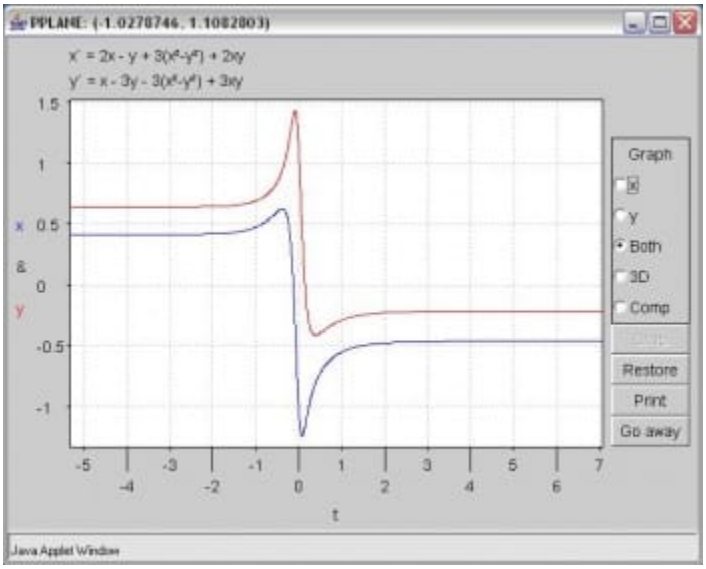
Note, if your differential equations contain constant parameters, you can enter them in the "Parameter Expressions" boxes below the differential equations as seen in the figure below (A, B, and C are used as example parameters). This is a convenient feature to use when considering the effect of changed parameters on the steady state of a system because it eliminates the redundancy of re-entering the parameter values multiple times within the differential equations.



In the PPlane Phase Plane window below you will see the vector fields for the system. By clicking on the field you will plot solution curves in the phase plane. If you are interested in a plot of your solution vs. time or a 3-D view, click on graph:



If you choose the x-t and y-t option, you have to pick a specific solution curve. The result will look like this:



Use the crop function to zoom in on a point of interest

Choose Find Equillibrium Point under the Solution pull down menu. Then when you click on an orbit in the phaseplane, the Pplane Messages window will display the eigenvalues and possible equillibrium points.

Additional Things you can Change in PPlane

Changing the Slope Field

By clicking on the "Options" tab, then by selecting "Direction Field Settings," you can change the number of rows and columns plotted, the way the field is made up, as well as the computational settings of PPlane.

Erasing Made Orbits

On the "Edit" tab, there are options that say "Delete Orbit" or "Delete All Orbits". These options act as their names imply.

Changing the Direction of Graphing

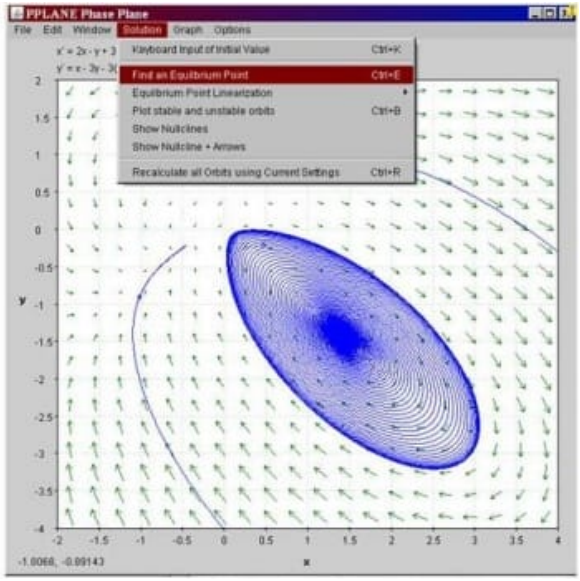
By clicking on the "Options" tab, then by selecting the "Solution Direction" option, you can then change the way that PPlane graphs a line when you click on the field. You can change the graphing to plot forwards (for values of t>0), backwards (for values of t<0) or in both directions of t.

More Uses for PPLANE

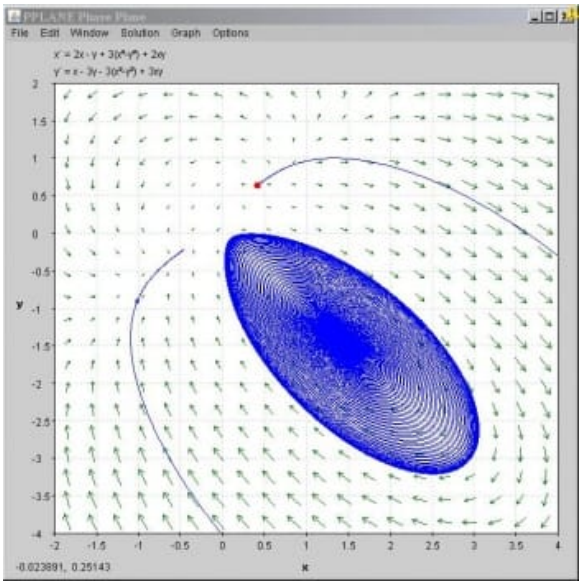
Of the many uses that PPLANE has to offer, some of the most helpful functions involve:

Finding eigenvalues/eigenvectors for an equilibrium point.

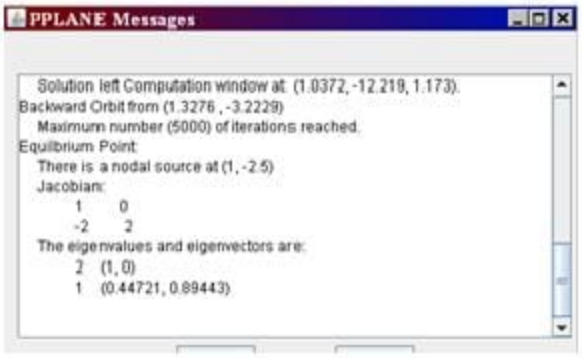
After graphing a series of differential equations, find a particular equilibrium point within the graphed data (“Find an Equilibrium Point,” under the Solution tab)



Now, by selecting a point on the field that has ben graphed by pplane, pplane will find the closest equilibrium point on the graph, and highlight this point on the graph in red.



The PPLANE Messages box in the upper left hand corner of the screen should pop up with some new information. This information provides eigenvalues and the corresponding eigenvectors to the selected equilibrium value:



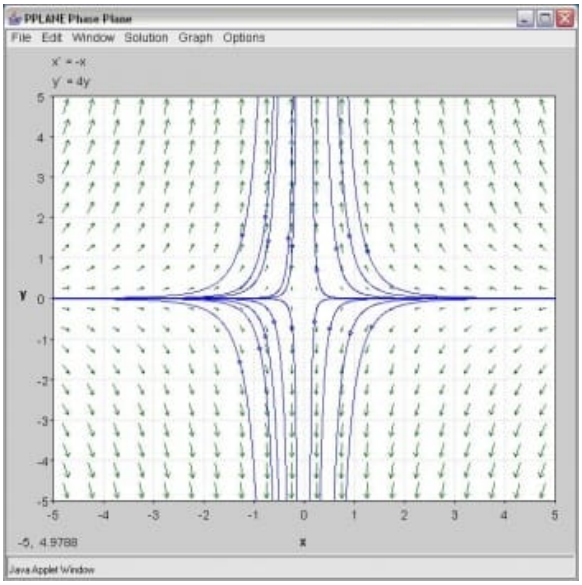
Stability of a Equilibrium Point

Similarly to before, the Messages Box will provide the stability features (ie: is it a nodal sink?) of the found equilibrium point, immediately after using the “Find an Equilibrium Point.”

Other concepts of phase plane analysis

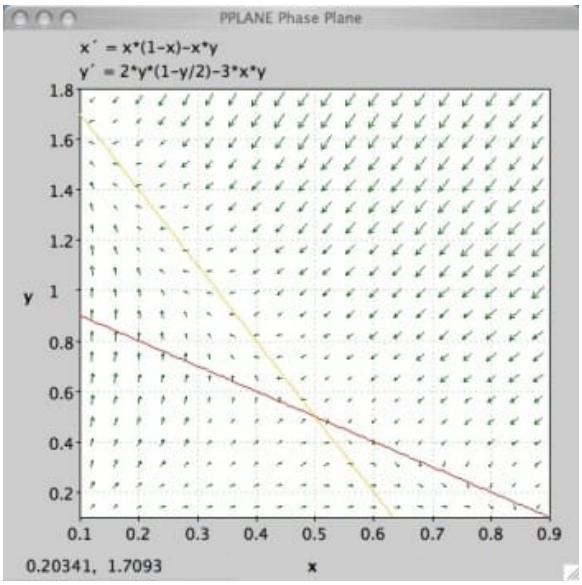
Separatrix

A separatrix is any line in the phase-plane that is not crossed by any trajectory. The unstable equilibrium point, or saddle point, below illustrates the idea of a separatrix, as neither the x or y axis is crossed by a trajectory. If you picture a topographic map, the seperatrix would be a mountain ridge; if you fall a little of the edge, you will never come back. Plotting your phase plane in Pplane would be useful to identify impossible set points, for example.



Nullclines

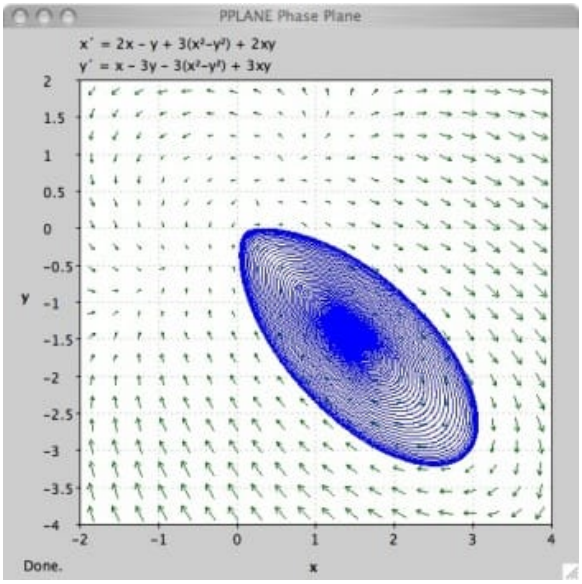
A nullcline is a curve where x'=0 or y'=0, thus indicating where the phase plane is completely horizontal or completely vertical. The point at which two nullclines intersect is an equilibrium point. Nullclines can also be quite useful for visualization of a phase plane diagram as they split the phase plane into regions of similar flow. To display nullclines on the Phase Plane window, select Nullclines under the Solutions drop down menu. The screenshot below is an example.



Notice that the red nullcline shows where the flow is completely vertical (x'=0) and the yellow nullcline shows where the flow is completely horizontal (y'=0).

Limit Cycle

Below you will find a solution curve for a limit cycle. The limit cycle contains the response in a set range, which is something you may want to take advantage for certain engineering applications. On the other hand it is always rotating and may not be stable enough for your purposes.



Taking Screen Shots to copy Pplane phase portraits

With the introduction of Windows Vista, the Snipping Tool was introduced. This tool allows much greater flexibility with taking screen shots and editing them. This article will talk about the Snipping Tool as well as the Windows Print Screen key which can be used to take photos of your computer screen. When pressing the key, your computer copies the image of your screen and onto your computer's clipboard. The image can then be pasted into multiple programs. There are many instances throughout the CHE 466 course in which taking a screen shot of your work will come in handy. Examples include copying phase portraits created in Pplane, graphs created in Mathematica, or your Mathematica code.

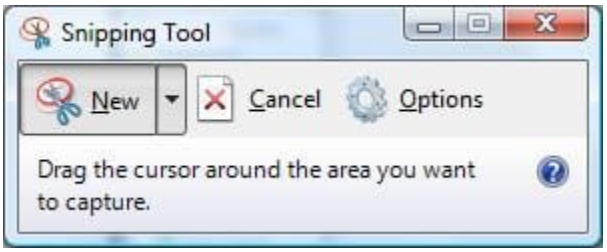
To enable the Snipping Tool on your Vista computer go to the Windows button in the bottom left of your screen and click Accessories -> Snipping Tool.

Figure 1. How to enable the Snipping Tool

A window will appear asking if you would like to add the Snipping Tool to your Quicklaunch. This provides a simple and quick way to take screenshots.

To take a picture of your graph, just press the Snipping Tool button in the Quicklaunch area and a window like this will appear:

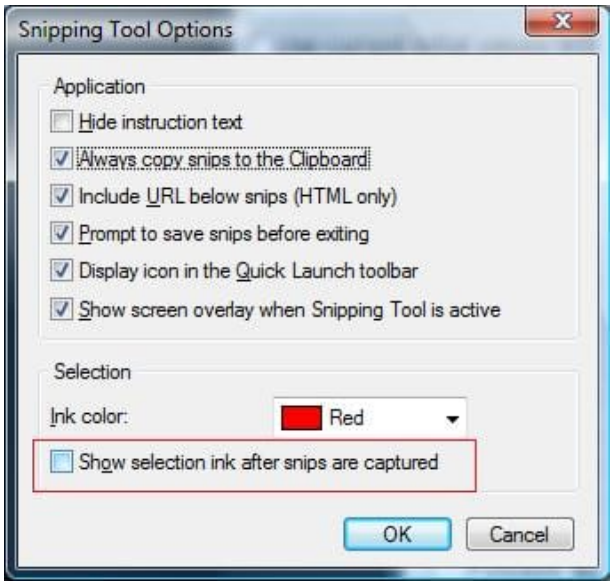
Figure 2. The Snipping Tool Window



Automatically, the Snipping Tool will default to a crosshair from which you can click and drag to make a selection of the section of the screen you would like represented by a red rectangle.

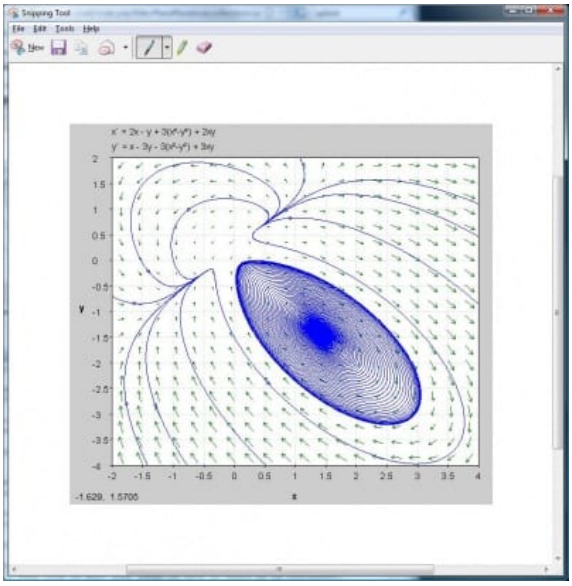
WARNING: In the Options section you should uncheck "Show selection ink after snips are captured" in order to eliminate the red edge around your photos.

Figure 3. Snipping Tool Option Menu (Uncheck the selection ink)



The Snipping Tool will open up a new window with your selection and copy the image to your clipboard. Feel free to edit your image or save it where it is convenient.

Figure 4. Snipping Tool Editing Window



If not using Windows Vista you can still use Print Screen:

Follow these simple steps to copy and paste your phase portrait into a Microsoft Word document:

1. Pull up the window containing your phase portrait so that it is displayed on the screen.
2. Find the Print Screen or PrtSc button in the upper-right hand portion of your keyboard. (The key may appear slightly different depending on your Windows keyboard manufacturer).
3. Open Microsoft Word to the document of your choice (i.e. CHE 466 Homework 7).
4. Paste the image into the Word document. Figure 1 below indicates how your phase portrait will look in Word.
5. To crop or resize the image as you like, you may use the Picture toolbar (seen in Figure 2) by selecting View -> Toolbars -> Picture.

If you prefer to take a screen shot of just your phase portrait rather than the entire computer screen, follow these simple steps:

1. Pull up the window containing your phase portrait so that it is displayed on the screen.
2. Press Alt-Print Screen to capture a photo of the window you selected.
3. Open Microsoft Word to the document of your choice (i.e. CHE 466 Homework 7).
4. Paste the image into the Word document. Figure 3 below indicates how your phase portrait image will look.

Figure 5. Initial screen shot

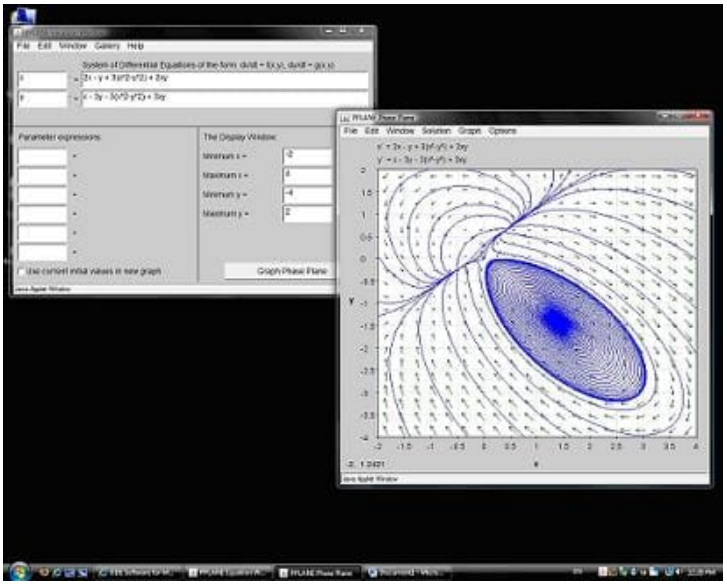


Figure 6. Microsoft Word document containing the screen shot

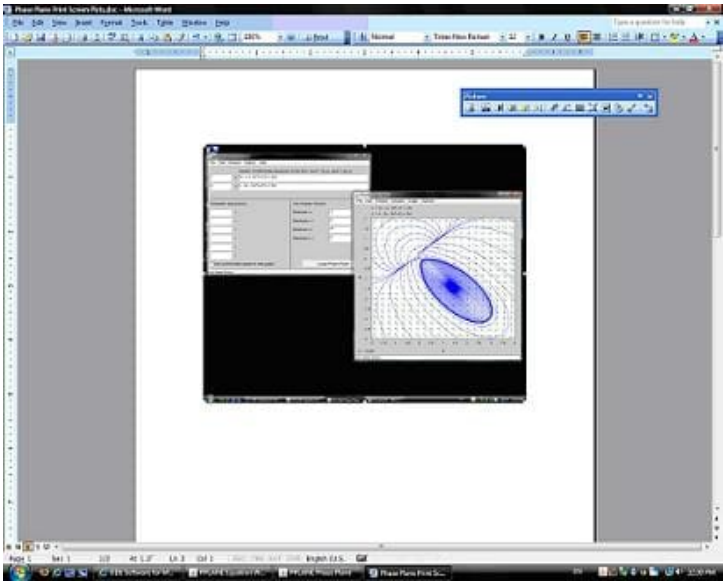
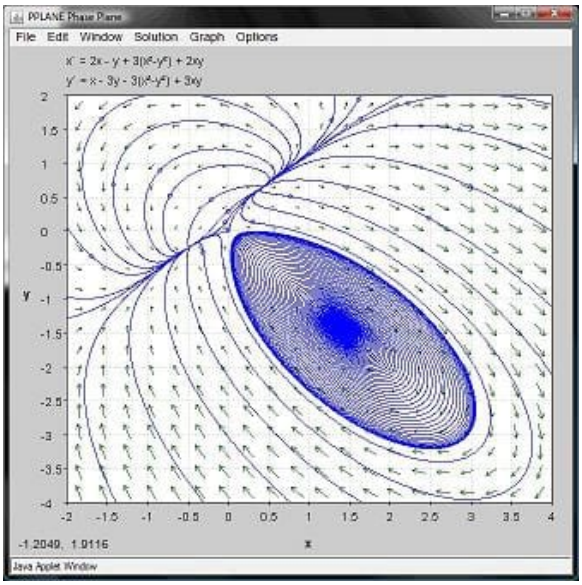


Figure 7. Screen shot of phase portrait window



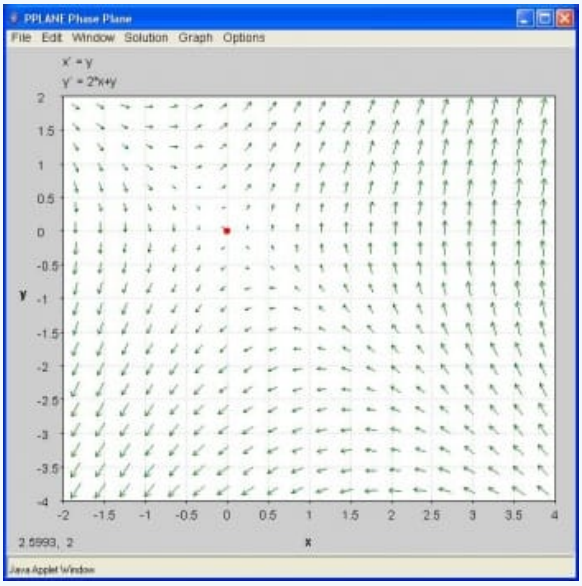
Example 10.5.1: Linear System of Equations

Use PPLANE to calculate the following information of the system given below: location and type of equilbrium point, Jacobian matrix, eigenvectors, and eigenvalues.

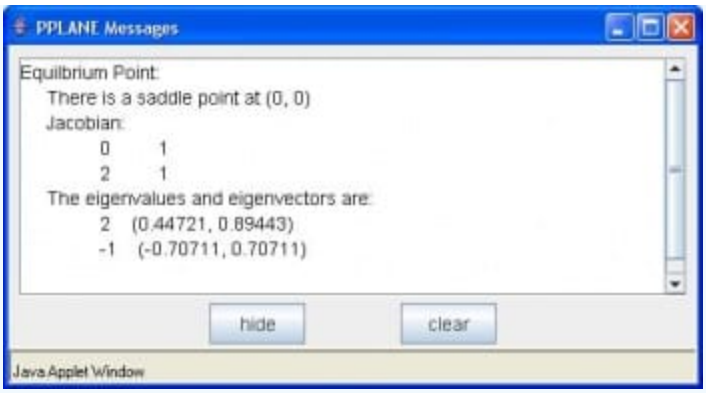
$$x' = y$$
$$y' = 2x + y$$

Solution

Using PPLANE's "Find an Equlibrium Point" feature and clicking on the phase plane, the following equilibrium point will be indicated:



The location, and type of the equilibrium point is given in the "PPLANE Messages" window, along with the Jacobian matrix, eignvectors and eigenvalues.



For our second example problem we would like you to try a non-linear system of equations.

Example \({10{.}5{.}\eatSpaces{2\): Nonlinear System of Equations

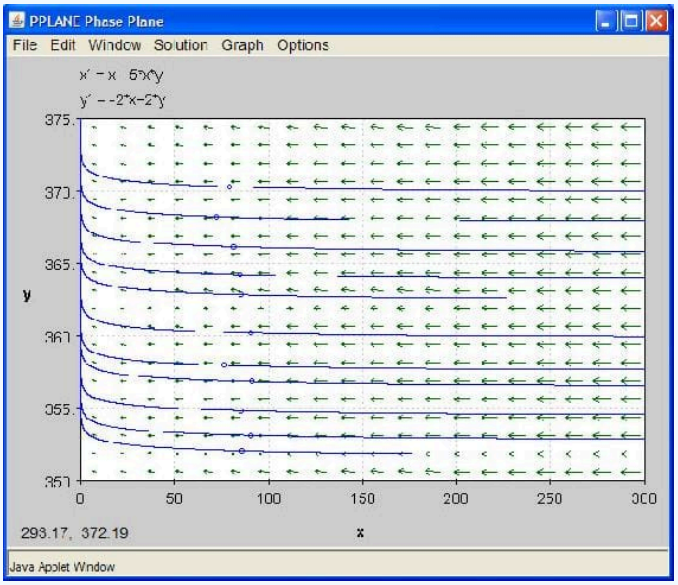
Solve for the set of equations on PPlane. Consider the trends of change in rate of the differential equations and subsequently solve the equations on Mathematica to compare the trends. The following two differential equations are going to be used to walk through the solutions on PPlane and Mathematica:

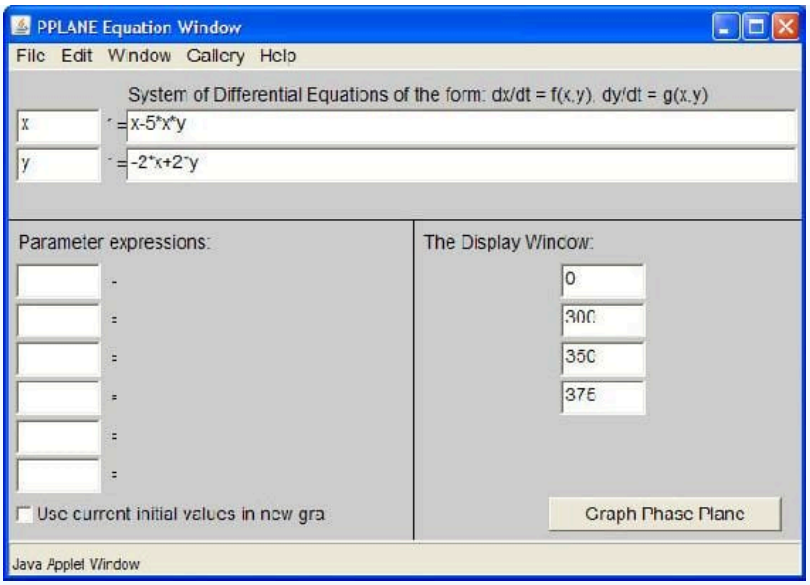
Extra close brace or missing open brace

$$\frac{dy}{dt} = -2x + 2y\)$$

Solution

The above equations were put in to the PPlane window and solved. The following two windows show the solution for the set of differential equations:



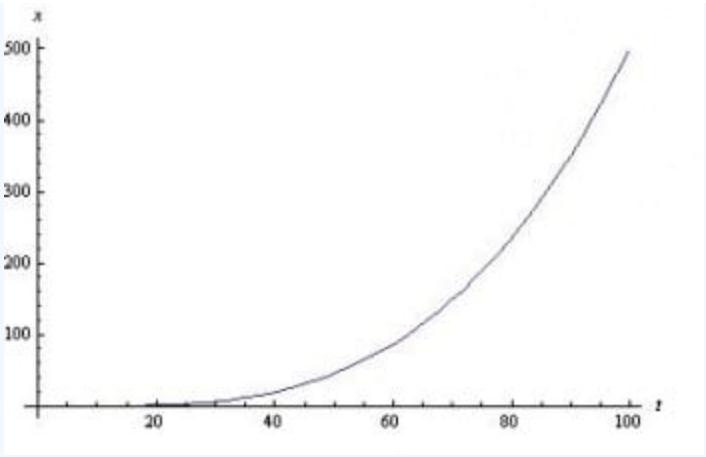


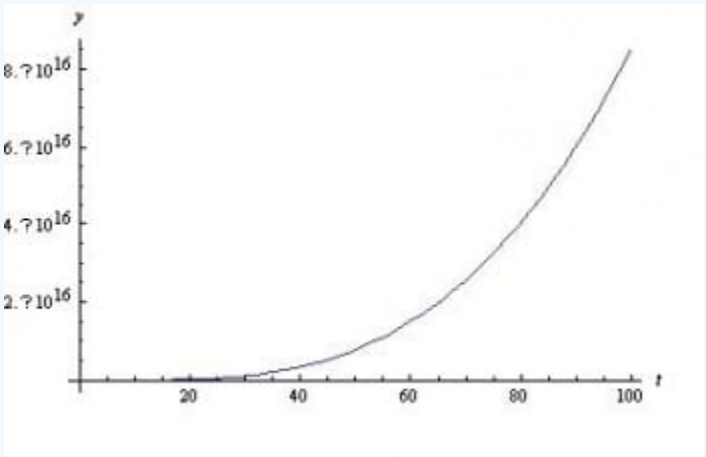
Looking at the phase-plane plot, at low values of x and y, t increases slowly. However at higher values of y, the increase in t is rapid. When x is high and y is low, however, t increases slowly. Mathematica will help us visualize the relative rates of change better.

The following is the code used in Mathematica to solve and plot the set of differential equations:

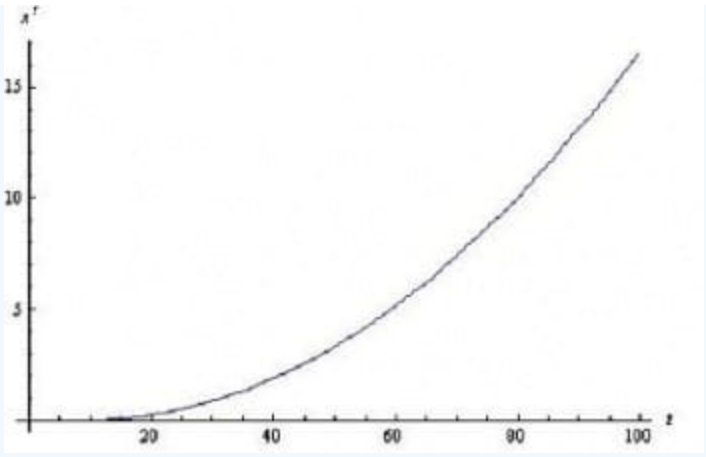
ODEs = {x'[t] == (x[t] - 5*x[t]*y[t]), y'[t] == (-2*x[t]) + (2*y[t]), x[0] == 9, y[0] == 370} numericalSol = NDSolve[ODEs, {x[t], y[t]}, {t, 1, 100}] Plot[y[t] /. numericalSol, {t, 1, 100}, PlotRange -> All]

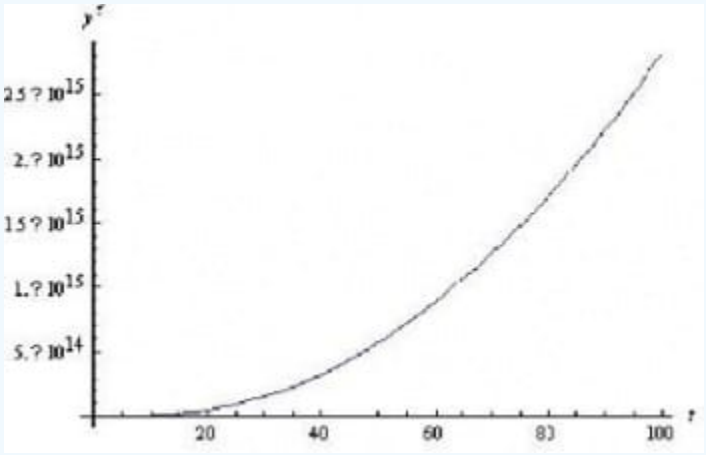
The graphs obtained on Mathematica for x versus t and y versus t are shown below. Please note the difference between the axes scales.





Also, for a closer look, here are the plots of the slopes of graphs above.





Comparing the increase in t with increase in x, we see a similar trend in the PPlane diagram. Mathematica shows a constantly increasing x' with t. At higher values of x, the value for t increases in the PPlane diagram. However, as seen clearly in the PPlane diagram and the graphs of the slope of x with respect to t and the slope of y with respect to t, the slope of x does not compare to the very large slope shown for y' versus t. Therefore the results using Mathematica and PPlane are consistant.

This modeling system could be used to view trends of variables in a CSTR or any other system which can be modeled using differential equations.

Multiple Choice Questions

Question 1

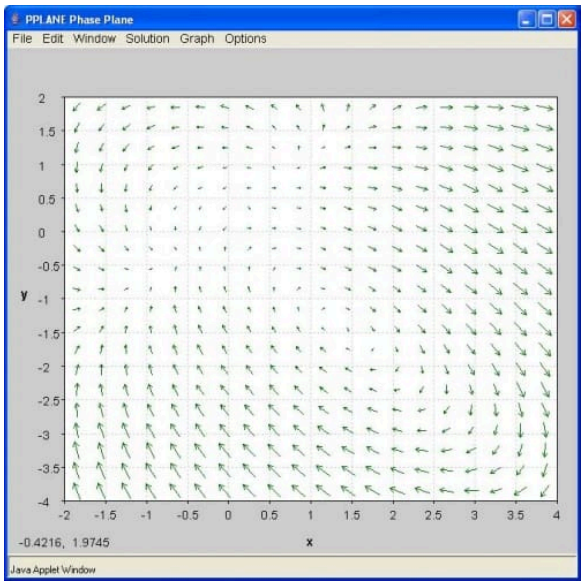
Open PPLANE and enter the following equations into the PPLANE Equation Window:

x '= sin(x)

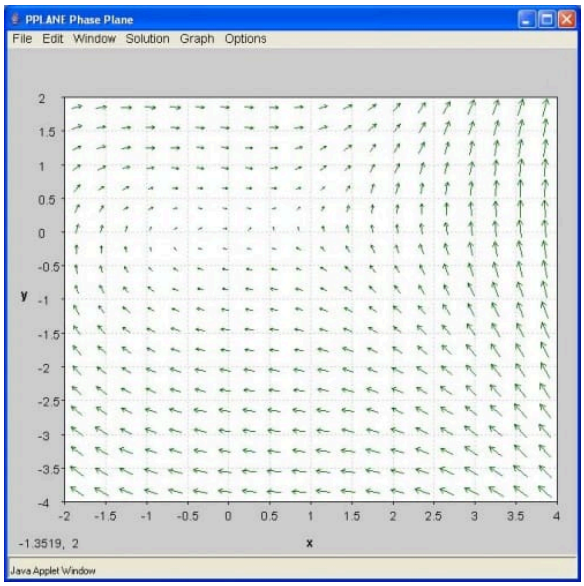
y' = cos(y)

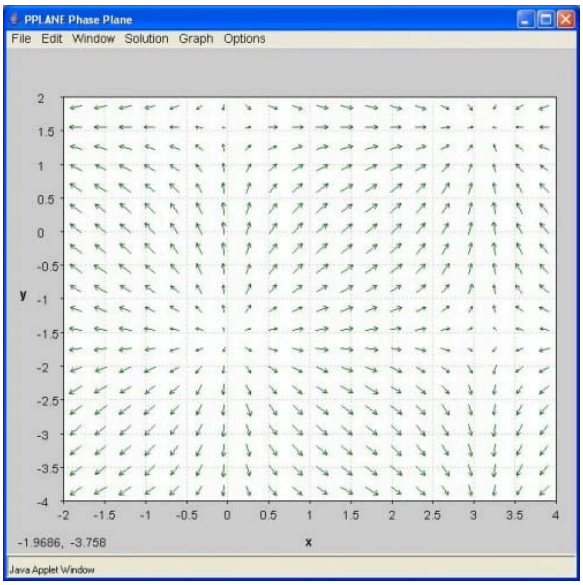
What does the resulting phase plane look like? (Note: Click on image to enlarge)

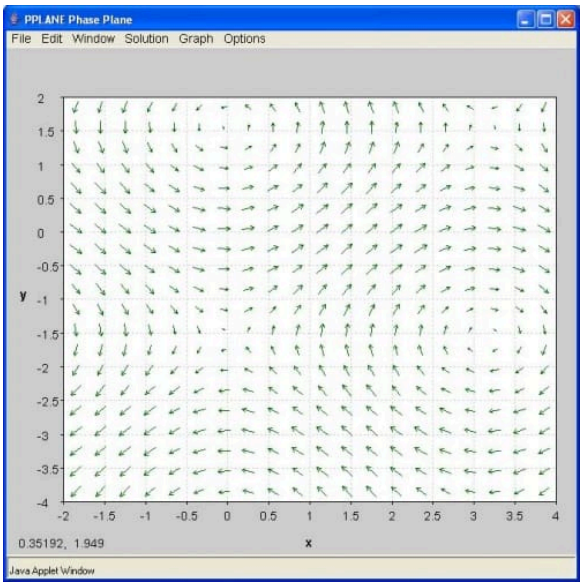
A.


PPLANE Phase Plane
File Edit Window Solution Graph Options
2
1.5
1
0.5
0
-0.5
y -1
-1.5
-2
-2.5
-3
-3.5
-4
-2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3 3.5 4
-0.4216, 1.9745
x
Java Applet Window

B.


PPLANE Phase Plane
File Edit Window Solution Graph Options
2
1.5
1
0.5
0
-0.5
y -1
-1.5
-2
-2.5
-3
-3.5
-4
-2 -1.5 -1 -0.5 0 0.5 1 1.5 2 2.5 3 3.5 4
-1.3519, 2
x
Java Applet Window

C. 

D. 

Question 2

If you have a disturbance in your system and the system is driven right back to equilibrium, that fixed point's eigenvalue is most likely a:

A. complex number with negative real number component
B. 0
C. negative real number
D. positive real number

Answers to the Multiple Choice Questions

Question 1: C

Question 2: A

Contributors and Attributions

- Authors: Erin Knight, Dipti Sawalka, Matt Russell, Spencer Yendell
- Stewards: Eric Black, Megan Boekeloo, Daniel Carter, Stacy Young

This page titled 10.5: Phase Plane Analysis - Attractors, Spirals, and Limit cycles is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

10.6: Root Locus Plots - Effect of Tuning

Introduction

Root locus plots show the roots of the systems characteristic equation, (i.e. the Laplacian), as a function of the control variables such as Kc. By examining these graphs it is possible to determine the stability of different values of the control variable. A typical transfer function is of the form $G(s) = Y(s)/U(s)$.

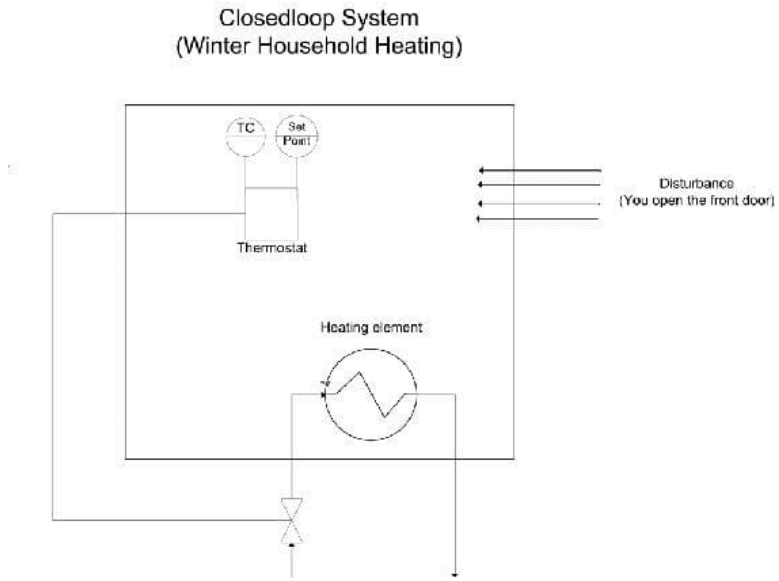
- Poles: U(s) = 0
- Zeros: Y(s) = 0

In other words, after factorization the poles are the roots of the denominators and the zeros are the roots of the numerator. Stability only depends on the value of the poles. The system is stable for all values of the control variables that result in the value of the real part of every pole value being less than zero.The lines of a Root locus plot display the poles for values of the control variable(s) from zero to infinity on complex coordinate system. These plots will always have a line of symmetry at $i = 0$.

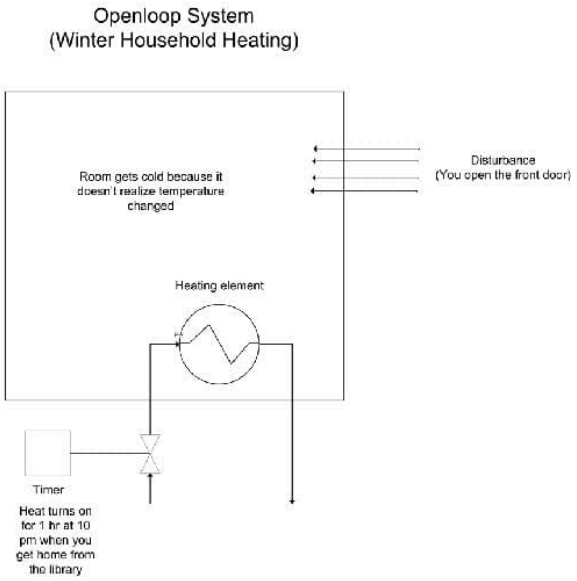
Closed-loop vs. Open-loop

A closed-loop system uses feedback control where the output has an effect on the input. With a closed-loop, oscillations are usually introduced, and therefore can become unstable. Unlike the open-loop systems, the closed-loop incorporates valves and controllers.

In an open-loop system, the output is not compared with and has no effect on the input. In an open-loop system, oscillations are not introduced and therefore cannot become unstable. An open-loop system, however, can be inaccurate because it does not take into account the control dynamics. Open-loop systems will include feed forward control schemes or timed control schemes. The two diagrams below depict this difference when trying to control the temperature in your apartment. (Click on images for increased resolution).



In this diagram a thermostat presents feedback to the heater to turn it on or off.



Since the heater only turns on at 10 pm, it assumes that the amount of time you open the door will be the same each day. It is considered an open system since the temperature in the room is independent of the heater controller.

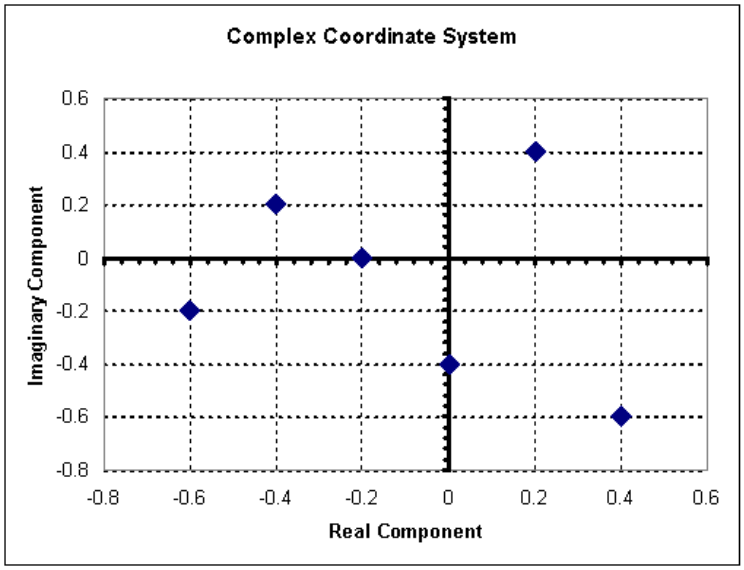
Note that all the examples presented in this web page discuss closed-loop systems because they include all systems with feedback.

Complex Coordinate Systems

Root locus plots are a plot of the roots of a characteristic equation on a complex coordinate system. A complex coordinate system allows the plotting of a complex number with both real and imaginary parts. The real component is plotted on the x-axis and the imaginary component is plotted on the y-axis. When creating root locus plots imaginary roots must be solved for. These imaginary roots come in complex conjugate pairs (this can be seen below in the section on "Plotting Poles on a Complex Coordinate System to make Root Locus Plot").

For example a plot of the following complex numbers is shown below. These complex numbers can be broken into the real and imaginary components to make it easier to plot.

Complex Number	Real Component	Imaginary Component
$-0.4 + 0.2i$	-0.4	0.2
$-0.6 - 0.2i$	-0.6	-0.2
$-0.2 + 0i$	-0.2	0
$0 - 0.4i$	0	-0.4
$0.2 + 0.4i$	0.2	0.4
$0.4 - 0.6i$	0.4	-0.6



Developing a Characteristic Equation

Although the focus of this article is to discuss root locus plots, it is necessary to mention briefly how to determine the characteristic equation for a system in order to obtain the root locus plot. In general, most chemical engineering processes can be described by a system of ordinary differential equations. Follow the following steps to determine the characteristic equation for the system (which will allow you to develop a root locus plot).

1. If the ODEs are not linear, linearize them. (See this link for more information on linearizing ODEs)
2. After linearizing the ODEs, use matrix algebra to find the eigenvalues of your system. Be careful here not to insert numerical values for your control parameters, (e.g. leave Kc as Kc, not Kc=1). (For more information on finding eigenvalues)
3. The polynomial equation obtained for the eigenvalue should contain lambda and the control parameters. This equation is the characteristic equation. Obtain solutions to this equation by setting values for the control parameters and solving for the eigenvalues. The roots obtained will be used to create the root locus diagram.

This 3-step process is valid to obtain a characteristic equation for any closed loop control system. A more traditional method to develop characteristic equations is by applying Laplace transforms.

Laplace Transforms

Laplace transforms are a method to change linear ordinary differential equations into transfer function. All transfer functions used in root locus plots are independent of time because the \(L[f(t)] ⇳ F(s)\). Formally, the equation below shows that the time function is integrated, leaving only the variable s.

$$\mathbf{F(s)} = \int_0^\infty f(t)e^{-st}dt$$

s is a complex number, therefore allowing us to construct complex coordinate system graphs. The exact solution to most disturbances and controllers can be found in any controls book. Please always refer to your Ordinary differential equations mathematics book for more information on Laplace transforms.

Example 10.6.1

The stability of the series chemical reactors is to be determined. The reactors are well mixed and isothermal, and the reaction is first-order in component A. The outlet concentration of the second reactor is controlled with a PI feedback algorithm that manipulates the flow of the reactant, which is very much smaller than the flow of the solvent. The sensor and final element are assumed fast, and process data is as follows.

Process

- V= 5m^3
- Fs=5m^3/min >> FA
- k = 1 min^{-1}
- vs= 50% open
- CA$_0$ = 20 mole/m^3
- CA$_0$(s)/v(s) = Kv= 0.40 (mole/ m^3) (%open)
- PI Controller Kc = ?? Tf = 1 min

Formulation The transfer function for the process and controller are

$$G_y(s) = \frac{K_y}{(\tau s + 1)(\tau s + 1)}$$

$$G_c(s) = K_c\left(1 + \frac{1}{T_f s}\right)$$

$$K_y = K_v\left(\frac{F}{F + Vk}\right) = 0.1\frac{mole/m^3}{\%}$$

$$\tau = \frac{V}{F + VR} = 0.5\text{min}$$

The individual transfer functions can be combined to give the closed-loop transfer function for a set point change, which includes the characteristic equation. (where CV= Control variable & SP = set-point signal)

$$\frac{CV(s)}{SP(s)} = \frac{G_p(s)G_v(s)G_c(s)}{1 + G_r(s)G_v(s)G_c(s)G_s(s)} = \frac{K_p\left(1 + \frac{1}{s}\right)\frac{0.1}{(0.5s+1)^2}}{1 + K_p\left(1 + \frac{1}{s}\right)\frac{0.1}{(0.5s+1)^2}}$$

Characteristic equation

$$0 = 1 + K_p\left(1 + \frac{1}{s}\right)\frac{0.1}{(0.5s + 1)^2}$$

Root Locus Diagrams

Root Locus plots are a method of evaluating the behavior of a control system. The creation of a root locus plot begins by determining the poles of the control system for a given set of control parameters. These poles are then plotted on a complex coordinate system as seen in the previous section and analyzed to determine the behavior of the system.

Determining the Poles of a Control System

The “poles” of a system are the roots of the demoninator of the transfer function. In other words the poles are the values of "s" when the transfer function go to infinity (when the demoninator equals zero). For a system of differential equations, poles are the eigenvalues of the equation system. Consider the following solution to a system of differential equations:

$$f(s) = 48s^3 + 44s^2 + 12s + 1 \tag{10.6.1}$$

This equation is a third order polynomial, therefore it will have three poles (be aware that some of these poles may be imaginary numbers). These poles can be obtained by factoring the expression or using a computer program such as Maple. The three poles, or roots, of this equation are s = -0.167, -0.25, -0.5.

For a system of differential equations, finding the eigenvalues can be time consuming and the use of Matlab, Maple, or Mathematica is more efficient. One method would be to use the Mathematica **eigenvalues[]** function to solve the system for you. This function is explained in Eigenvalues and Eigenvectors section of the wiki text.

Plotting Poles on a Complex Coordinate System to make Root Locus Plot

A root locus plot is created by plotting the resulting poles on a complex coordinate system. For system of P-only control, the governing differential equations will depend on the proportional gain, Kc. Consider the following solution to a system of differential equations:

$$f(s) = 48s^3 + 44s^2 + 12s + 1 + 6Kc \tag{10.6.2}$$

Notice that Kc is a term in this equation. Therefore, there exists a set of poles for each value of Kc. If Kc = 0, then equation #2 reduces to equation #1, and poles are as listed above. Table 1 lists the three poles of the system for given values of Kc. These poles were calculated using a computer algebra system because it is impractical to try and evaluate these functions with analytical methods.

Table 1- Poles of Characteristic Equation

Kc	Pole 1	Pole 2	Pole 3
-0.028	-0.406	-0.406	-0.105
0.000	-0.167	-0.25	-0.500
0.004	-0.205	-0.205	-0.506
0.050	-.182+.126i	-.182-.126i	-0.553
0.200	-.140+.229i	-.140-.229i	-0.637
0.500	-.093+.325i	-.093-.325i	-0.731
1.000	-.045+.417i	-.045-.417i	-0.880
1.670	0-.500i	0+.500i	-0.917
2.000	.018-.532i	.018+.532i	-0.952

To create a root locus plot, each pole is broken down into its real (x-axis) and imaginary (y-axis) component (Table 2):

Kc	Real Part	Imaginary Part
0	-0.167	0
0	-0.25	0
0	-0.5	0
0.004	-0.205	0
0.004	-0.205	0
0.004	-0.506	0
0.05	-0.182	0.126
0.05	-0.182	-0.126
0.05	-0.553	0
0.2	-0.14	0.229
0.2	-0.14	-0.229
0.2	-0.637	0
0.5	-0.093	0.325
0.5	-0.093	-0.325
0.5	-0.731	0
1	-0.045	0.417
1	-0.045	-0.417
1	-0.88	0
1.67	0	0.5
1.67	0	-0.5
1.67	-0.917	0
2	0.018	0.532
2	0.018	-0.532
2	-0.952	0

Table 2- Real and Imaginary Components of each Pole

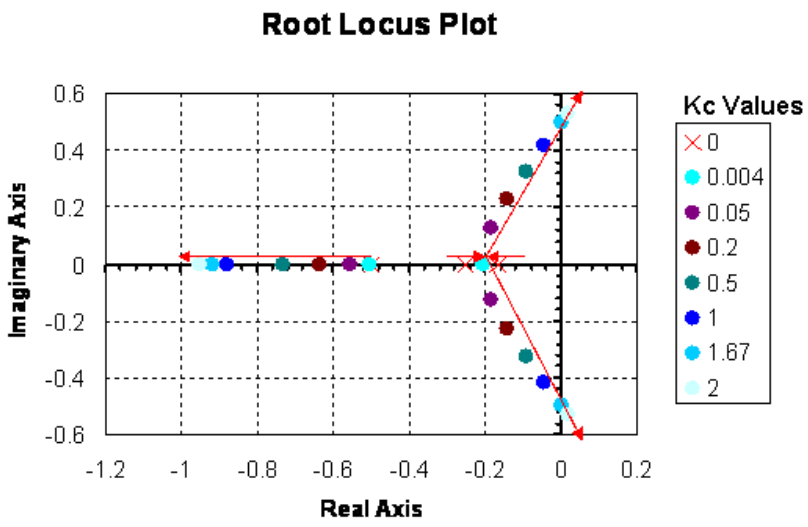


Figure 1- Root Locus Diagram of the Characteristic Equation

By convention, red arrows are drawn on the plot in the direction of increasing Kc values. They help to illustrate how the roots of the system vary by changing the Kc value. Also by convention, the points in which Kc = 0 are represented with 'x' marks instead of dots.

Interpreting a Root Locus Diagram

The primary use of a Root Locus Diagram is to evaluate how differing values of Kc affect the stability and behavior of a control system.

The stability of the control system depends on the sign of the real component of the pole. If the real components of all poles are negative, then the system is said to be stable for that value of Kc. If the real component of the pole is positive, the system is unstable for that value of Kc, meaning the output signal will diverge from the set point.

The behavior of the control system depends on the presence of an imaginary component of the pole. If any of the three poles contains an imaginary number component, then that value of Kc will cause the output signal to oscillate. If all of the poles are real (contain no imaginary components), the output signal will not oscillate at that Kc value.

Please refer back to the Root Locus diagram in Figure 1. The system becomes unstable (i.e. the real component of the poles becomes positive) for Kc > 1.67 and Kc < 0. Within the range of stability, no oscillations (i.e. no imaginary component) are observed when 0 < Kc < 0.004.

Root Locus Diagrams for PID Control

Root Locus diagrams are much more difficult to create for PID control. The characteristic equation will contain unknown variables Kc, Ti, and Td. Therefore, each point on the Root Locus Diagram will represent a set of tuning parameters. In order to show the progression as each tuning parameter changes, the resulting diagram will be a three-dimensional surface plot. Due to the complexity of this diagram, we will not create one, as it is beyond the scope of this text. However, stability analysis can still be applied to the characteristic equation. Take for example Equation 10.6.3:

$$f(s) = \frac{48s^3 + 44s^2 + 12s + 1 + 6Kc}{Ti + 3Td} \tag{10.6.3}$$

In this theoretical instance, equation 3 is the characteristic equation governing the same system mentioned above, only now with PID control. Notice the additional presence of the Ti and Td terms. Under P- only control, this system was stable for Kc values between 0 and 1.67. Will the same be true if this system were tuned with PID control?

Let's suppose we want to test the conditions Kc = 1.0, Ti = 0.3 and Td = 0.1. These values were randomly chosen. Using a computer

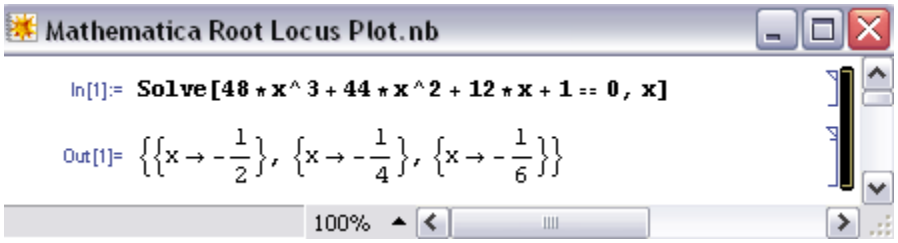
software package, the three roots were determined to be: r1 = -0.827, r2 = -0.044 + 0.417i, r3 = -0.044 – 0.417i. Therefore, since the real components are all negative, the system is still stable. Since complex roots are present, the response is expected to oscillate around the set point.

Creating Root Locus Plots with Mathematica

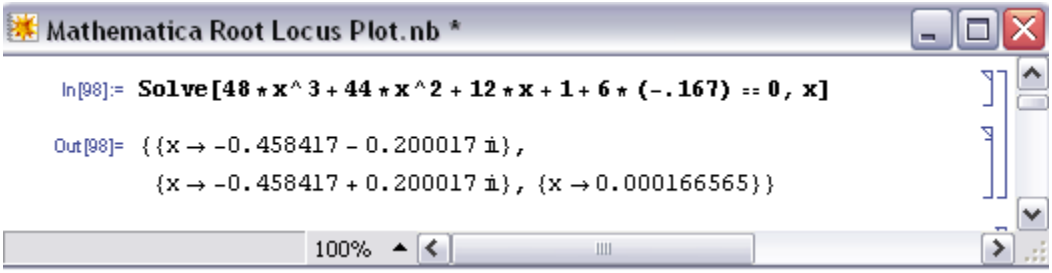
Mathematica allows you to develop root locus plots for polynomials since the math involved in solving for the solutions can become very tedious. Before you are able to obtain the root locus plot, you need to solve for the roots. We will use the same equation as was used in the "Determining the Poles of a Control System" section.

$$f(s) = 48s^3 + 44s^2 + 12s + 1 + 6K_c$$

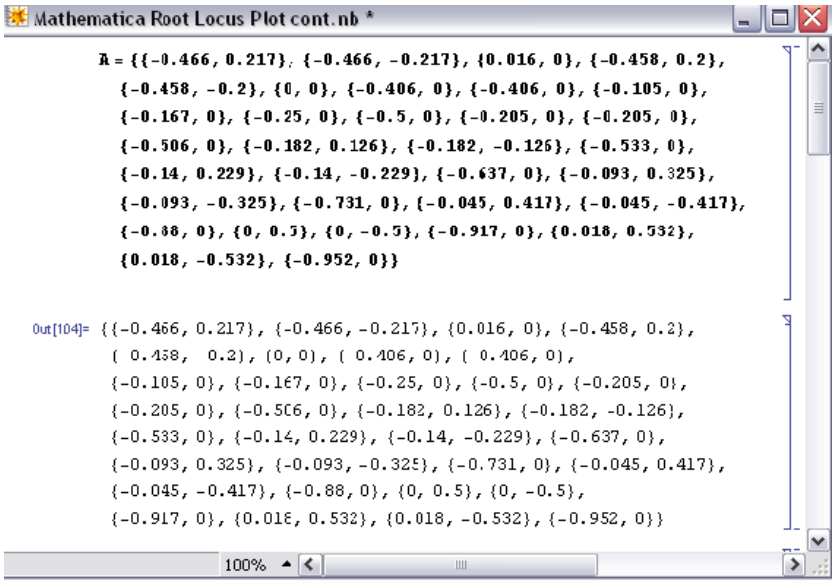
The **Solve[ ]** function can be used to determine the roots, both real and imaginary, for each corresponding Kc value. The syntax in Mathematica is shown below for the specified equation when Kc = 0. It is important to note that in order to obtain an output you must press "Shift Enter."

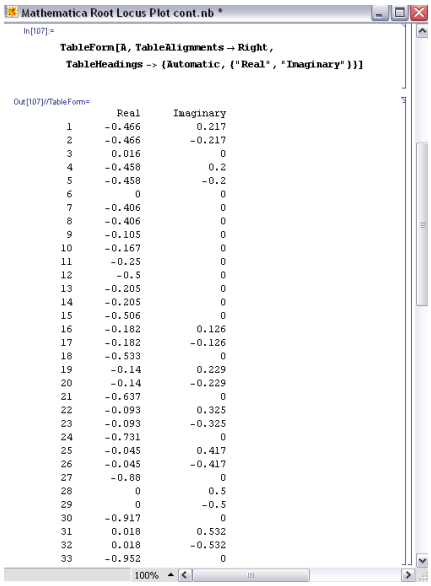


The following is just to show how Mathematica formats the output when $Kc \neq 0$ and when there are imaginary roots. In this case, $Kc = -0.167$.



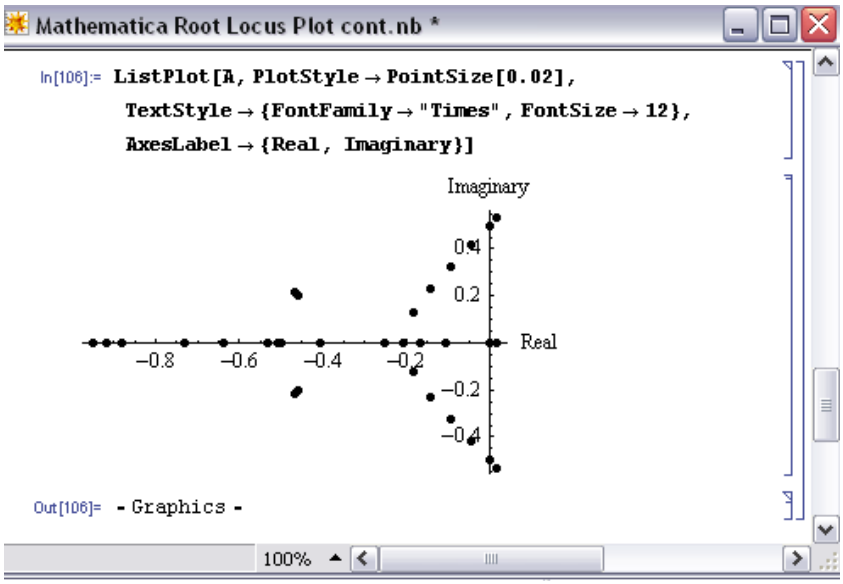
This can be done for all Kc values to obtain the corresponding real and imaginary roots. After all of the roots have been calculated, a table can be made to format the real roots (x axis) verses imaginary roots (y axis). When the syntax, **A={{x,y}{a,b}...}**, is used, you are inputting all of the x and y values and naming those values **A**. When you call **A** in the **TableForm[ ]** function, the table will be made according to all of the x and y values input into **A**. The syntax and corresponding outputs are shown below.





In order to produce the table shown above, you only need **TableForm[A]** and nothing else. The other descriptions (ie. TableAlignments and TableHeadings) are just used for formatting purposes.

Once the real (x) and imaginary (y) roots have been determined and put into table format for your equation, the **ListPlot[ ]** function can be used to develop your root locus diagram by calling **A**.



As you can see, the **ListPlot[ ]** function returned the same plot as shown in the "Complex Coordinate System to make Root Locus Plot" section (this plot was created in Excel). The arrows are not shown in this Mathematica plot; however the Excel plot shows the direction of increasing Kc values to show how the roots of the system vary by changing the Kc values.

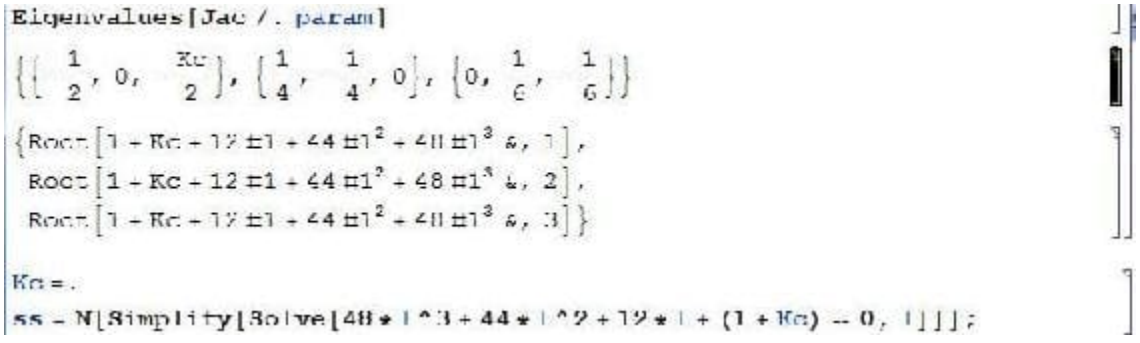
The Mathematica Help Browser is a very useful tool for understanding the syntax. For example, when using the **ListPlot[ ]** function you can search how to label the axes and how to make the points larger (more visible) on the plot. The following examples are used to illustrate the use of root locus plots. Click the following link if you would like to access the actual Mathematica file: Media:RLPFinalaa.nb

Second Plot Method Using Arrays

The goal here is to be able to understand exactly what is happening with these Mathematica inputs in order to get a better grasp on the outputs. Basically, once you have found the Eigenvalues of the system, which examples can be found for in other areas of the wiki, we then want to solve explicitly for them. After we solve for the Eigenvalue in terms of the variables, we can create an array by varying this value and solving for the real and imaginary terms of the Eigenvalues.

Eigenvalue Outputs in Mathematica

The Eigenvalues can either be returned as a function of 'l', which is the designation of the Eigenvalue or simply as a complex function in terms of variables. When returned as a function of 'l' there will be a polynomial input that has to be simplified to find 'l' as a function of the variable. In this case: Kc.



This example is taken directly from lecture 17 and its supplementary Mathematica file. What is happening here is that the Jacobian was inputted, and there were multiple eigenvalue arrays that were found. The last array has three complex polynomial functions that equal to 0. The last line is simplifying them, solving for 'l'. This is putting the simplified Eigenvalues into the array

'ss'.

Another output of Eigenvalues is simply in the form of complex numbers themselves. Rather than worry about the simplification steps, these Eigenvalues can be directly inputted as a complex number into an array, again termed 'ss':

$$ss = \left\{\left\{l \to \left(-5 - 10 * Kc - \sqrt{150 - 1590 * Kc + 1240 * Kc}\right)\right\},\ \left\{l \to \left(-5 - 10 * Kc + \sqrt{150 - 1590 * Kc + 1240 * Kc}\right)\right\}\right\}$$

This is a variation of the sample in lecture 17.

Plot Array Creation

Now that we have seen the two ways of Eigenvalue array creation, we can use these to create another array that will be the points for the plot. Here we will be substituting in values of the variable Kc to find different values of the Eigenvalues, which will give us information on stability changes as Kc changes.

```
Kc = .
ss = N[Simplify[Solve[18 * l^3 + 44 * l^2 + 12 * l + (1 + Kc) == 0, l]]];

a = {};
Kc = 0;
a = Append[a, {Re[l /. ss[[2]]], Im[l /. ss[[2]]]}];
a = Append[a, {Re[l /. ss[[1]]], Im[l /. ss[[1]]]}];
a = Append[a, {Re[l /. ss[[3]]], Im[l /. ss[[3]]]}];
Kc = 0.5;
a = Append[a, {Re[l /. ss[[2]]], Im[l /. ss[[2]]]}];
a = Append[a, {Re[l /. ss[[1]]], Im[l /. ss[[1]]]}];
a = Append[a, {Re[l /. ss[[3]]], Im[l /. ss[[3]]]}];
```

In the example above, 'ss' is in the form :

ParseError: EOF expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

Each solution in this array is in the form l -> a+bi. For example, ss refers to solution2, which is the object in the second position in ss. Re[l/.ss] applies solution2 to the variable l and then returns only the real part. Im[l/.ss] applies solution2 to the variable l and then returns only the imaginary part. The effect of the Append function above is that the ordered pair {a,b} (from a solution of the form a+bi) is added to the array called 'a'.

However, not all Eigenvalue arrays will have the same number of Eigenvalues. For example, as mentioned in the above paragraph, 'ss' is in the form :

ParseError: EOF expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

There could be 2, or 3, or more separate 'solution' values in the array. Depending on the number of the 'solution' values, there will be a different format for your 'append' function in Mathematica. For example, in the image below, there are only two eigenvalues in our arrat that we want to solve for. Thus, there should only be two rows of the addition to the array for the this method:

```
a = {};
F = 0;
a = Append[a, {Re[l /. ss[[2]]], Im[l /. ss[[2]]]}];
a = Append[a, {Re[l /. ss[[1]]], Im[l /. ss[[1]]]}];
```

These different methods both output arrays of numbers that act essentially as 'x' and 'y' values, or in this case Real and Imaginary, for the plot. The array 'a' is plotted as such:

ListPlot[a, PlotStyle -> PointSize[0.02]]

This is inputted directly to Mathematica, which creates the Root Locus Plot.

Differential Equation Example of Root Locus Plots in Mathematica

An example problem pulled from the Fall 2008 second exam:

"You have just been put in charge of designing a large-scale bioreactor for the production of a blockbuster protein based drug. After some research, you have come up with the following model to describe the reactor system:"

$$\frac{dX}{dt} = -2X + \frac{10XS}{S+2}$$

$$\frac{dS}{dt} = 4Fin - S - \frac{3XS}{S+2}$$

$$\frac{dP}{dt} = 9X - 2P$$

where X is the cell concentration, S is the nutrient concentration, and P is the protein product. Fin is the flow rate of nutrient into the system.

In this system there are two fixed points found by setting all the differential equations equal to zero and solving in Mathematica.

```
eqn = {0 == -2 * X + (10 * X * S) / (S + 2), 0 == 4 * Fin - S - (3 * X * S) / (S + 2), 0 == 9 * X - 2 * P};
eqns = {-2 * X + (10 * X * S) / (S + 2), 4 * Fin - S - (3 * X * S) / (S + 2), 9 * X - 2 * P};
Solve[eqn, {X, S, P}]
```

Out[4]= $\left\{\{P \to 0, X \to 0, S \to 4\ Fin\}, \left\{P \to \frac{15}{4}(-1 + 8\ Fin), X \to \frac{5}{6}(-1 + 8\ Fin), S \to \frac{1}{2}\right\}\right\}$

The two fixed points depend on the value of Fin and with that varying values of Fin will change eventual eigenvalues. Since this problem will eventually want root locus plots for both fixed points it is easier to save both points as parameters to be applied to general equations later. Overall it will help save the need for redundant code. The entire mathematica file being referenced in this example is available here: Media:Exam2Prob1.nb

From this point a Jacobian matrix should be created to linearize the system around the fixed points (as applied to the Jacobian). Once created the stored fixed point values for X, S, and P can be applied. Once applied, the eigenvalues for each fixed point can be solved for. These will still contain a variable 'Fin' since Fin has not been defined yet.

In[192]:=
```
Clear[Fin]
eqn = {0 == -2 * X + (10 * X * S) / (S + 2), 0 == 4 * Fin - S - (3 * X * S) / (S + 2), 0 == 9 * X - 2 * P};
eqns = {-2 * X + (10 * X * S) / (S + 2), 4 * Fin - S - (3 * X * S) / (S - 2), 9 * X - 2 * P};
Solve[eqn, {X, S, P}]
par1 = {P → 0, X → 0, S → 4 * Fin};
par2 = {P → 15/4 (-1 + 8 Fin), X → 5/6 (-1 + 8 Fin), S → 1/2};
Jac = {{D[eqns[[1]], X], D[eqns[[1]], S], D[eqns[[1]], P]},
   {D[eqns[[2]], X], D[eqns[[2]], S], D[eqns[[2]], P]},
   {D[eqns[[3]], X], D[eqns[[3]], S], D[eqns[[3]], P]}};
a = Jac /. par1;
b = Jac /. par2;
MatrixForm[a]
MatrixForm[b]
c = N[Eigenvalues[a]];
d = N[Eigenvalues[b]];
```

Out[201]//MatrixForm=

$$\begin{pmatrix} -2 + \frac{40\ Fin}{2+4\ Fin} & 0 & 0 \\ -\frac{12\ Fin}{2+4\ Fin} & -1 & 0 \\ 9 & 0 & -2 \end{pmatrix}$$

Out[202]//MatrixForm=

$$\begin{pmatrix} 0 & \frac{8}{3}(-1+8\ Fin) & 0 \\ -\frac{3}{5} & -8\ Fin + \frac{1}{5}(-1+8\ Fin) & 0 \\ 9 & 0 & -2 \end{pmatrix}$$

From here the Table[] function can be used to tabulate eigenvalues for various Fin values for visual sake of inspection. This is not necessary since the eigenvalue results have already been solved for in terms of Fin. To create the root locus plots, values of Fin need to be defined and applied to the eigenvalue equations with the results stored. That step is next. In this table step it is for the user's sake to see with increasing Fin values what the eigenvalues look like. The tabulated eigenvalues are stored as a matrix

```
a = Jac /. par1;
b = Jac /. par2;
MatrixForm[a]
MatrixForm[b]
c = N[Eigenvalues[a]];
d = N[Eigenvalues[b]];
e = Table[c, {Fin, 0, 2, 0.1}];
MatrixForm[e]
f = Table[d, {Fin, 0, 2, 0.1}];
MatrixForm[f]
```

and displayed as such with the function 'MatrixForm[

Out[206]//MatrixForm=

-2.	-1.	-2.
-2.	-1.	-0.333333
-2.	-1.	0.857143
-2.	-1.	1.75
-2.	-1.	2.44444
-2.	-1.	3.
-2.	-1.	3.45455
-2.	-1.	3.83333
-2.	-1.	4.15385
-2.	-1.	4.42857
-2.	-1.	4.66667
-2.	-1.	4.875
-2.	-1.	5.05882
-2.	-1.	5.22222
-2.	-1.	5.36842
-2.	-1.	5.5
-2.	-1.	5.61905
-2.	-1.	5.72727
-2.	-1.	5.82609
-2.	-1.	5.91667
-2.	-1.	6.

Out[208]//MatrixForm=

-2.	-1.36886	1.16886
-2.	-1.12456	0.284557
-2.	-0.74 - 0.642184 i	-0.74 + 0.642184 i
-2.	-1.06 - 1.0566 i	-1.06 + 1.0566 i
-2.	-1.38 - 1.27106 i	-1.38 + 1.27106 i
-2.	-1.7 - 1.38203 i	-1.7 + 1.38203 i
-2.	-2.02 - 1.41407 i	-2.02 + 1.41407 i
-2.	-2.34 - 1.37273 i	-2.34 + 1.37273 i
-2.	-2.66 - 1.25076 i	-2.66 + 1.25076 i
-2.	-2.98 - 1.01961 i	-2.98 + 1.01961 i
-2.	-3.3 - 0.556776 i	-3.3 + 0.556776 i
-2.	-4.41019	-2.82981
-2.	-5.26801	-2.61199
-2.	-6.02284	-2.49716
-2.	-6.73787	-2.42213
-2.	-7.4318	-2.3682
-2.	-8.11282	-2.32718
-2.	-8.78524	-2.29476
-2.	-9.4516	-2.2684
-2.	-10.1135	-2.2465
-2.	-10.772	-2.228

These tables make it easy to see with increasing Fin (goes from 0 to 2 down the table) what the changes are in the stability of each fixed point. The first table is the first fixed point, the second table the second FP.

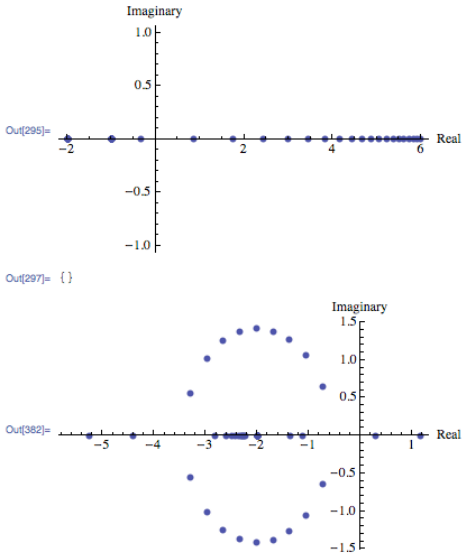
Now, to make the root locus plots, a really long string of values needs to be created in order for the real and imaginary roots to be plotted. This is a long string of code that is basically brute force. The overall action occurring here is setting Fin to a value, evaluating that Fin through the eigenvalues, and saving the real portion separate from the imaginary portion by different columns. This is repeated for as many Fin values as you desire. It is extremely repetitive, but works. After evaluating as many Fin values as you desired you can display the two columns through the ListPlot[] function. Example snippets of the code look like such:

```
Fin=.
rootc = {}
Fin = 0;
rootc = Append[rootc, {Re[c[[2]]], Im[c[[2]]]}];
rootc = Append[rootc, {Re[c[[1]]], Im[c[[1]]]}];
rootc = Append[rootc, {Re[c[[3]]], Im[c[[3]]]}];
Fin = 0.1;
rootc = Append[rootc, {Re[c[[2]]], Im[c[[2]]]}];
rootc = Append[rootc, {Re[c[[1]]], Im[c[[1]]]}];
rootc = Append[rootc, {Re[c[[3]]], Im[c[[3]]]}];
Fin = 0.2;
rootc = Append[rootc, {Re[c[[2]]], Im[c[[2]]]}];
rootc = Append[rootc, {Re[c[[1]]], Im[c[[1]]]}];
rootc = Append[rootc, {Re[c[[3]]], Im[c[[3]]]}];
Fin = 0.3;
rootc = Append[rootc, {Re[c[[2]]], Im[c[[2]]]}];
rootc = Append[rootc, {Re[c[[1]]], Im[c[[1]]]}];
rootc = Append[rootc, {Re[c[[3]]], Im[c[[3]]]}];
Fin = 0.4;
rootc = Append[rootc, {Re[c[[2]]], Im[c[[2]]]}];
rootc = Append[rootc, {Re[c[[1]]], Im[c[[1]]]}];
rootc = Append[rootc, {Re[c[[3]]], Im[c[[3]]]}];
Fin = 0.5;
rootc = Append[rootc, {Re[c[[2]]], Im[c[[2]]]}];
rootc = Append[rootc, {Re[c[[1]]], Im[c[[1]]]}];
rootc = Append[rootc, {Re[c[[3]]], Im[c[[3]]]}];
```

And the ListPlot[] function can look like this:

```
ListPlot[rootc, PlotStyle → PointSize[0.02],
 TextStyle → {FontFamily → "Times", FontSize → 12}, AxesLabel → {Real, Imaginary}]
```

And would make graphs like these (Fixed Point 1 on top and FP2 on bottom in this example):



Looking back at the tables of eigenvalues the plots can be interpreted. For FP1 the values start real negative and increase to real positive. There are no imaginary values. For FP2, increasing Fin values go from real positive to negative with imaginary values occurring in the transition from low to high Fin values.

Alternative Mathematica Method

Some Mathematica 6 programs have an add-on or supplementary application known as **ANALOG INSYDES** which contains special options to create Root Locus Plots with ease. The Function is called *RootLocusPlot*. Enter the function that is to be analyzed, into Mathematica with the following format:

RootLocusPlot[*tfunc*,{*k*,*k*_0 ,*k*_1}]

tfunc is the transfer function in the frequency variable s and one real parameter *k*. k_0 and k_1 are the range for the real parameter *k* that is to be varied int he Root Locus Plot. Follow an example below.

Example: Equation to make a Root Locus Plot with:

$$H(s) = \draca + 2s + s210 + 3as + 4s^2 + s^3$$

To make a Root Locus Plot follow the example code below:

H4[s_, a_] := (a + 2*s + s^2)/(10 + 3*a*s + 4*s^2 + s^3)

RootLocusPlot[H4[s, a], {a, 3, 5}]

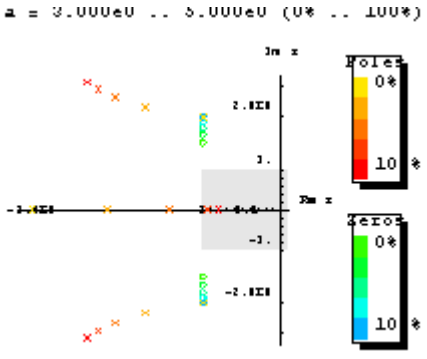


Image: Wolfram.com

Other Mathematica forms for using RootLocusPlot[]:

RootLocusPlot[*func*] This form displays the a pole/zero diagram of a function i.e. *func* without parameters *k* and *k_1*.

RootLocusPlot[*rootloc*] This form displays a root locus calculated with function RootLocusByQZ[].

Refer to Mathematica link for further discuss about RootLocusByQZ

Many options for the plot can be varied including the number of points, plot style, and plot color. All of these options and more can be viewed online with a Mathematica Tutorial

Creating Root Locus Plots with Matlab

Root Locus Generation in Matlab

Three matlab files have been given to obtain the root locus plot and poles of the root locus plot at specified Kc values for a specific transfer function with relative ease. These three files are polesfunction.m, transferfunction.m, and Locusplotpoles.m .[1]. The transferfunction file is where the specific transfer function should be input. As written the only line needing variation is line 4 (Gs). The polesfunction file finds the poles of the given transfer function using built in matlab utilities and needs no variation, even after changing the transfer function. The Locusplotpoles file is what is called in matlab to generate the root locus plot as well as the value of the poles and integrates the two previous files. To use this file you type in the values of the specific Kc you require the pole values for as well as a single value of Kc for which you would like to see the root locus plot. An example of the inputs to and outputs of this file are given below.

>> Locusplotpoles([0 50 100 200],0)

ans =

1.0e+002 *

Columns 1 through 3

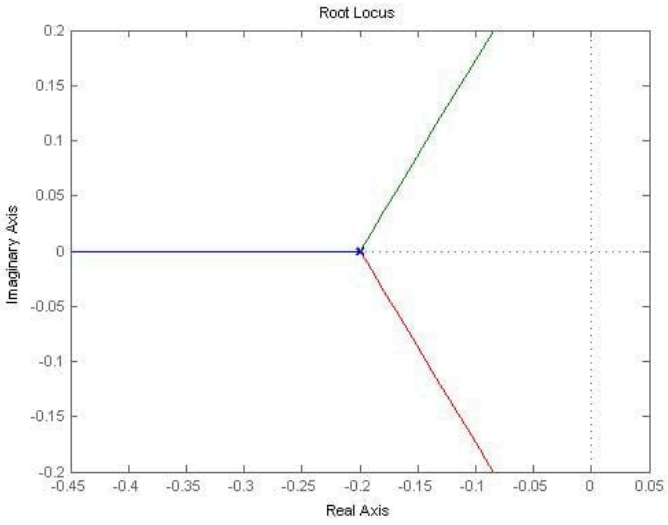
0 -0.0020 -0.0020 + 0.0000i 0.5000 -0.0045 -0.0008 + 0.0022i 1.0000 -0.0051 -0.0004 + 0.0027i 2.0000 -0.0060 -0.0000 + 0.0034i

Columns 4 through 6

-0.0020 - 0.0000i 0 0 -0.0008 - 0.0022i -0.0020 -0.0020 + 0.0000i -0.0004 - 0.0027i -0.0020 + 0.0000i -0.0020 - 0.0000i -0.0000 - 0.0034i -0.0020 -0.0020 + 0.0000i

Column 7

0 -0.0020 - 0.0000i -0.0020 -0.0020 - 0.0000i



The values inside the brackets are the specific Kc values for which poles are desired. The last value is the Kc which the root locus plot will be plotted for. Notice that the value of Kc for the root locus plot to be generated in this example is zero. This is so that the entire range of Kc can be examined using the interactive graph produced by matlab. If other specific Kc plots would like to be observed this value can be changed to generate these as well. The column lines give the values of Kc input and then value of the poles for each Kc. Note that the actual values are the values matlab gives multiplied by 100.

Now that a basic understanding of root locus plots has been obtained here are a few websites which go further in depth into the underlying principles and generation of these plots. www.facstaff.bucknell.edu/mastascu/eControlHTML/RootLocus/RLocus1A.html

This site gives animated examples which progress the value of the control variable through a root locus plot to better explain the fundamental function of these plots. It would be beneficial to go through a few of these examples to observe patterns of fluctuation of the control variable. The parameters for each animation are given to the left of the frame. www.engin.umich.edu/group/ctm/rlocus/rlocus.html

This page gives examples of how you generate a root locus plot of a given transfer function using matlab. While you go through the content on this page take special notice of the fact that none of the given transfer functions have an integrated Kc variable for which to generate a locus plot. To see how these plots can are generated when a value of Kc is specified see the code of Locusplotpoles.m. For clarification on the application of root locus plots see the given examples.

Creating Root Locus plots with Excel and PPLANE

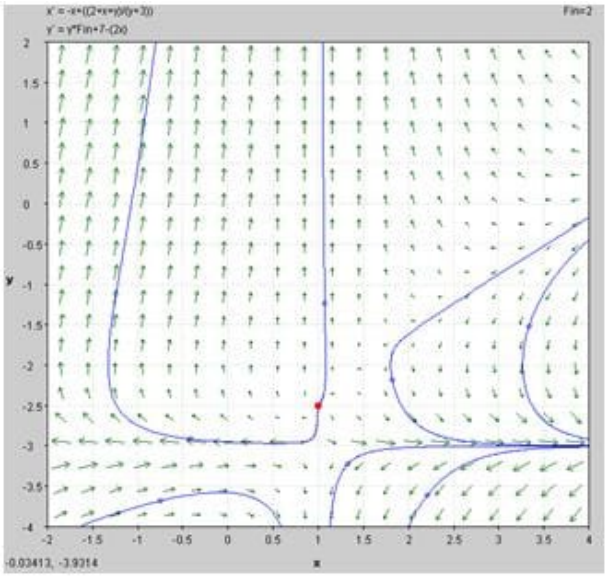
Though this tactic may remain slightly more time consuming and work involved, there is a method in computing the Root Locus plot using PPLANE and Excel. This particular model really aids in better understanding of how to read and acknowledge a Root Locus Plot, and is also useful if alternative options for making a Root Locus Plot are not currently available.

This method involves using PPLANE in order to find the eigenvalues for each equilibrium value, and then plotting these points with Excel software. In order to do this, suppose that you have the following differential equations that represent a reactor:

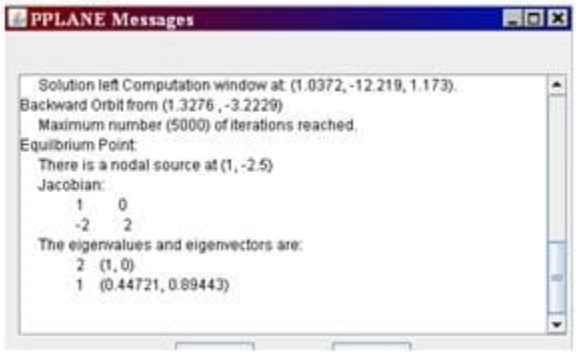
$$\frac{dX}{dt} = -X + \frac{2 + X + Y}{Y + 3}$$

$$\frac{dY}{dt} = Y \times Fin + 7 - (2X)$$

Graphing the following differential equations provides something similar to the following chart (a few lines were drawn, just to better show the graphs flow):



Using the ability to “Find an Equilibrium Point” in PPLANE (if you need help with PPLANE, click PhasePlaneAnalysis), one can select a particular equilibrium point on the field, and be able to provide the following results:



This window “pops up” in the higher left corner of the screen, when the equilibrium point is found. This window provides the eigenvalues for this equilibrium point.

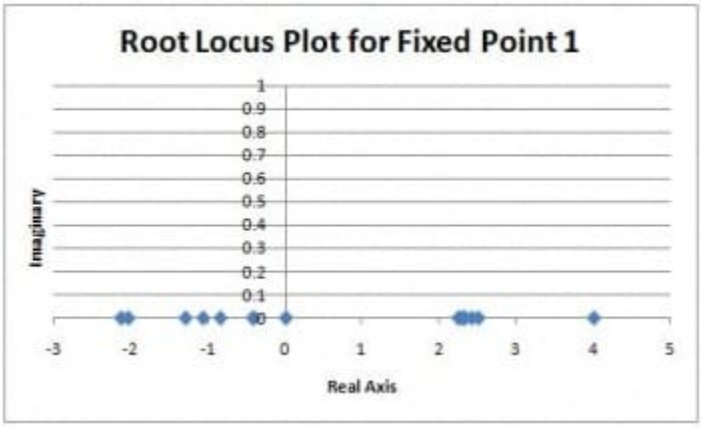
Depending what you are varying (in this case, Fin), all you would have to do is enter various values of this Fin into PPLANE, and plot the new set of differential formulas. After this is done, just find the same equilibrium point, and record the new given eigenvalues in excel. These values should be entered with an x-component (in one column) and a y-component (in another column) into excel, with the following rules:

- Assuming the example 3+3i, 3-3i were given as eigenvalues.
- For any given eigenvalue, the x-value of a particular eigenvalue is the real part of that number (ie: the “3” part).
- For any given point, the y-value of a particular eigenvalue is the imaginary part of that number (ie: the “3i” part). If there is an imaginary component, ignore the “i,” though the number should be recorded in the y-component. If there is no imaginary number, the y component is zero.
- For both cases, a positive is a positive, and a negative is a negative. Remember: imaginary numbers have 2 parts, a positive and negative “i” value.

These points are simply recorded according to these rules, then the points are plotted against each other, with x components on the x-axis, and the y in the y-axis. Here is an example of an excel diagram, taken from the proposed diagram:

X-axis	Y-axis (imaginary)
2.3028	0
-1.3028	0
2.2415	0
-2.1415	0
2.2471	0
-2.0471	0
4	0
-0.42857	0
2.3058	0
-0.85078	0
2.3252	0
-1.0752	0
2.4142	0
-0.41421	0
2.5	0
0	0

And, the excel data is then plotted, x against y, and is shown below:



As you can see, this particular model did not have any imaginary data. The root locus plot has been formed, and for all accounts of Fin, the eigenvalues are along the x-axis for the root locus plot.

Practical Application

In the past, it was necessary for engineers to master the techniques required to efficiently construct root locus diagrams. In today's engineering world, this is not the case for one of two reasons. In many instances, root locus diagrams are not used industrially because they require models of the system which are generally not available. If a model is available to develop a root locus diagram, there are computer applications that can develop the diagrams much faster than a person. Thus, energy and effort should be placed on understanding and interpreting a root locus diagram and understanding the general rules of stability for a root of the characteristic equation.

Example 10.6.1

Given the following table of roots to the characteristic equation for a system:

K_c	r_1	r_2	r_3
0	-0.45	-0.60	-0.78
0.33	-0.51	-0.51	-0.83
0.65	-0.50 + 0.40i	-0.50 − 0.40i	-0.90
1.33	-0.41 + 0.50i	-0.41 - 0.50i	-0.95
1.67	-0.35 + 0.57i	-0.35 - 0.57i	-1.01
2.00	-0.10 + 0.65i	-0.10 - 0.65i	-1.32
2.33	0.02 + 0.71i	0.02 - 0.71i	-1.50
2.67	0.23 + 0.76i	0.23 - 0.76i	-1.69
3.00	0.31 + 0.78i	0.31 - 0.78i	-1.85

1. Construct a root locus plot (Making the plot by hand initially may give some a better understanding of what is going on as you will see the changes in the plot as you add each set of roots).
2. Determine the ranges of Kc that would lead to the following system responses: no oscillations, dampened oscillations, stable oscillations & unstable oscillations. Also, sketch what the output response would look like for a given step change in input for each range of Kc.
3. How would the roots to the characteristic equation be different if this were PI control?

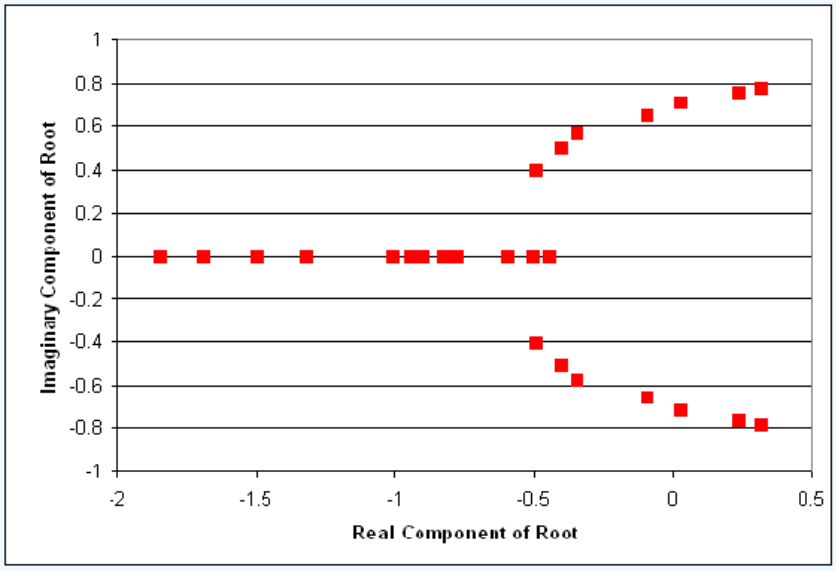
Solution

1) Root locus plot:

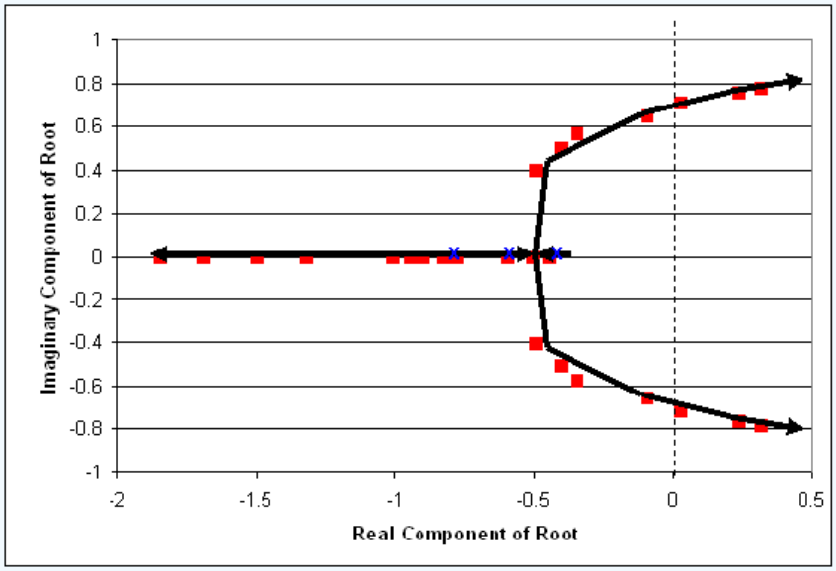
First, we break the roots into their imaginary and real components:

Kc	r1 real	r1 imaginary	r2 real	r2 imaginary	r3 real	r3 imaginary
0	-0.45	0	-0.6	0	-0.78	0
0.33	-0.51	0	-0.51	0	-0.83	0
0.66	-0.5	0.4	-0.5	-0.4	-0.9	0
1.33	-0.41	0.5	-0.41	-0.5	-0.95	0
1.67	-0.35	0.57	-0.35	-0.57	-1.01	0
2	-0.1	0.65	-0.1	-0.65	-1.32	0
2.33	0.02	0.71	0.02	-0.71	-1.5	0
2.67	0.23	0.76	0.23	-0.76	-1.69	0
3	0.31	0.78	0.31	-0.78	-1.85	0

Then, we plot each point on a complex coordinate system (x-axis = real componebt, y-axis = imaginary component).



Also, it is useful to add lines tracing the path of increasing Kc. These lines are added below. Also, the points where Kc= zero are generally labeled with an x. This has also been done in the figure below.



2) Ranges of Kc with explanations:

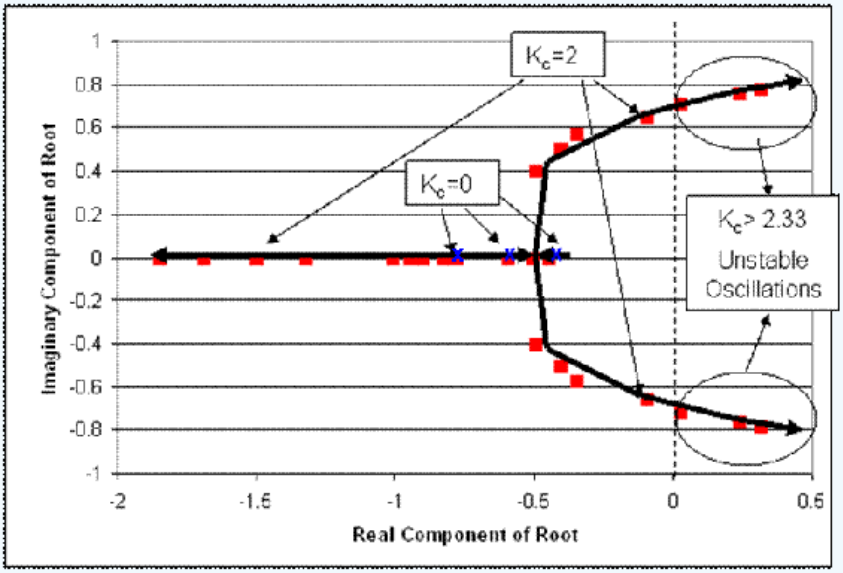
We know that the range where there will be no oscillatory response is when Kc provides all real roots. This occurs when Kc is between 0 and 0.66. (Sketch will look like figure a in the "interpreting a root locus diagram" section above).

Also, Kc yields dampened oscillations when the two of the roots of the characteristic equation are complex with negative real components. This occurs when Kc is greater than 0.66 but less than 2.33. (Sketch will look like figure b in the "interpreting a root locus diagram" section above).

Unstable oscillations are achieved when two of the roots yield complex numbers with positive real components. This occurs when Kc is greater than 2.33. Graphically, we see that when the points move past the dashed line, we will have unstable oscillations. This is represented in the figure below. (Sketch will look like figure c in the "interpreting a root locus diagram" section above).

Based on the data provided, our best assumption is that at a Kc just below 2.33, we will be able to achieve stable oscillations. When Kc = 2.33 the real component of the complex root is 0.02 (slightly greater than zero). Ideally, we will achieve stable oscillations when the real component of the complex number is zero. (Oscillations should be consistent in magnitude).

Lastly, it is a good idea to label some values of Kc on the root locus plot. This will help orient the reader when they are trying to determine which point correspond to different values of Kc.



3) If this were PI control, each root to the characteristic equation would be accompanied by not only a Kc value but also a value for integral control parameter, KI.

Example 10.6.2

You've been recently hired at MichChem Controls. In your first week at the company, a virus spreads through the computer network, effectively disabling most of the computational programs. Fortunately, the virus was halted before it was able to disable Mathematica. Shortly after the chaos from the virus subsides, your boss needs to know how to use Mathematica to create a root locus plot for several new Proportional only level control systems that are being tested. Being the highly skilled engineer that you are, he doesn't want you to waste time developing the root locus plots, rather, he wants you to write out example code, with a brief description that someone else could use to develop the root locus plot.

The general characteristic equation for the new level control systems is:

$$f(s) = Ax^3 + Bx^2 - Cx + D + EK_c$$

Please, provide your boss with an example of the exact script that would need to be entered into Mathematica in order to develop the root locus plot. You should also brief descriptions explaining what the code does and what steps would need to be repeated to develop the plot.

Solution

First, the user should develop a table of roots for the characteristic equation. This can be done using the solve function in Mathematica. The code would appear like this for Kc = 0.

(note: the In[#] appears for each new line of code. This will not be typed in by the user.)

```
In[1]:= Solve[Ax^3 +Bx^2-Cx+D+E*(0)==0,x]
```

This should be repeated for Kc = 0.1, 0.15, 0.225, 0.5, 0.75, 1.5, etc. (until the roots of the equation begin to yield two complex solutions whose real components are positive). After a sufficient number of roots have been obtained for the characteristic equation at different values of Kc, these will need to be tabulated. To develop a table of data in Mathematica input the following code:

In[2]:= A =

ParseError: EOF expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Control
```

This tabulates the real components and the imaginary components at each value of Kc. Once all of the values are tabulated, the root locus plot can be developed using the ListPlot function.

In[3]:= ListPlot[A, PlotStyle -> PointSize[0.02]

This will plot all of the tabulated data in In[2] which is the root locus plot.

Exercise 10.6.1

A root locus plot is used to:

a. Identify the solutions to the characteristic transfer function.
b. Show the response time of the feedback system to a disturbance.
c. Graphically illustrate the effect of Kc on the dynamic characteristics of the feedback system.
d. Choose the optimal Kc for the feedback system.

Answer available at the end of this section.

Exercise 10.6.2

When a root is complex with positive real parts, the oscillations are:

a. Decreasing
b. Increasing
c. Unstable
d. A & C
e. B & C

Answer available at the end of this section.

References

- Luyben, Michael L.; Luyben, William L. "essential of Process Control". McGraw-Hill:New York, 1997.
- Marlin, Thomas E. "Process Control: Designing Processes and Control Systems for Dynamic Performace 2nd Edition". McCGraw-Hill: New York, 2000.
- Ogunnaike, Babatunde A.; Ray, W. Harmon. *Process Dynamics, Modeling, and Control*. New York Oxford: Oxford UP, 1994. 488-507.
- Riggs, James B.; Karim, M. Nazmul. *Chemical and Bio-Process Control*. 3rd ed. Ferret. 279-295.
- Luyben, Michael L.; Luyben, William L.; Tyreus, Bjorn D. "Plantwide Process Control". Mcgraw-Hill:New York, 1999.
- Seborg, Dale E.; Edgar, Thomas F.; Mellichamp, Duncan A. "Process Dynamics and Control". John Wiley and Sons, Inc. 2004.
- www.engin.umich.edu/group/ctm/rlocus/rlocus.html as viewed 10/23/2007
- www.facstaff.bucknell.edu/mastascu/eControlHTML/RootLocus/RLocus1A.html as viewed 10/23/2007

Contributors and Attributions

- Authors: Halley Crast, Andrew Laskowski, Maurice Telesford, Emily Yatch
- Stewards: Victoria Cardine, Anthony Campell, David Hines, Stephen Kerns

Answers

Exercise 10.6.1

c

Exercise 10.6.2

e

This page titled 10.6: Root Locus Plots - Effect of Tuning is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

10.7: Routh Stability - Ranges of Parameter Values that are Stable

Introduction

The stability of a process control system is extremely important to the overall control process. System stability serves as a key safety issue in most engineering processes. If a control system becomes unstable, it can lead to unsafe conditions. For example, instability in reaction processes or reactors can lead to runaway reactions, resulting in negative economic and environmental consequences.

The absolute stability of a control process can be defined by its response to an external disturbance to the system. The system may be considered stable if it exists at a consistent state or setpoint and returns to this state immediately after a system disturbance. In order to determine the stability of a system, one often must determine the eigenvalues of the matrix representing the system's governing set of differential equations. Unfortunately, sometimes the characteristic equation of the matrix (the polynomial representing its eigenvalues) can be difficult to solve; it may be too large or contain unknown variables. In this situation, a method developed by British mathematician Edward Routh can yield the desired stability information without explicitly solving the equation.

Recall that in order to determine system stability one need only know the signs of the real components of the eigenvalues. Because of this, a method that can reveal the signs without actual computation of the eigenvalues will often be adequate to determine system stability.

To quickly review, negative real eigenvalue components cause a system to return to a steady state point (where all partial derivatives equal zero) when it experiences disturbances. Positive real components cause the system to move away from a stable point, and a zero real component indicates the system will not adjust after a disturbance. Imaginary components simply indicate oscillation with a general trend in accordance with the real part. Using the method of Routh stability, one can determine the number of each type of root and thus see whether or not a system is stable. When unknown variables exist in the equation, Routh stability can reveal the boundaries on these variables that keep the roots negative.

The Routh Array

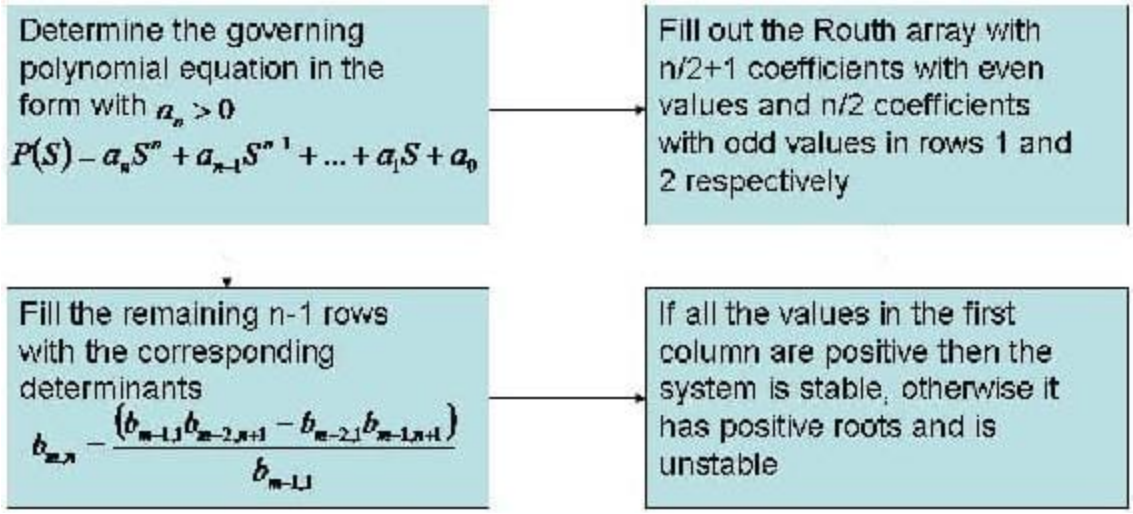
The Routh array is a shortcut to determine the stability of the system. The number of positive (unstable) roots can be determined without factoring out any complex polynomial.

Generating the Array

The system in question must have a characteristic equation of a polynomial nature. as shown below:

$$P(S) = a_n S^n + a_{n-1} S^{n-1} + \cdots + a_1 S + a_0$$

In order to examine the roots, set P(S)=0, which will allow you to tell how many roots are in the left-hand plane, right hand plane, and on the j-omega axis. If the system involves trigonometric functions it needs to be fit to a polynomial via a Taylor series expansion. One necessary condition for stability is that a_n 0 \qquad " src="/@api/deki/files/18713/image-273.png">. (If $a_n < 0$, all coefficients may be multiplied by -1 before checking). The other condition is that all values in column 1 of the Routh array must be positive for the system to be stable.



This flow diagram shows the generation of a Routh array for an idealized case with m,n representing the location in the matrix.

The coefficients of the polynomial are placed into an array as seen below. The number of rows is one more than the order of the equation. The number of sign changes in the first column indicate the number of positive roots for the equation.

Row 1	a_n	a_{n-2}	a_{n-4}	...
Row 2	a_{n-1}	a_{n-3}	a_{n-5}	...
Row 3	b_1	b_2	b_3	...
Row 4	c_1	c_2	c_3	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Row 5	p_1	p_2		
Row 6	q_1			
Row 7	v_1			

In the array, the variables b1,b2,c1,c2,etc. are determined by calculating a determinant using elements from the previous two rows as shown below:

$$b_1 = \frac{a_{n-1}a_{n-2} - a_n a_{n-3}}{a_{n-1}}, b_2 = \frac{a_{n-1}a_{n-4} - a_n a_{n-5}}{a_{n-1}}, b_3 = \frac{a_{n-1}a_{n-6} - a_n a_{n-7}}{a_{n-1}}, \cdots$$

$$c_1 = \frac{b_1 a_{n-3} - a_{n-1} b_2}{b_1} \; ; \; c_2 = \frac{b_1 a_{n-5} - a_{n-1} b_3}{b_1} \; , c_3 = \frac{b_1 a_{n-7} - a_{n-1} b_4}{b_1} \; , \ldots$$

The general expression for any element *x* after the first two rows with index (m,n) is as follows:

For $A = \begin{bmatrix} x_{m-2,1} & x_{m-2,n+1} \\ x_{m-1,1} & x_{m-1,n+1} \end{bmatrix}$

$$x_{m,n} = \frac{-det(A)}{x_{m-1,1}}$$

Note, that if the Routh array starts with a zero, it may still be solved (assuming that all the other values in the row are not zero), by replacing the zero with a constant, and letting that constant equal a very small positive number. Subsequent rows within that column that have this constant will be calculated based on the constant choosen.

Once the array is complete, apply the following theorems to determine stability:

1. If all of the values in the fist column of the Routh array are >0, then P(S) has all negative real roots and the system is stable.
2. If some of the values in the first column of the Routh array are <0, then the number of times the sign changes down the first column will = the number of positive real roots in the $P(S)$ equation.
3. If there is 1 pair of roots on the imaginary axis, both the same distance from the origin (meaning equidistant), then check to see if all the other roots are in the left hand plane. If so, then the imaginary roots location may be found using $AS^2 + B = 0$, where A and B are the elements in the Routh array for the 2nd to last row.

To clarify even further, an example with real numbers is analyzed.

Example Array

The following polynomial was generated from a sample system.

$$P(S) = 5S^3 - 10S^2 + 7S + 20$$

The preceding polynomial must be investigated in order to determine the stability of the system. This is done by generating a Routh array in the manner described above. The array as a result of this polynomial is,

Row 1	$5 \quad 7$
Row 2	$-10 \quad 20$
Row 3	17
Row 4	20

In the array shown above, the value found in the third row is calculated as follows. $b_1 = \frac{-10 * 7 - 5 * 20}{-10}$ The array can now be analyzed. When looking down the first column, it can be seen that 5 is positive in magnitude, then the sign changes in the -10 entry, and the sign changes a second time to positive 17. This counts as two changes in sign, which corresponds to two positive roots, making the system unstable.

Finding Stable Control Parameter Values

Often, for a unit operation, a PID parameter such as controller gain (K_c), the integral time constant (T_i), or the derivative time constant (T_d) creates an additional variable in the characteristic equation. This can be carried through the computations of the Routh array to indicate which values of the variable will provide stability to the system through by preventing positive roots from occuring in the equation. For example, if a controller output is governed by the function:

$$10s^3 + 5s^2 + 8s + (T_d + 2)$$

The stable values of T_d can be found via a Routh array:

Row 1	10 8
Row 2	5 $(T_d + 2)$
Row 3	$4 - 2T_d$
Row 4	$(T_d + 2)$

We reveal $-2 < T_d < 2$ in order to keep the first column elements positive, so this is the stable range of values for this parameter.
If multiple parameters were in the equation, they would simply be solved for as a group, yielding constraints along the lines of "T_i + K_c > 2" etc, so any value chosen for one parameter would give a different stable range for the other.

Special Cases

There are a few special cases that one should be aware of when using the Routh Test. These variances can arise during stability analysis of different control systems. When a special case is encountered, the traditonal Routh stability solution methods are altered as presented below.

One of the coefficients in the characteristic equation equals zero

If the power of the 0*S^n is ≥ 1, replace the zero with a quantity, ε, which would be positive and will approach zero. Then continue with your analysis as normal. Essentially this gives the limit of the roots as that coefficient approaches zero. (If the power = 0, see Case 3). For example,

Equation: $2S^3 - 24S + 32 = 0$

Working Equation: $2S^3 + \epsilon S^2 - 24S + 32 = 0$

Row 1	2 -24
Row 2	ϵ 32

Row 3	$frac-24\epsilon - 64\epsilon \qquad 0$
Row 4	$qquad32$

Since ε is positive we know that in the first column row 2 will be positive, row 4 will be positive, and row 3 will be negative. This means we will have a sign change from 2 to 3 and again from 3 to 4. Because of this, we know that two roots will have positive real components. If you actually factor out the equation you see that $(S-2)^2(2S+8)=0$, showing that we do have 2 positive roots. Both of these roots are equal to 2, so there is technically only one root, but in any case we know the system is unstable and must be redesigned.

One of the roots is zero

This case should be obvious simply from looking at the polynomial. The constant term will be missing, meaning the variable can be factored from every term. If you added an ε to the end as in case 1, the last row would be ε and falsely indicate another sign change. Carry out Routh analysis with the last zero in place.

Equation: $S^3 - S^2 - 2S = 0$

Row 1	$1 \quad -2$
Row 2	$: -1 \qquad 0$
Row 3	$: -2$
Row 4	0

As you can see in column one we have row 1 positive, row 2 and 3 negative, and row 4 zero. This is interpreted as one sign change, giving us one positive real root. Looking at this equation in factored form,

$$(S+1)(S-2)S=0$$

we can see that indeed we have only one positive root equal which equals 2. The zero in the last row indicates an additional unstable root of zero. Alternatively, you may find it easier to just factor out the variable and find the signs of the remaining eigenvalues. Just remember there is an extra root of zero.

Row Full of Zeros

When this happens you know you have either a pair of imaginary roots, or symmetric real roots. The row of zeros must be replaced. The following example illustrates this procedure.

Equation: $S^4 - 6S^3 + 10S^2 - 6S + 9 = 0$

Row 1	$1 \qquad 10 \qquad 9$

Row 2	$: -6 \qquad -6$
Row 3	$9 \qquad 9$
Row 4	$0 \qquad 0$

Row 4 contains all zeros. To determine its replacement values, we first write an auxiliary polynomial *A* determined by the entries in Row 3 above.

$$A(S) = 9S^2 + 9$$

Notice that the order decreases by 1 as we go down the table, but decreases by 2 as we go across.

We then take the derivative of this auxiliary polynomial.

$$A'(S) = 18S$$

The coefficients obtained after taking the derivative give us the values used to replace the zeros. From there, we can proceed the table calcuations normally. The new table is

Row 1	$1 \qquad 10 \qquad 9$
Row 2	$: -6 \qquad -6$
Row 3	$9 \qquad 9$
Row 4	$18 \qquad 0$
Row 5	9

In fact, the purely imaginary or symmetric real roots of the original polynomial are the same as the roots of the auxiliary polynomial. Thus, we can find these roots.

$$9S^2 + 9 = 0$$

$$S = \pm i$$

Because we have two sign changes, we know the other two roots of the original polynomial are positive.

In fact, after factoring this polynomial, we obtain

$$(S^2 + 1)(S - 3)^2 = 0$$

Therefore, the roots are $S = \pm i, 3$, where in this case, the root 3 has multiplicity 2.

Limitations

Routh arrays are useful for classifying a system as stable or unstable based on the signs of its eigenvalues, and do not require complex computation. However, simply determing the stability is not usually sufficient for the design of process control systems. It is important to develop the extent of stability as well as how close the system is to instability. Further stability analysis not accounted for in the Routh analysis technique include finding the degree of stability, the steady state performance of the control system, and the transient response of the system to disturbances.

More involved techniques, such as those discussed in Eigenvalues and Eigenvectors, must be used to further characterize the system stability (with the exception of system polynomials resulting in pure imaginary roots). Another limitation of the Routh method occurs when the polynomial in question becomes so large that Routh stability is too computationally time consuming (a personal judgment). For this situation another method, such as a root locus plotmust be used.

Note that for defining stability, we will always start out with a polynomial. This polynomial arises from finding the eigenvalues of the linearized model. Thus we will never encounter other functions, say exponenential functions or sin or cos functions in general for stability analysis in control theory.

Advantages Over Root Locus Plots

Routh stability evaluates the signs of the real parts of the roots of a polynomial without solving for the roots themselves. The system is stable if all real parts are negative. Therefore unlike root locus plots, the actual eigenvalues do not need to be calculated for a Routh stability analysis. Furthermore, sometimes the system has too many unknowns to easily construct and interpret a root locus plot (e.g. with two PID controllers there are the variables Kc1, Kc2, τi1, τi2, τd1, and τd2).

Example 10.7.1

Assume the following polynomial describes the eigenvalues of a linearized model of your process. For this polynomial, complete a Routh array and determine the system's stability?

$$P(x) = x^4 + 10x^3 + 35x^2 + 50x + 24$$

Answer
Since P(X) is a fourth-order polynomial, the Routh array contains five rows.

Row 1:	a_n	a_{n-2}	a_{n-4}
Row 2:	a_{n-1}	a_{n-3}	a_{n-5}
Row 3:	b_1	b_2	b_3
Row 4:	c_1	c_2	c_3
Row 5:	d_1	d_2	d_3

Rows 1 and 2 correspond to the coefficients of the polynomial terms:

Row 1:	1	35	24
Row 2:	10	50	

Rows 3, 4, and 5 contain the determinants using elements from the previous two rows.

Row 3	$: \quad b_1 = \dfrac{a_{n-1}a_{n-2} \quad a_n a_{n-3}}{a_{n-1}} \qquad b_2 = \dfrac{a_{n-1}a_{n-4} \quad a_n a_{n-5}}{a_{n-1}} \qquad b_3 = \dfrac{a_{n-1}a_{n-6} \quad a_n a_{n-7}}{a_{n-1}}$
Row 4	$: \quad c_1 = \dfrac{b_1 a_{n-3} - a_{n-1} b_2}{b_1} \qquad c_2 = \dfrac{b_1 a_{n-5} - a_{n-1} b_3}{b_1}$
Row 5	$: \quad d_1 = \dfrac{c_1 b_2 - b_1 c_2}{c_1}$

Therefore,

Row 3	$: \quad b_1 = \dfrac{10(35) - 1(50)}{10} = 30 \qquad b_2 = \dfrac{10(24) - 1(0)}{10} = 24 \qquad b_3 = \dfrac{10(0) - 1(0)}{10} = 0$
Row 4	$: \quad c_1 = \dfrac{30(50) - 10(24)}{30} = 42 \qquad c_2 = \dfrac{30(0) - 10(0)}{30} = 0$
Row 5	$: \quad d_1 = \dfrac{42(24) - 30(0)}{42} = 24$

The complete Routh array:

Row 1	$: \quad 1 \qquad 35 \qquad 24$
Row 2	$: \quad 10 \qquad 50$
Row 3	$: \quad 30 \qquad 24$
Row 4	$: \quad 42$
Row 5	$: \quad 24$

Since all the values in the first column are positive, the equation P(x) has all negative roots and the system is stable.

Example 10.7.1

Consider a system with the following characteristic equation:

$$20s^3 + 59s^2 + 46s + (4 + K_c) = 0$$

Using a P-only controller, find the range of controller gain that will yield a stable system.

Answer

Since the equation is a third-order polynomial, the Routh array has four rows. Routh Array:

Row 1	20 46
Row 2	59 $(4+K_c)$
Row 3	$46-\frac{20}{59}(4+K_c)$ 0
Row 4	$(4+K_c)$ 0

For the system to be stable, all the elements of the first column of the Routh array have to be positive.

The first column will contain only positive elements if:

- 6 - \frac{20}{59}(4+K_c)0" src="/@api/deki/files/18855/image-349.png">

$$K_c < \frac{2634}{20}$$

- 4+K_c0 \qquad" src="/@api/deki/files/18859/image-351.png">

_c-4 \qquad" src="/@api/deki/files/18861/image-352.png">

Acceptable stable range $-4 < K_c < \frac{2634}{20}$

Example 10.7.3

Consider a system with the following characteristic equation:

$$s^5 - 3s^4 + s^3 + s^2 + 4 = 0$$

Determine the stability of this system.

Answer

One of the coefficients in the characteristic equation equals 0. We replace the zero with a quantity which would be positive (approach zero from the right-hand side) and continue with the analysis as normal.

Working equation:

$$s^5 - 3s^4 + s^3 + s^2 + \epsilon s + 4 = 0$$

Row 1	1 1 ϵ

Row 2	$: -3 \qquad 1 \qquad 4$
Row 3	$\frac{4}{3} \qquad \epsilon + \frac{4}{3}$
Row 4	$4 + \frac{9\epsilon}{4} \qquad 4$
Row 5	$\epsilon + \frac{4}{3} - \frac{64}{27\epsilon + 48} \qquad 0$
Row 6	4

Since ε is positive, in the first column, there are two sign changes, from row 1 to row 2 and from row 2 to row 3. Thus, we know that the roots will have two positive real components. If you actually factor out the equation you will see that, $(s^2 + 1)(s - 2)^2(s + 1) = 0$, showing that we do have two positive roots, both equal to 2.

Additional complication exists because at row 5, as ε goes to zero, the term also goes to zero, which means that for row 5, we are getting a row full of zeros. This means that we have a pair of imaginary roots, and this situation can be solved using the equation, $a_2 S^2 + a_0 = 0$.

In this case, the working equation is, $S^2 + 4 = 0$

The imaginary roots are, $S = \pm 2i$

Example 10.7.4

You are an engineer at an icecream factory. There is a storage vat with a cooling system that has a PI controller scheme. It has the following characteristic equation:

$$10s^4 + 3s^3 - K_c s^2 + 3s + (T_i + 6) = 0$$

Your job is to determine the constraints on the values K_c and T_i such that the system is stable.

Answer

The goal is to make the matrix such that the first column has no sign changes. Since the first two entries in the first column are numbers and positive, therefore all other values in this column must be positive.

Working equation: $10s^4 + 3s^3 - K_c s^2 + 3s + (T_i + 6) = 0$

Row 1	$10 \qquad -K_c \qquad T_i + 6$
Row 2	$3 \qquad 3$

Row 3	$10 - K_c$ $T_i + 6$
Row 4	$3 - \frac{3(T_i + 6)}{(10 - K_c)}$ 0
Row 5	$T_i + 6$

Since $10 - K_c$ is in the first column, it gives the constraint $K_c < 10$. Likewise, $T_i + 6$must be positive, giving the constraint _i-6\," src="/@api/deki/files/18907/image-375.png">. Which brings the final entry of $3 - \frac{3(T_i + 6)}{(10 - K_c)}$. By dividing by three and rearrangement of the terms, it is seen that $\frac{(T_i + 6)}{(10 - K_c)} < 1$. The terms are then rearranged to arrive at the inequality: $T_i + K_c < 4$.

Aug 23, 2021, 12:13 PM

Thus the stable conditions for the constants are:
$K_c < 10$
_i-6\," src="/@api/deki/files/18907/image-375.png">
$T_i + K_c < 4$

References

- Bequette, W.B. *Process Control Modeling Design and Simulation*., New Jersey: Prentice Hall, pp 170-178.
- Foecke, H.A. and Weinstein, A. "Complex roots via real roots and square roots using Routh's stability criterion." *arxiv.org,* January 5, 2007.
- Liptak, Bela G., *Process Control and Optimization*. Vol. II. New York: Taylor & Francis.
- Ogunnaike, Babatunde A.; Ray, W. Harmon. *Process Dynamics, Modeling, and Control*. New York Oxford: Oxford University Press, 1994.

Contributors

- Authors: John D'Arcy, Matt Hagen, Adam Holewinski, and Alwin Ng
- Stewards: Jeff Falta, Taylor Lebeis, Shawn Mayfield, Marc Stewart, Tom Welch

This page titled 10.7: Routh Stability - Ranges of Parameter Values that are Stable is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

11: CONTROL ARCHITECTURES

11.1: Feedback Control
11.2: Feed Forward Control
11.3: Cascade Control
11.4: Ratio Control
11.5: Common control loops and model for liquid pressure and liquid level
11.6: Common Control Loops and Model for Temperature Control
11.7: Common control architectures and model for reactors

This page titled 11: Control Architectures is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

11.1: Feedback Control

Introduction

There are many different control mechanisms that can be used, both in everyday life and in chemical engineering applications. Two broad control schemes, both of which encompass each other are feedback control and feed-forward control. Feedback control is a control mechanism that uses information from measurements to manipulate a variable to achieve the desired result. Feed-forward control, also called anticipative control, is a control mechanism that predicts the effects of measured disturbances and takes corrective action to achieve the desired result. The focus of this article is to explain application, advantages, and disadvantages of feedback control.

Feedback control is employed in a wide variety of situations in everyday life, from simple home thermostats that maintain a specified temperature, to complex devices that maintain the position of communication satellites. Feedback control also occurs in natural situations, such as the regulation of blood-sugar levels in the body. Feedback control was even used more than 2,000 years ago by the Greeks, who manufactured such systems as the float valve which regulated water level. Today, this same idea is used to control water levels in boilers and reservoirs.

Feedback Control

In feedback control, the variable being controlled is measured and compared with a target value. This difference between the actual and desired value is called the error. Feedback control manipulates an input to the system to minimize this error. Figure 11.1.1 shows an overview of a basic feedback control loop. The error in the system would be the *Output - Desired Output*. Feedback control reacts to the system and works to minimize this error. The desired output is generally entered into the system through a user interface. The output of the system is measured (by a flow meter, thermometer or similar instrument) and the difference is calculated. This difference is used to control the system inputs to reduce the error in the system.

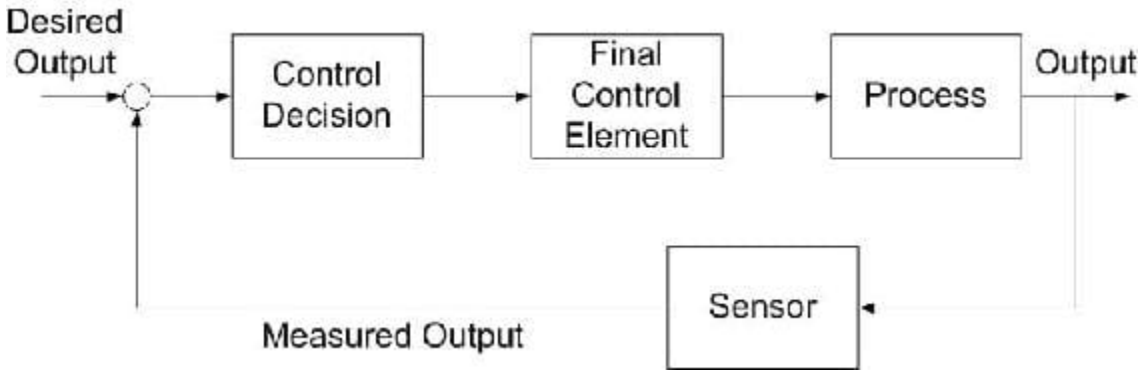


Figure 11.1.1: Feedback control loop (Adapted from Lee, Newell, and Cameron 1998: 6)

To understand the principle of feedback control, consider Figure 11.1.2 In order to bake cookies, one has to preheat an electric oven to 350°F. After setting the desired temperature, a sensor takes a reading inside the oven. If the oven is below the set temperature, a signal is sent to the heater to power on until the oven heats to the desired temperature. In this example, the variable to be controlled (oven temperature) is measured and determines how the input variable (heat into oven) should be manipulated to reach the desired value.

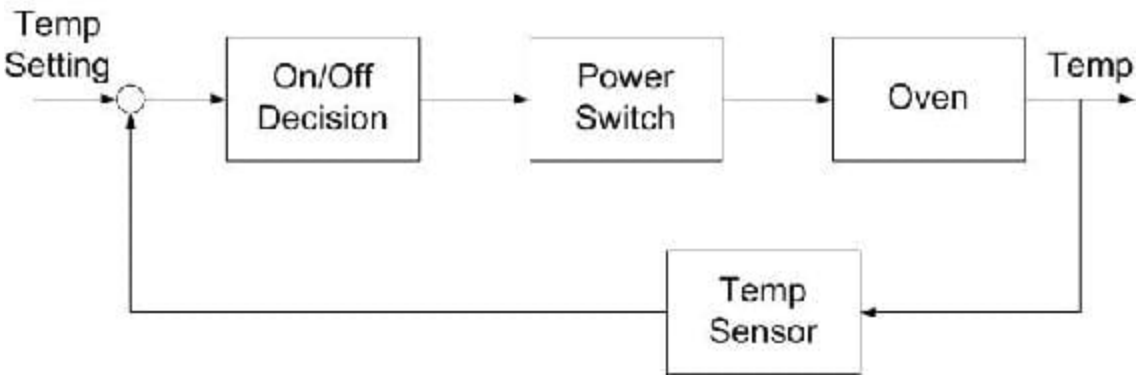
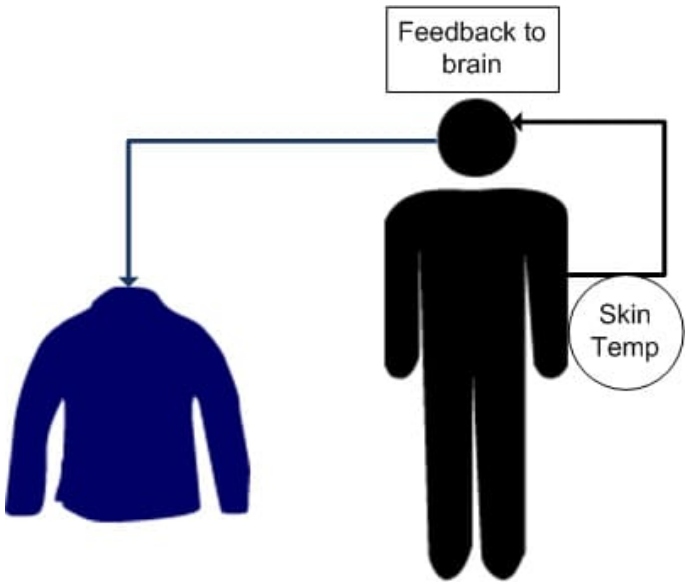


Figure 11.1.2: Feedback control in an electric oven (Adapted from Lee, Newell, and Cameron 1998: 6)

Feedback control can also be demonstrated with human behavior. For example, if a person goes outside in Michigan winter, he or she will experience a temperature drop in the skin. The brain (controller) receives this signal and generates a motor action to put on a jacket. This minimizes the discrepancy between the skin temperature and the physiological set point in the person. The example is illustrated below:



There are two types of feedback control: negative and positive. Negative feedback is the most useful control type since it typically helps a system converge toward an equilibrium state. On the other hand, positive feedback can lead a system away from an equilibrium state thus rendering it unstable, even potentially producing unexpected results. Unless stated explicitly, the term feedback control most often refers to negative feedback.

Negative Feedback

By definition, negative feedback is when a change (increase/decrease) in some variable results in an opposite change (decrease/increase) in a second variable. This is demonstrated in Figure 11.1.3 where a loop represents a variation toward a plus that triggers a correction toward the minus, and vice versa. Negative feedback leads to a tight control situation whereby the corrective action taken by the controller forces the controlled variable toward the set point, thus leading the system to oscillate around equilibrium.

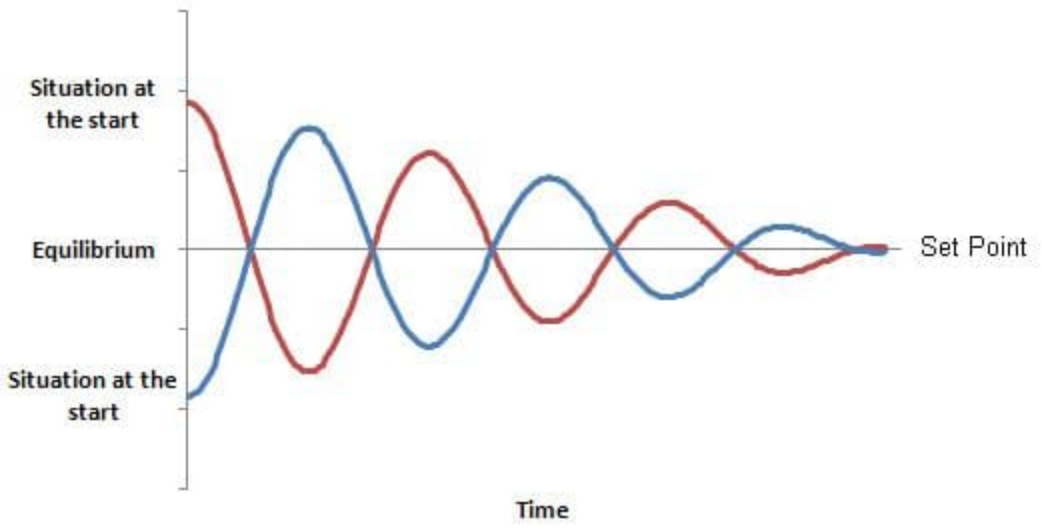


Figure 11.1.3: Negative Feedback: Maintenance of equilibrium and convergence

Figure 11.1.4 shows the negative feedback mechanism between Duck Population and Duck Death Rate. For a given flock, we say that if the death rate increases, the duck flock will decrease. On the contrary, if the duck flock increases, the death rate of the flock will decrease.

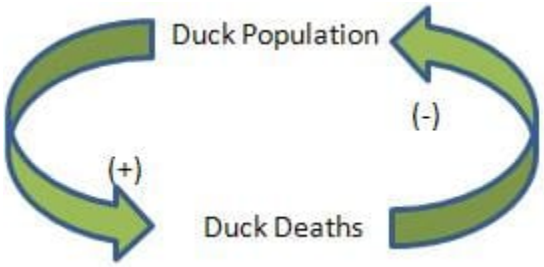


Figure 11.1.4: Duck population negative feedback loop

Positive Feedback

As opposed to negative feedback, positive feedback is when a change (increase/decrease) in some variable results in a subsequently similar change (increase/decrease) in a second variable. In some cases, positive feedback leads to an undesirable behavior whereby the system diverges away from equilibrium. This can cause the system to either run away toward infinity, risking an expansion or even an explosion, or run away toward zero, which leads to a total blocking of activities (Figure 11.1.5).

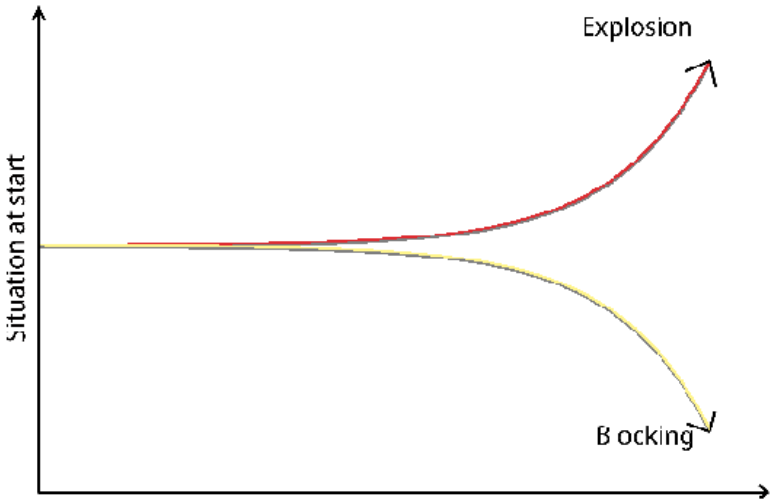


Figure 11.1.5: Positive Feedback: Exponential Growth and divergent behavior, no intermediate situation

Figure 11.1.6 shows the feedback mechanism responsible for the growth of a duck flock via births. In this example, we consider two system variables: Duck Births and Ducks Population. For a given flock, we state that if the birth rate increases, the duck flock will increase. Similarly, if the duck flock increases, the birth rate of the flock will increase.

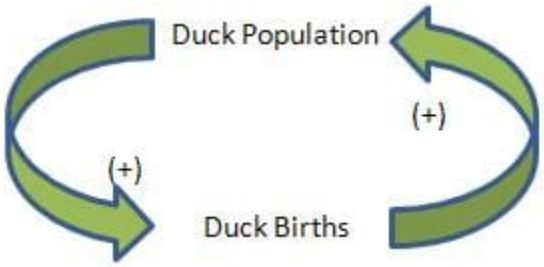


Figure 11.1.6: Duck population positive feedback loop

Applications

Control mechanisms are used to achieve a desired output or operating condition. More specifically, feedback control is a useful mechanism in numerous applications across many engineering fields. In chemical engineering, feedback control is commonly used to manipulate and stabilize the conditions of a CSTR. Figure 11.1.7 shows how feedback control can be effectively used to stabilize the concentrations of reactants in a CSTR by adjusting the flow rates.

CSTR with Feedback Control

Several types of feedback control can be used to manipulate the conditions in a CSTR: positive feedback, negative feedback, or a combination of both. Figure 11.1.7 illustrates each of these possible situations. As depicted below, each CSTR is equipped with two electrodes that measure the voltage of the solution contained inside the reactor. A computer adjusts the flow rates of the pump(s) in response to any changes in the voltage.

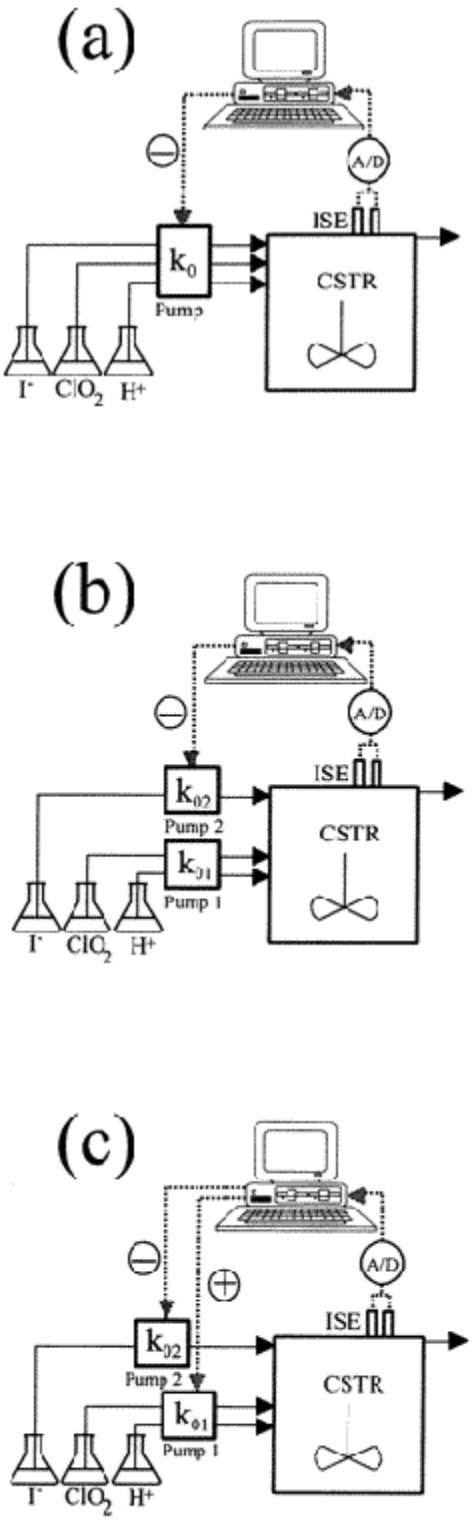


Figure 11.1.7: CSTR with feedback control: equipment and control configuration (Dolnik et al. 1997)

- **(a)** All of the reagents are pumped into the reactor by the same pump. The flow rate through the pump is adjusted constantly by a negative feedback mechanism; when level of the iodide solution is low, the computer detects the insufficiency and increases the flow rate of all the reactants.

- **(b)** The iodide solution is pumped into the reactor by one pump, while the other two reactants are pumped in at a constant flow rate. The flow rate of the iodine solution is controlled by a negative feedback mechanism; when the computer detects an insufficient amount of iodine in the CSTR, it signals the pump. The flow rate of iodine into the CSTR is therefore increased.
- **(c)** Two pumps are used to feed the reactor: one pump for the iodine solution and one for the ClO_2 and H^+ solutions. The flow of the iodine solution is controlled by a negative feedback mechanism; if the computer detects an iodine deficiency, it signals Pump 2 to increase the flow rate of the iodide solution. The flow rate of the ClO_2 and H^+ solutions is controlled by a positive feedback mechanism; if the computer detects an iodide deficiency, it will signal Pump 1 to decrease the flow rates of ClO_2 and H^+,thereby increasing the concentraion of iodide.

It is easy to see that by combining feedback controls, such as in Figure $11.1.1c$, output concentrations and operating conditions can be maintained at the desired state much more effectively than by only using one type of feedback control.

Advantages and Disadvantages

The unique architecture of the feedback control provides for many advantages and disadvantages. It is important to look at the exact application the control will be used for before determining which type of control will be the best choice (see Cascade Control Systemsand Feed Forward Control).

Advantages: The advantages of feedback control lie in the fact that the feedback control obtains data at the process output. Because of this, the control takes into account unforeseen disturbances such as frictional and pressure losses. Feedback control architecture ensures the desired performance by altering the inputs immediately once deviations are observed regardless of what caused the disturbance. An additional advantage of feedback control is that by analyzing the output of a system, unstable processes may be stabilized. Feedback controls do not require detailed knowledge of the system and, in particular, do not require a mathematical model of the process. Feedback controls can be easily duplicated from one system to another. A feedback control system consists of five basic components: (1) input, (2) process being controlled, (3) output, (4) sensing elements, and (5) controller and actuating devices. A final advantage of feedback control stems from the ability to track the process output and, thus, track the system's overall performance.

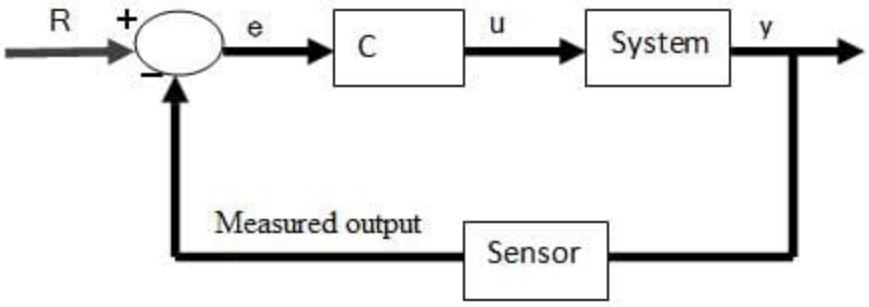
Disadvantages: Time lag in a system causes the main disadvantage of feedback control. With feedback control, a process deviation occurring near the beginning of the process will not be recognized until the process output. The feedback control will then have to adjust the process inputs in order to correct this deviation. This results in the possibility of substantial deviation throughout the entire process. The system could possibly miss process output disturbance and the error could continue without adjustment. Generally, feedback controllers only take input from one sensor. This may be inefficient if there is a more direct way to control a system using multiple sensors. Operator intervention is generally required when a feedback controller proves unable to maintain stable closed-loop control. Because the control responds to the perturbation after its occurrence, perfect control of the system is theoretically impossible. Finally, feedback control does not take predictive control action towards the effects of known disturbances.

Closed Loop Control versus Open Loop Control

Although there are various types of controllers, most of them can be grouped into one of the two broad categories: closed loop and open loop controllers. The subsections below summarize the differentiation.

Closed Loop System

In a closed loop control system, the input variable is adjusted by the controller in order to minimize the error between the measured output variable and its set point. This control design is synonymous to feedback control, in which the deviations between the measured variable and a set point are fed back to the controller to generate appropriate control actions.The controller C takes the difference e between ther reference r and the output to change the inputs u to the system. This is shown in figure below. The output of the system y is fed back to the sensor, and the measured outputs go to the reference value



Open Loop System

On the other hand, any control system that does not use feedback information to adjust the process is classified as open loop control. In open loop control, the controller takes in one or several measured variables to generate control actions based on existing equations or models. Consider a CSTR reactor that needs to maintain a set reaction temperature by means of steam flow (Figure 11.1.8): A temperature sensor measures the product temperature, and this information is sent to a computer for processing. But instead of outputting a valve setting by using the error in temperature, the computer (controller) simply plugs the information into a predetermined equation to reach output valve setting. In other words, the valve setting is simply a function of product temperature.

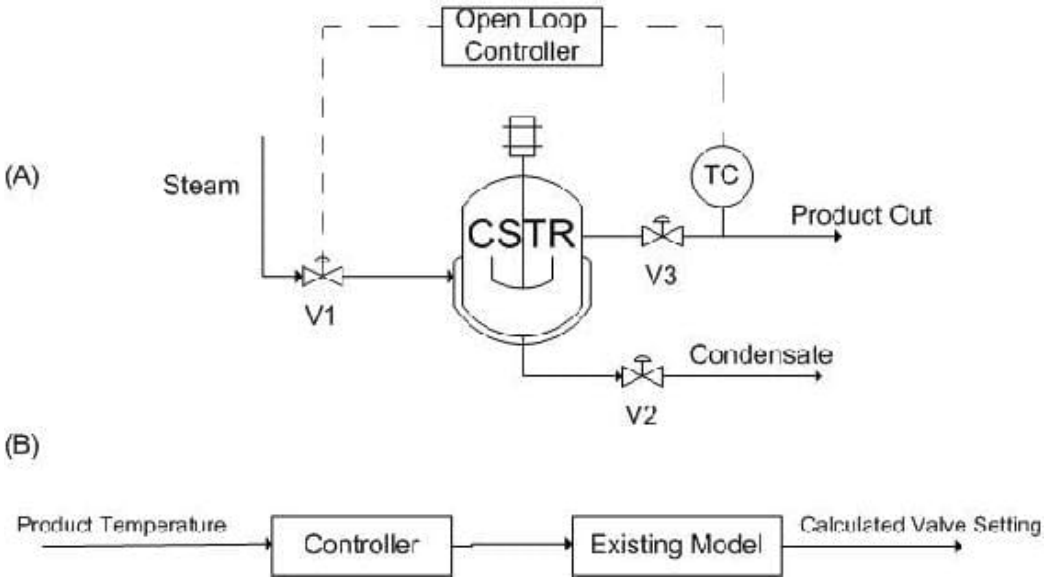


Figure 11.1.8: Example of CSTR open loop controller. (A) System diagram. (B) Open control loop.

Note that the open loop controller only uses the **current state** of the measured variable (product temperature) and a model to generate its control output (valve setting), as opposed to monitoring errors that have already taken place. As the result, the quality of the control system depends entirely upon the accuracy of the implemented model, which is challenging to develop. For this reason, feedback, or closed loop, controllers are generally recognized as the more reliable control system.

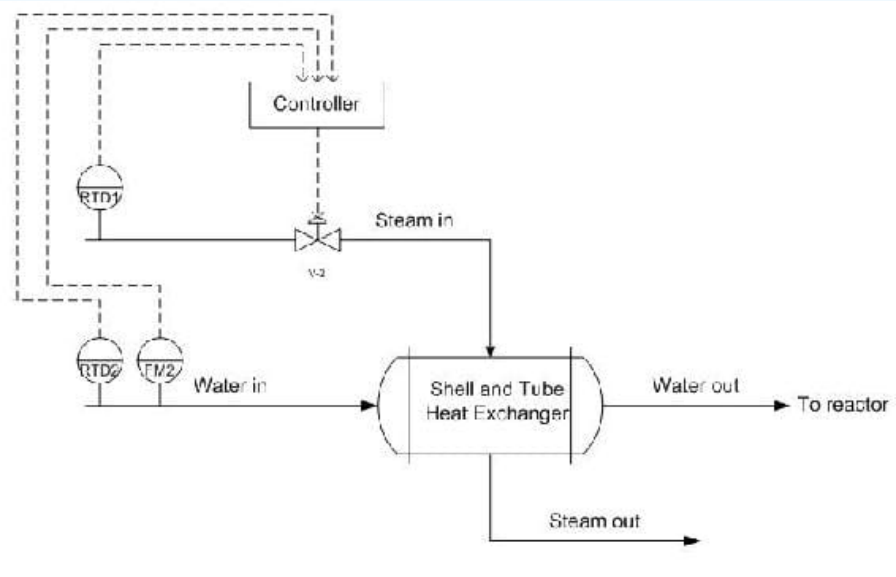
Short Summary on Closed and Open Loop Controllers

- Feedback Controller = Closed Loop Controller
- Non-Feedback Controller = Open Loop Controller

Example 11.1.1

Your company is designing a plant to produce biodiesel. Several of the unit processes require heated water that is under tight temperature control. The desired temperature of the water is 65°C ± 1°C. To accomplish this, steam will be used in a shell and tube heat exchanger to heat the incoming water.

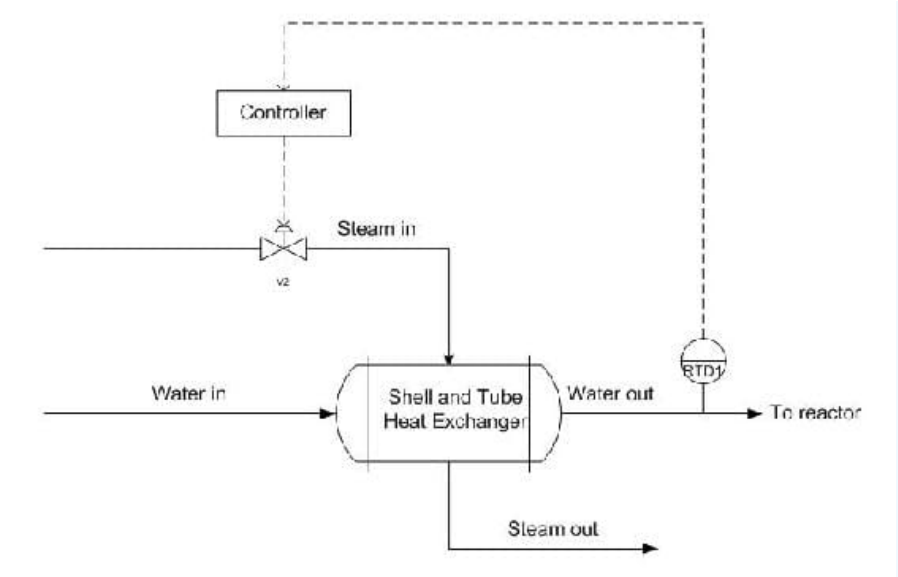
A member on the project team suggests the following feed forward control system to regulate the temperature of the water stream. Temperature and flow rate data for the incoming steam and water streams will be used along with energy balance calculations to determine the flow rate of steam required to keep the exiting water stream at the appropriate temperature.



You on the other hand suggest a feedback control system. Place the appropriate sensors, actuators, and controllers to implement a feedback control system that will allow for the control of the exiting water temperature. What are some of the advantages and disadvantages of using a feedback control system over a feed forward control system?

Solution

In the feedback control system, the temperature of the variable to be controlled is measured directly. If the exiting water temperature is too high, the controller will decrease the amount of steam entering the shell side of the heat exchanger. Likewise, if the exiting water temperature is too low, the controller will increase the amount of steam entering the shell side of the heat exchanger. This is an example of negative feedback at work to keep the exiting temperature at the setpoint.



Advantages

- Effective regardless of the origin of the disturbance.
- Relatively simple to implement. A mathematical model of the exiting water temperature is not needed.

Disadvantages

- Corrective action taken only after there is a deviation in the output from the set point.
- Can cause instability if incorrectly tuned

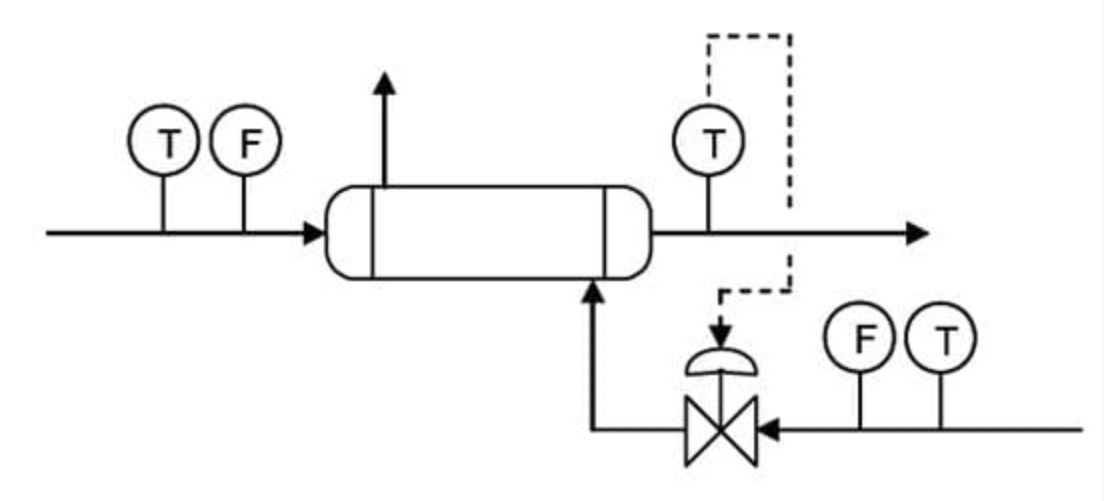
Example 11.1.1

Of the following two scenarios, which would be the best to use a feedback control?

- **Solution**
  - A feed-forward control or feed-forward/feedback system would most appropriate for the first scenario. In this situation, it is critical that the pH of the waste stream not drop below the setpoint of pH 6. It is also desirable to minimize the use of lime, since this reduces the cost of the waste treatment. Therefore very tight process controls are desired, for which a feed-forward or feed forward/feedback system would be most appropriate. A feed-forward system can adjust to changes in inputs before they cause deviations in the output stream. Additionally, since there are important fluctuations two process inputs (pH variations in both the lime slurry and the acidic waste), feed-forward controls would be favored over feedback control.
  - A feedback control would be most useful for the second scenario. There is only one monitored variable (concentration of reduced wine) and only one manipulated variable (steam). Since minor fluctuations about the wine concentration set point are not critical, the performance of a typical feedback control system would be acceptable. It would be hard to implement a feed-forward control system since the concentration of the wine feed to the evaporator may fluctuate and is not monitored.

Example 11.1.3

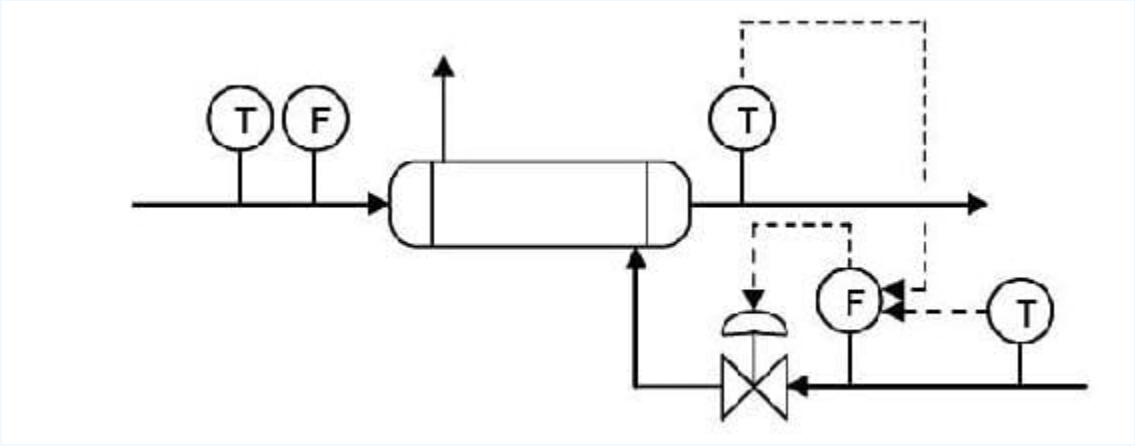
Using a shell and tube heat exchanger, a fuel stream is pre-heated using a steam stream from the utilities part of the plant. The plant is located in Ann Arbor, Michigan. Originally, a feedback controller is implemented to control the amount of steam coming into the shell and tube exchanger based on the temperature of the flow out of the exchanger. This is depicted below.



Upon implementation in the process, it was found that the feed temperature was not the only factor in the variance of the outlet stream. Because the plant is operated in Michigan, the steam temperature varies between seasons, colder in the winter and warmer in the summer. Propose a control scheme to account for both of these variances.

Solution

Here a cascade control scheme would be optimal to control the outlet temperature of the stream to a desired temperature, not just a feedback controller. The scheme is depicted below.



The feedback control previously used to control the system is still implemented, however it is now reporting to the flow controller before the steam control valve. Also, there is a temperature controller now incorporated in the control scheme. This is a feed forward controller that also controls the flow controller for the steam control valve. Thus the primary control loops, the feed forward and feedback control loops, comprise the secondary control loop in the cascade control. This incorporates both the temperature of the incoming feed stream, which is controlled by feedback control via the temperature sensor after the exchanger. This is because the temperature of the outlet stream would vary with the temperature of the inlet stream if the steam stream were not accounted for. But accounting for the temperature of the steam stream, via a feed forward controller that is based on a temperature sensor before the control valve, will give a better approximation and more control over the temperature of the outlet stream. This is also a very good example of combining several control architectures and how they function in a real-world application, and why a feedback controller in not necessarily the most accurate control.

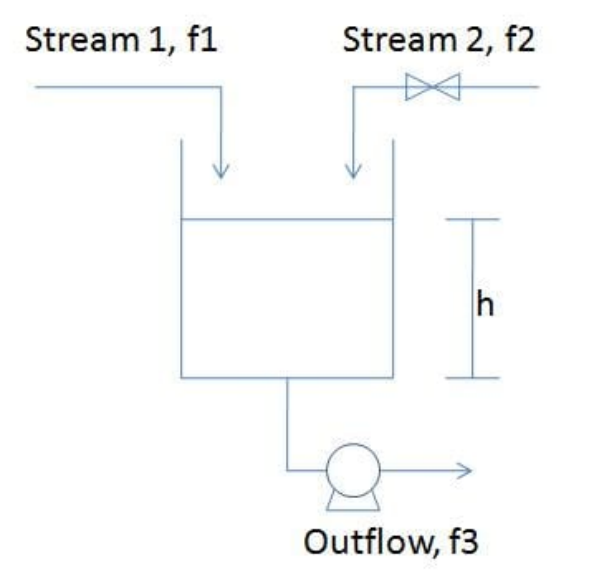
Example 11.1.4

A blending system is used to mix a process stream (Stream 1, pure A) with liquid catalyst (Stream 2, pure B). The mixture then flows into a reactor. The flowrate of stream 1 (f1) can vary but it is possible to manipulate the flowrate of stream 2 (f2). The outflow rate from the reactor (f3) also can be manipulated via a pump. The level in the measures by a level transmitter that sends its signal to a level controller. All three flow rates can be measured, but measuring the composition of the streams is impossible. In order to control the level of the tank, and engineer has two options:

i. manipulate f2 while holding f3 constant
ii. manipulate f3 while holding f2 constant

Could method (i) utilize feedforward or feedback control of the level? If so, explain how it can be done.

Could method (ii) utilize feedforward or feedback control of the level? If so, explain how it can be done.



Solution

For configuration (i)

- Feedback: If you measure h, when it goes up, you can reduce f2, and when it goes down, you can increase f2
- Feedforward: If you measure f1, when it goes up, you can reduce f2, and and when it goes down, you can increase f2. You are indirectly controlling h due to the mass balance.

For configuration (ii)

- Feedback: If you measure h, when it goes up, you can increase f3, and when it goes down, you can decrease f3.
- Feedforward: If you measure f1, when it goes up, you can reduce f3, and when it goes down, you can reduce f3.

This problem was meant to explore different strategies for both feedback and feedforward control configurations. Additionally, it demonstrates how controlling the manipulated variables (either f2 for configuration (i) or f3 for configuration (ii)) affects the desired outcome (maintaining consistent tank height).

References

- Lee, P., Newell, R., Cameron, I.(1998).Process Control and Management, London:Blackie Academic & Professional. ISBN 0-7514-0457-8
- Marlin, T.(2000). Process Control: Designing Processes and Control Systems for Dynamic Performance, Boston:Mcgraw Hill.

ISBN 0-07-039362-1
- Palazoglu, A., Romagnoli, J.(2006).Introduction to Process Control, Boca Raton:Taylor & Francis. ISBN 0-8493-3696-9
- Perry,R.H.(1997). Perry's Chemical Engineer's Handbook, 7th Edition, McGraw-Hill. ISBN 978-0-07-049841-9
- National Lime Association Article. Using Lime for Acid Neutralization. September 2000.
- MIT Open Courseware. (2006). "Lesson 8: Cascade and Feedforward Control Schemes," [1]

Contributors and Attributions

- Authors: Eddy Kim, Bing Liu, Tyler Roehm, Samah Tout
- Stewards: Gillian Berberich, Katie Feldt, Christopher Mark, Jason Wong
- Cara Canady, David Carpenter, Che Martinez, Jeremy Minty, Bradley Novak

This page titled 11.1: Feedback Control is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

11.2: Feed Forward Control

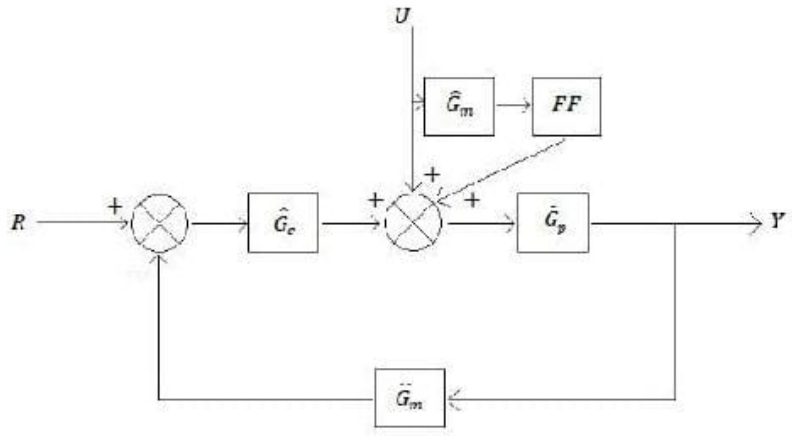
Introduction

Feed-forward control is a useful tool in the field of chemical engineering when there is a known set of deviations occurring upstream of the system. This would allow engineers to account for that particular deviation within the controller and reduce the effects of the deviation on the system. An example would be a car's cruise-control system. If it has feedback control, when there is a slope and therefore a speed reduction, the feedback controller would compensate by applying additional throttle to the engine. If it uses a feed-forward controller instead, the controller would calculate the slope beforehand and throttle the engine before any speed reduction occurs. In this sense, the controller predicts the incoming deviation and compensates for it.

The following block diagram shows a feed-forward controller implemented on a idealized process with a setpoint R and load U:

where:

- $\hat{G}_P$ represents the process operator,
- $\hat{G}_M$ represents the measurement operator,
- $\hat{G}_C$ represents the controller operator, and
- FF is the feed-forward controller.



The perfect feed-forward controller is the inverse of the process operator, $\hat{G}_P^{-1}$

For example:

$$\hat{G}_P^{-1} = K_P\left(\tau\frac{\delta}{\delta t}+1\right) = \hat{G}_{FF}$$

In General:

$$\hat{G}_P^{-1}Y = \left(U+\hat{G}_{FF}U\right)+\hat{G}_C(R-Y)$$

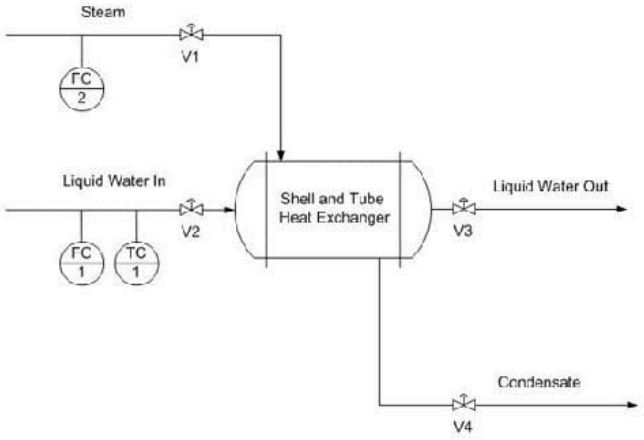
(in the language of operator or state space).

Feed-Forward Control

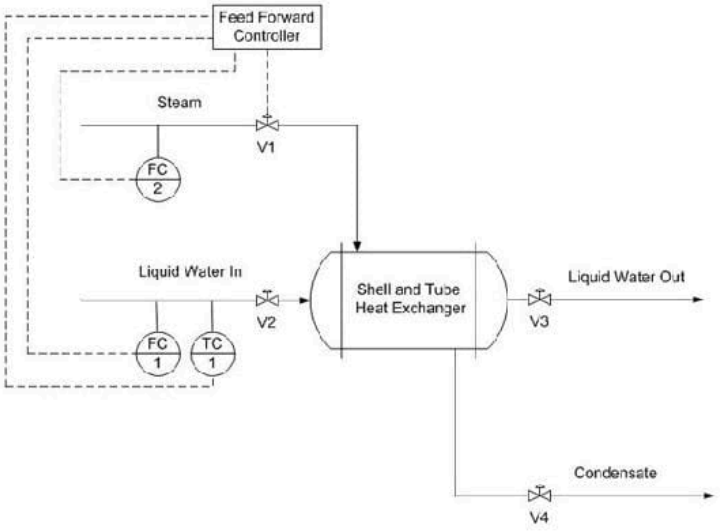
The objective of feed-forward control is to measure disturbances and compensate for them before the controlled variable deviates from the setpoint. Feed-forward control basically involves a control equation which has certain corrective terms which account for predicted disturbances entering the system. The equation is only effective for gains in a steady state process. Dynamic compensation should be used in the control equation if there are any dynamic deviations with the process response to the control action. This dynamic compensation ability will be discussed further in the next section.

One form of a feed-forward control would be a Derivative (D) control which calculates the change in error and compensates proportionately. But a D-control can't perform by itself and usually requires working in conjunction with a Proportional (P) or Proportional Integral (PI) control. More information can be found in the PID Intro article.

Simulated below is a typical shell and tube heat exchanger which heats up liquid water using steam.



Adding a feed-forward control into the system manipulates the amount of steam required to compensate for the varying amounts of liquid feed coming in.



Overall heat balance:

Heat into the system is equal to the heat leaving the system.

Heat gained by liquid:

$$q_{\text{out}} = m_l C_p \left(T_2 - T_1\right) \tag{11.2.1}$$

where:

- m_l -- mass flow rate of the liquid
- C_p -- the heat capacity of the liquid
- T_1 -- the input temperature of the liquid
- T_2 -- the desired temperature setpoint for the liquid

Heat lost by steam:

$$q_{in} = m_s \lambda \tag{11.2.2}$$

where:

- λ -- heat of vaporization
- m_s -- mass flow rate of the steam

Therefore:

$$m_s = \frac{C_p}{\lambda} m_l \left(T_2 - T_1\right) \tag{11.2.3}$$

To compensate for a deviation from the desired inlet temperature of the liquid water, the amount of steam must be adjusted. Equation (3) acts as the control equation for the feed-forward controller. This controller actuates the inlet valve of the steam and acquires temperature and flow rate data from the inlet streams. Based on the desired temperature setpoint, T2, and the actual input temperature, T1, we can use this equation to calculate the mass flow rate of the steam required to compensate for the temperature deviation.

Accounting for System Non-Idealities

Often, one of the most difficult tasks associated with created a functional feed forward controller is determining the necessary equations that govern the system. Even more complexities arise when the system is not, and cannot be treated as, ideal. This is the case with many real and practical systems. There is simply too much heat lost or too many unforeseen effects to safely assume ideal conditions. The pure math of the example above does not account for these effects.

The equations will output a value to a control valve (often in voltage). That voltage will be derived from some previously determined relationship between voltage, valve %open, and steam flow rate. A very simple way to begin managing the issue of non-ideality is by including a "non-ideality constant". This can be an additive or a multiplicative constant that adjusts the voltage output determined by the equations.

- Voltage output is some function of the calculated required steam flow

  $$V = f\left(m_s\right)$$

- Voltage output is adjusted by some constant

  $$V = c_N * f\left(m_s\right)$$

  or

  $$V = c_N + f\left(m_s\right)$$

This non-ideality constant c_N often must be determined by trial. One way to accomplish this is to use manual control to determine the output voltage needed at various inlet conditions. Using the data from the manual trials, and the resulting voltage that your unadjusted feed forward controller would output, it is possible to determine by what factor your feed forward voltage needs to be

adjusted. For example, if by manual control you determine that for inlet conditions X the required voltage is 300 mV and your feed forward controller is only outputting 270 mV for conditions X, you need some factor to adjust.

It may also happen that your "non-ideality constant" will not turn out to be constant when you begin to look at the data. In this situation, consider using a linear relationship between the non-ideality factor and some inlet condition or implementing CASE or IF statements covering discrete ranges of inlet conditions and giving different constants.

Dynamic Compensation

Dynamic compensation is a method to account for factors such as lead and lag times when using feed-forward control. For instance, in the example above, when the feed forward controller monitors a temperature decrease in the liquid feed, it will increase the steam flow rate - the manipulated variable - to compensate and thus maintain the temperature of the exiting liquid flow - the controlled variable. However, the steam may enter the heat exchanger faster than the liquid feed and this will cause a transient increase in the controlled variable from the setpoint. In an ideal case the steam and liquid feed would enter the heat exchanger at the same time and no deviation from the set exiting temperature would be observed. Therefore, dynamic compensation involves predicting non-ideal behavior and accounting for it. While perfect feed-forward control is nearly impossible, dynamic compensation is one step closer.

Open Loop System

Feed-forward control is a open loop system. In an open loop system, the controller uses current, or live, information of the system to generate appropriate actions by using predetermined models. The sensor providing the reference command to the closed loop actuator is not an error signal generated from a feedback sensor but a command based on measurements. This is the defining characteristic of an open loop system, in which the controller does not manipulate the system by trying to minimize errors in the controlled variable.

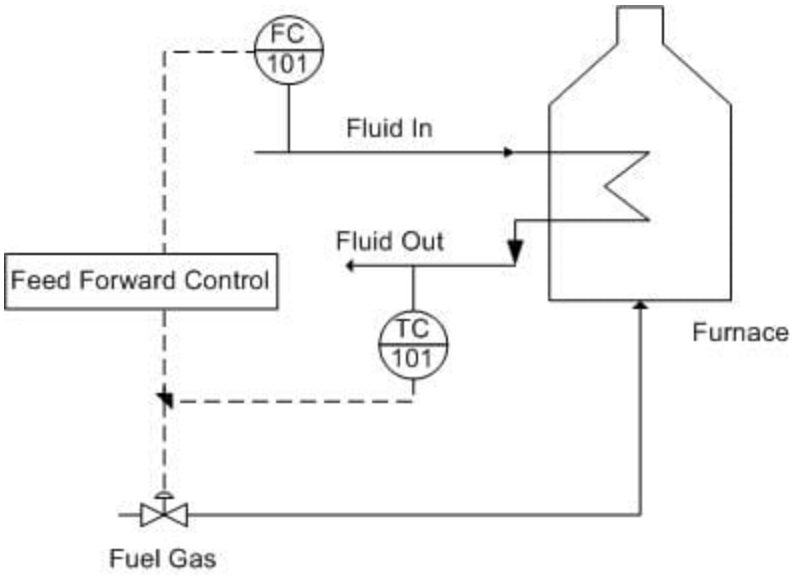
Because a feed-forward controller listens to the system and calculates adjustments without directly knowing how well the controlled variable is behaving (if it does, it is getting feedback), it is open loop in nature. In other words, the controller operates on "faith", that its mathematical models are able to accurately generate responses (valve setting, motor speed, etc.) that lead to expected results. Therefore, it is critical to implement good models in feed-forward systems. This is often the most difficult part of a feed-forward design, as mentioned previously.

However, not all open loop systems are feed-forward. Open loop simply means the system is not getting feedback information, in which the controlled variable tells the controller how well it is doing compared to its set point (i.e. error). Feed-forward control, on the other hand, takes an extra step by using non-feedback information to produce predictive actions.

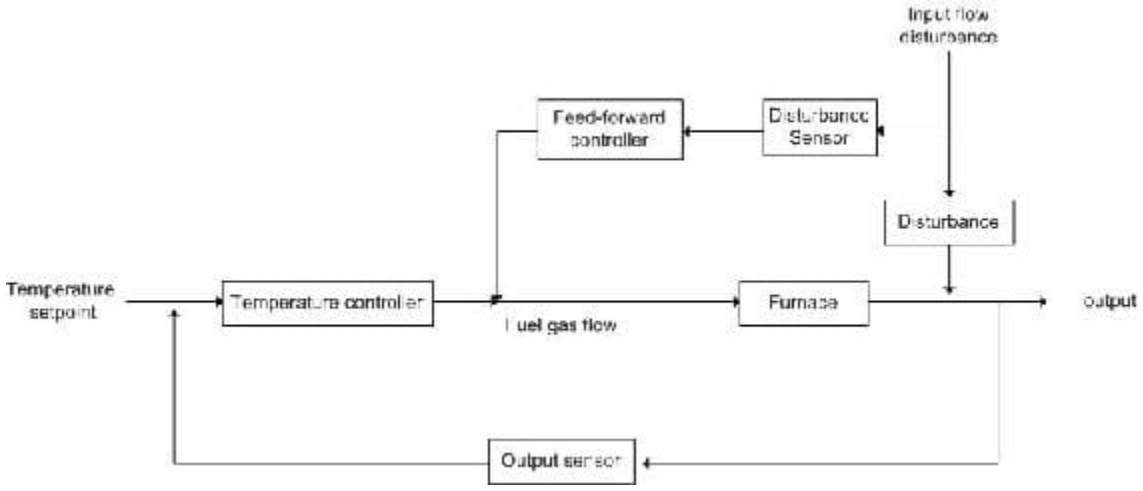
This simple open loop system is an example of a open loop system that is not feed-forward. The controller simply takes the current state of the controlled variable to generate a valve setting. It is not getting a feedback, because the product temperature is not compared to a set pont. It is not feed-forward either, because there is no mechanism that produces actions that may yield expected results in the future.

Feed-forward applications

Feed-forward control is used in many chemical engineering applications. These include heat exchangers, CSTRs, distillation columns and many other applications. A typical furnace, shown below, is heating up an input fluid using fuel gas.



One possible disturbance is the flow rate of the incoming fluid. For example: If the fluid input rate was increased by 10%, then the required heat duty should also be increased by approximately 10%. This feed-forward control strategy immediately changes the fuel gas flow rate and therefore the heat duty. The performance of feed-forward controls is limited by model uncertainty, and in practice feed-forward control is combined with feedback control in order to ensure better control of the system. Feed-forward control will compensate for fluid input disturbances while feedback control will compensate for other disturbances—such as the fuel gas flow rate or the temperature of the furnace—and model uncertainty. This system can also be represented by the following block diagram.



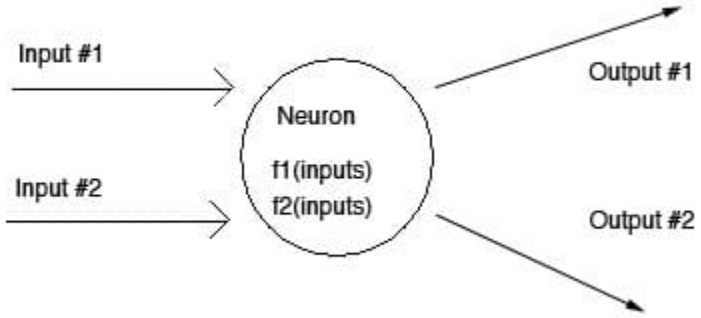
Pros & Cons of Feed-Forward Control

Different applications require different types of control strategies. Sometimes feed-forward solutions are required for proper system control; sometimes only feedback solutions are necessary. Feedback systems don't always maintain the setpoint as well because of the lag that comes with waiting for the disturbance to propagate through the system. As a result, many control systems use a combination of feed-forward and feedback strategies, such as PID controllers. PID controllers use the Proportional-Integral control for feedback and the Derivative control for feed-forward control. This forms a system with multiple loops, otherwise known as a cascading system. A critical advantage of running both forward and backward controls is that the system is still somewhat able to adjust a variable if one mechanism fails since the two loops use different sensors. As a result, PID controllers are great for controlling processes, however, they require a number of equations to determine feed-forward and feedback correction.

Feed-forward systems work by checking the conditions of an incoming stream and adjusting it before the system is adversely affected. If the controller is told the traits of an acceptable incoming stream, then it can compare that standard to whatever is

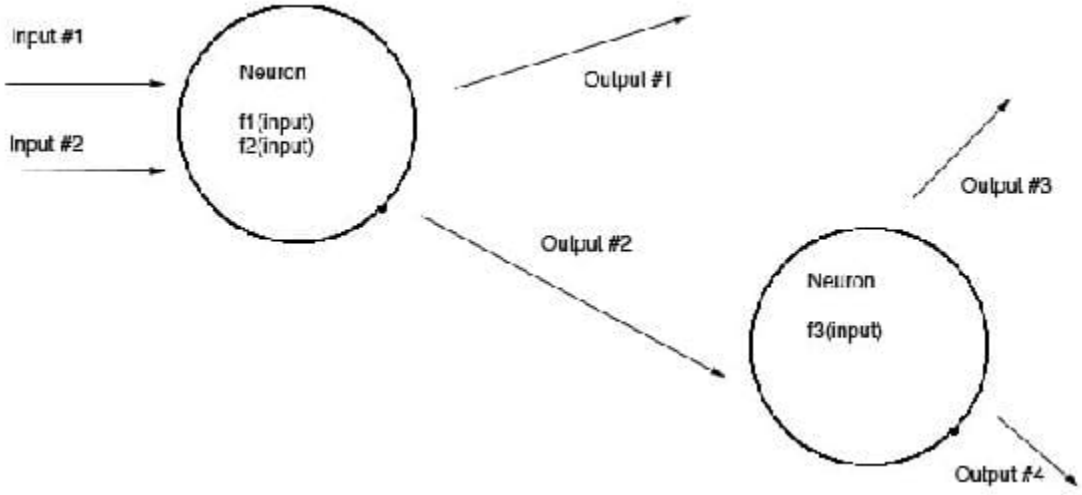
coming down the pipe. The feed-forward controller can look at this error and send a corrective signal to the automatic valve responsible for that pipe (or any other control device). In order to have this kind of predictive ability, the controller must have explicitly defined equations that account for the effects of a disturbance on the system. In addition, these equations must also then prescribe action to counter-act the disturbance. This can become even more difficult when there are several incoming stream traits that are being observed. When there are multiple inputs, the feed-forward system will require non-linear equations, leading to the development of neural networks.

Neural networks are based on “neurons”, which are representations of non-linear equations. This concept is based on the brain's use of neurons to process and transmit information. The neuron is actually comprised of a set of sigmoidal equations relating inputs to outputs. Sigmoid functions are non-linear equations that take inputs and apply constants, or weights, to transform the value to make an output. Below is a picture that illustrates the function that these neurons serve.



The picture above shows a multiple input-output system. Most feed-forward systems have to deal with more than one input.

Neurons can be “connected” in ways that allow the inputs to be transformed any number of times. Neurons that are connected indicate that one sigmoidal function's output becomes the input of another one. Although the concept of neurons is easily understandable, the difficulty lies in the potential complexity of real systems. For example: the number and type of inputs go to which neurons, the initial weights be, the number of neurons needed, etc. As you can see, there are a lot of design questions that make developing the neural network difficult.

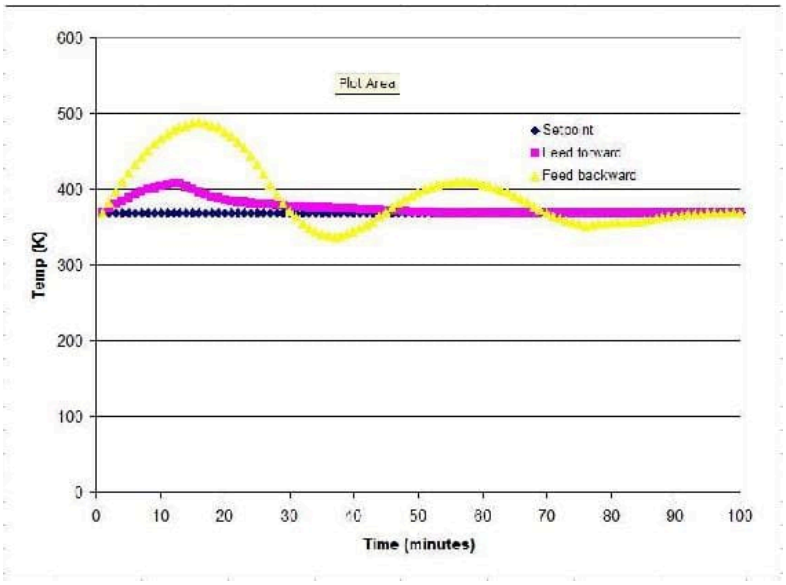


Coupled with an expected set of upstream conditions, the feed-forward system can continually adjust the method it uses to control an output variable. The system is capable of doing this by measuring sensor inputs coming into the controller, using neurons to transform the data, and comparing the resulting output(s) to a set of known or desired outputs. If the actual output is different from the desired output, the weights are altered in some fashion and the process repeats until convergence is achieved. This is how a controller is able to "learn". Learning is just discovering the weighting factors that allow the transformed ouputs to match the desired outputs.

The following table shows a list of feed-forward pros and cons:

Feed-Forward Pros	Feed-Forward Cons
• Does not require disturbance to propagate through system for detection • Able to predict incoming disturbance before they hit the system • Neutral network utilized to familiarize control system with a changing environment • Provides better set point management than feed backward control • Can be used in conjunction with teed-back control to provide multiple input single input MISO control	• Explicit non linear equation required for adequate disturbance prediction. Much harder to implement as a result • Cannot compensate for any disturbances that occur WITHIN the system • Requires infinitely accurate models • Requires infinitely accurate measurements

A CSTR with a given volume with heat-exchange capability has a coolant water system to maintain a specific system temperature (368K). To maintain this temperature, the flow of coolant water oscillates. Let a temperature disturbance of 100K be introduced over a period of 10 minutes. For this simulation, the Excel spreadsheet from PID-Tuning Optimization was used to create feed-forward and feed-backward data. In the following plot, observe the difference between using feed-forward control versus feed-backward control in an example.



As shown, the feed-forward strategy works more effectively than the feed-backward strategy at nullifying the disturbance (over the time frame). The temperatures reached in feedback mode vary more than the ones seen in feed-forward mode. However, the controller response is solely defined by the parameters used in the Excel spreadsheet. If one assumes that the PID parameters used by the **PID-optimization group** were optimized, then we can say feed-forward would be the best option here.

Feed-Forward Design Procedure

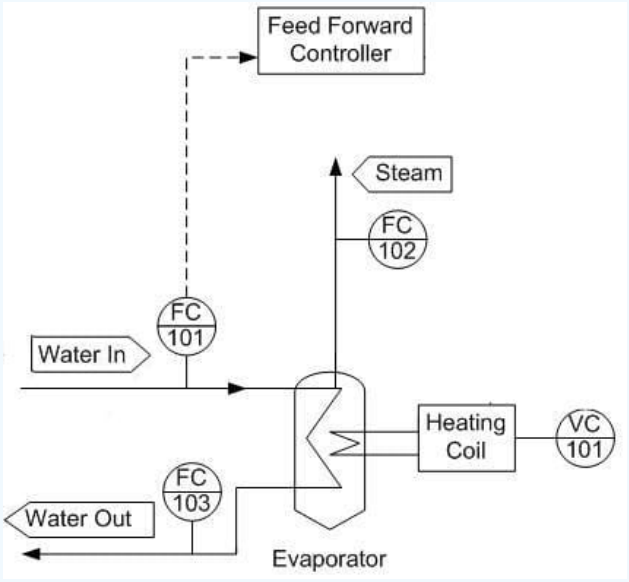
This section provides an outline of the different steps for designing a feed-forward control strategy. The steps pay close attention to designing a feed-forward controller where there are multiple disturbances. Feed-forward design can be broken down into eight steps (Corripio,1990).

- **Step 1.** State the control objective. This step includes defining which variable needs to be controlled and what the set point is. The setpoint should be adjustable by the operator.

- **Step 2.** List the possible measured disturbances. This step includes identifying which disturbances are easily measured and how fast each disturbance should be expected to vary.
- **Step 3.** State which variable is going to be manipulated by the feed-forward controller.
- **Step 4.** The feed-forward controller consists of two parts: steady-state and dynamic compensators. Develop the steady-state compensator first. The compensator should be an equation where the manipulated variable, identified in step 3, can be calculated from the measured disturbances, identified in step 2, and the control objective (set point), identified in step 1.
- **Step 5.** Reevaluate the list of disturbances. The effect of a disturbance on the controlled variable can be calculated from the equation. Three criteria will be used to determine which disturbance the feed-forward controller will correct: the effect the disturbance has on the controlled variable, the frequency and magnitude of variation, and the capital cost and maintenance of the sensor.
- **Step 6.** Introduce the feedback compensation. This depends on the physical significance assigned to the feedback signal.
- **Step 7.** Decide whether dynamic compensation, lead/lag, and/or dead time is required, and decide how to introduce it to the design.
- **Step 8.** Draw the instrumentation diagram from the feed forward control strategy. The details of the diagram depend largely on the control system being used.

Example 11.2.1

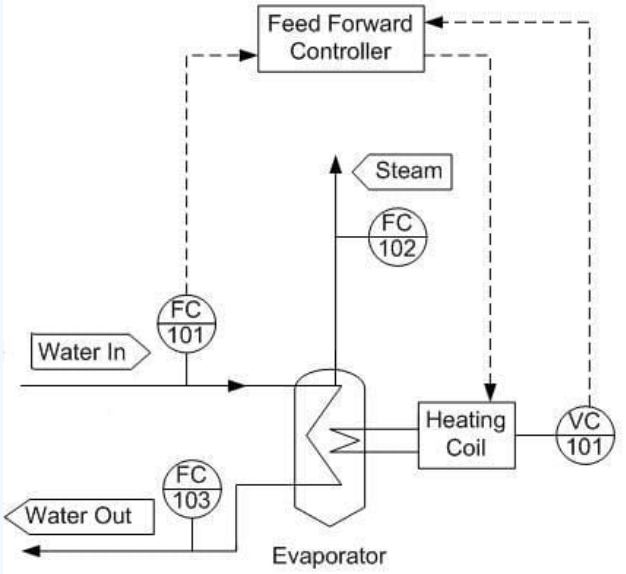
For the first example, a water stream enters an evaporator. A discrete amount of the water stream is to be evaporated. Below is the design diagram for the evaporator. The incoming stream is being used as the input to the feed forward control box.



Determine what additional input will be needed to regulate the amount of steam produced and draw this on the design diagram.

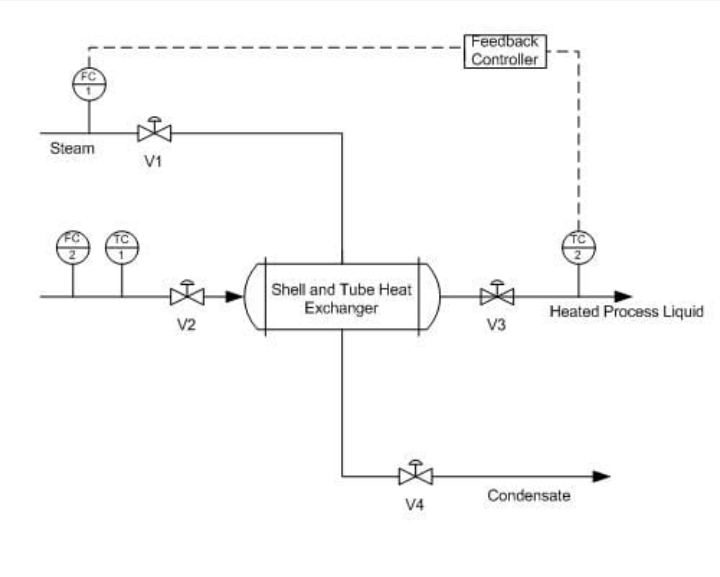
Solution

By regulating the heating coil, we can regulate how much water will boil off in accordance with any change of the feed stream.



Example 11.2.2

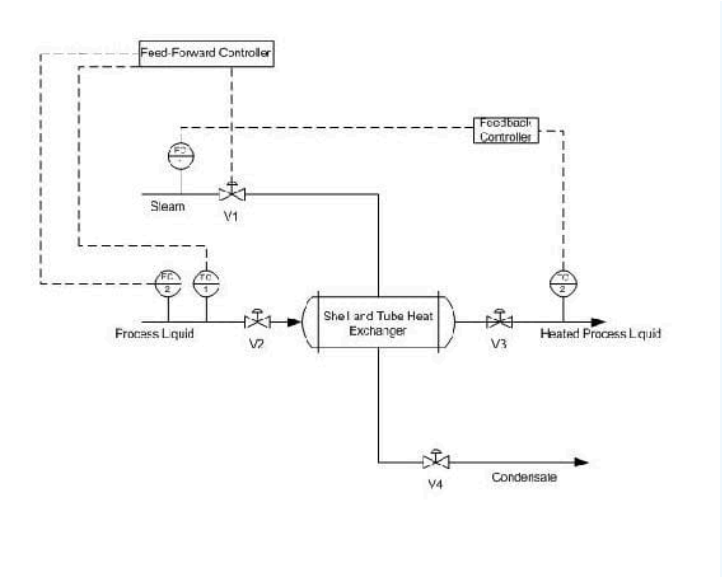
For the second example, a particular process liquid from a previous reaction will be heated before proceeding to the next phase of reactions. A shell and tube heat exchanger has been installed for this purpose with steam providing the heat duty. A feedback controller is used to control the output temperature of the process liquid stream, but the flowrate and the temperature of the input liquid stream vary too much for the feedback controller to be effective.



What sort of controller would be appropriate in addition to the feedback controller to account for the deviations of the input liquid stream? How would this controller be implemented? For simplicity it is assumed that all the steam condenses.

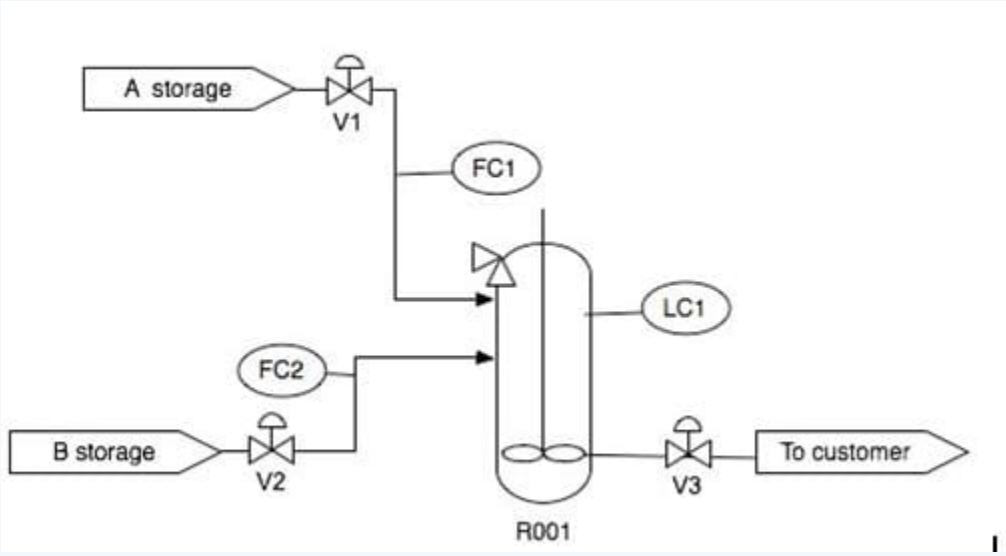
Solution

The solution is to implement a feed-forward controller that gathers the input temperature and flowrate data from TC1 and FC2 and adjusts the steam valve V1 in order to account for any deviation that may occur in the input liquid stream. This sort of feedback feed-forward combination is a typical setup for the majority of control systems. The feed-forward controller accounts for any upstream disturbances while the feedback controller accounts for the disturbances within the system.



Example 11.2.3

Using the P&ID shown below, write out the feed-forward controller expressions for v3. List any assumptions used in your controller expression.



Solution

To maintain LC1 by measuring FC1 and FC2 and preemptively adjust v3, we must use the controller expression below.

$$v_3 = k_{\text{valves}}(FC1 + FC2)$$

The controller expression above assumes that the valve is linear. Also, it assumes that FC1, FC2, and the model are infinitely accurate.

References

- Svrcek, William Y., Mahoney, Donald P. & Young, Brett R. (2006). *A Real-Time Approach to Process Control*. John Wiley & Sons, Ltd.
- Astrom, Karl J. & Hagglund, Tore (1998). *Advanced PID Control*. Automatic Tuning of PID Controllers. The Instrumentation Systems and Automation Society.
- Smith, Carlos A. & Corripio Armando B. (1997). *Principles and Practice of Automatic Process Control*. John Wiley & Sons, Inc.

Contributors

Authors: Anwar Stephens, Christopher Garcia, Winardi Kusumaatmaja, Meng Yang Ng

This page titled 11.2: Feed Forward Control is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

11.3: Cascade Control

Introduction

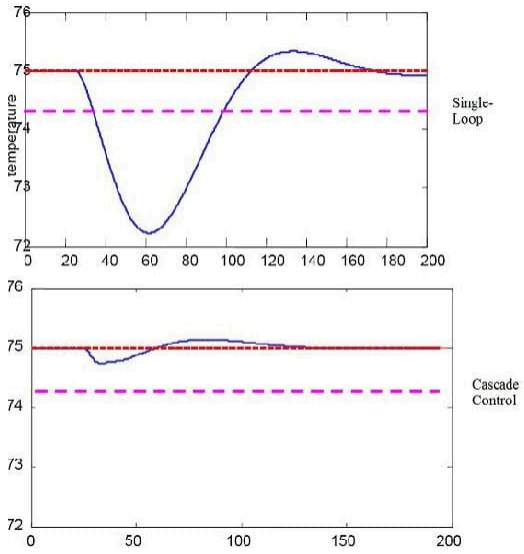
In the previous chapters, only single input, single output (SISO) systems are discussed. SISO involves a single loop control that uses only one measured signal (input). This signal is then compared to a set point of the control variable (output) before being sent to an actuator (i.e. pump or valve) that adjusts accordingly to meet the set point. Cascade controls, in contrast, make use of multiple control loops that involve multiple signals for one manipulated variable. Utilizing cascade controls can allow a system to be more responsive to disturbances.

Before venturing further into the topic of cascade controls, the terms 'manipulated variables', 'measured variables' and 'control variables' should be clarified. The definitions of these terms commonly found in literature are often interchangeable; but, they typically refer to either the input or output signal. For the purpose of this article, 'control variables' will refer to inputs like flow rates, pressure readings, and temperature readings. 'Manipulated variables' and 'measured variables' will refer to the output signals which are sent to the actuator.

The simplest cascade control scheme involves two control loops that use two measurement signals to control one primary variable. In such a control system, the output of the primary controller determines the set point for the secondary controller. The output of the secondary controller is used to adjust the control variable. Generally, the secondary controller changes quickly while the primary controller changes slowly. Once cascade control is implemented, disturbances from rapid changes of the secondary controller will not affect the primary controller.

Cascade Control

To illustrate how cascade control works and why it is used, a typical control system will be analyzed. This control system is one that is used to adjust the amount of steam used to heat up a fluid stream in a heat exchanger. Then an alternative cascade control system for the same process will be developed and compared to the typical single loop control. The figure below shows the performance of cascade control vs. single-loop control in CST heater



Cascade control gives a much better performance because the disturbance in the flow is quickly corrected

Example of Cascade Control

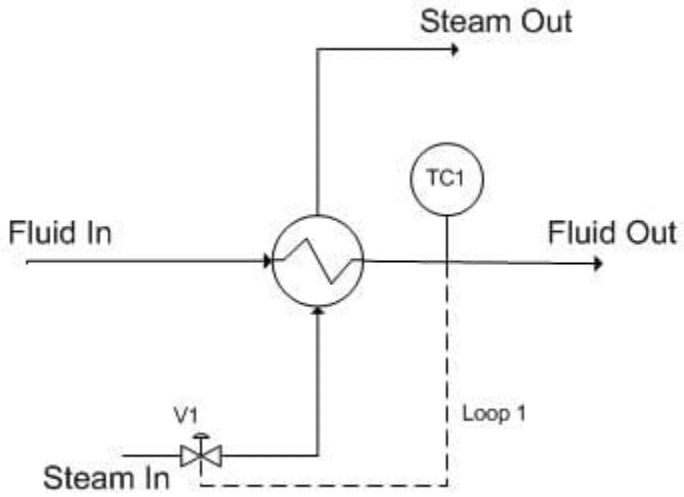


Figure 1. Single loop control for a heat exchanger

In the above process, the fluid is to be heated up to a certain temperature by the steam. This process is controlled by a temperature controller (TC1) which measures the temperature of the exiting fluid and then adjusts the valve (V1) to correct the amount of steam needed by the heat exchanger to maintain the specified temperature. Figure 2 shows the flow of information to and from the temperature controller.

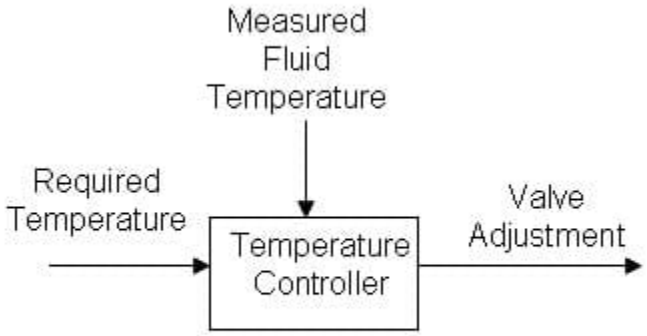


Figure 2. Flow of information when single loop feedback control is used for a heat exchanger

Initially, this process seems sufficient. However, the above control system works on the assumption that a constant flow of steam is available and that the steam to the heat exchanger is solely dependent on opening the valve to varying degrees. If the flow rate of the steam supply changes (i.e. pipeline leakage, clogging, drop in boiler power), the controller will not be aware of it. The controller opens the valve to the same degree expecting to get a certain flow rate of steam but will in fact be getting less than expected. The single loop control system will be unable to effectively maintain the fluid at the required temperature.

Implementing cascade control will allow us to correct for fluctuations in the flow rate of the steam going into the heat exchanger as an inner part of a grander scheme to control the temperature of the process fluid coming out of the heat exchanger. A basic cascade control uses two control loops; in the case presented below (see Figure 3), one loop (the outer loop, or master loop, or primary loop) consists of TC1 reading the fluid out temperature, comparing it to $TC1_{set}$ (which will not change in this example) and changing $FC1_{set}$ accordingly. The other loop (the inner loop, or slave loop, or secondary loop) consists of FC1 reading the

steam flow, comparing it to $FC1_{set}$ (which is controlled by the outer loop as explained above), and changing the valve opening as necessary.

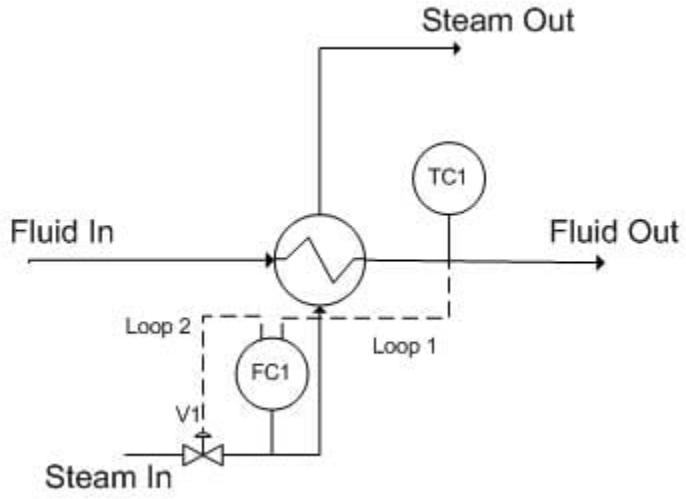


Figure 3. Cascade control for a heat exchanger

The main reason to use cascade control in this system is that the temperature has to be maintained at a specific value. The valve position does not directly affect the temperature (consider an upset in the stream input; the flow rate will be lower at the same valve setting). Thus, the **steam flow rate** is the variable that is required to maintain the process temperature.

The inner loop is chosen to be the inner loop because it is prone to higher frequency variation. The rationale behind this example is that the steam in flow can fluctuate, and if this happens, the flow measured by FC1 will change faster than the temperature measured by TC1, since it will take a finite amount of time for heat transfer to occur through the heat exchanger. Since the steam flow measured by FC1 changes at higher frequency, we chose this to be the inner loop. This way, FC1 can control the fluctuations in flow by opening and closing the valve, and TC1 can control the fluctuations in temperature by increasing or decreasing $FC1_{set}$.

Thus, the cascade control uses two inputs to control the valve and allows the system to adjust to both variable fluid flow and steam flow rates. The flow of information is shown in figure 4.

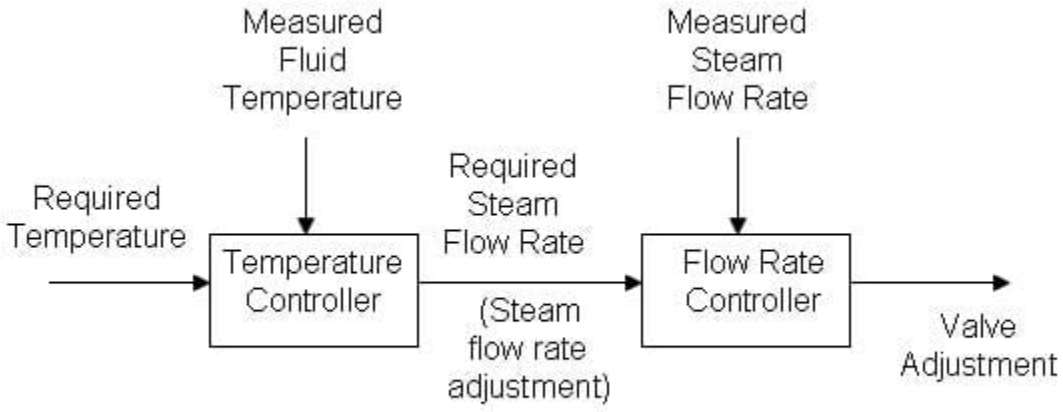


Figure 4. Flow of information when cascade control is used for a heat exchanger

In order to accomplish this, relationships between the primary and secondary loops (see definitions below) must be defined. Generally, the primary loop is a function of the secondary loop. A possible example of such relations is:

$$TC_1 = f(FC_1);\ \frac{d(TC_1)}{d(FC_1)} = \ldots$$

$$FC_1 = f(V_1);\ \frac{d(FC_1)}{d(V_1)} = \ldots$$

$$\frac{d(V_1)}{d(t)} = \ldots$$

Primary and Secondary Loops

In Figure 3, there are two separate loops. Loop 1 is known as the primary loop, outer loop, or the master, whereas loop 2 is known as the secondary loop, inner loop, or the slave. To identify the primary and secondary loops, one must identify the control variable and the manipulated variable. In this case, the control variable is the temperature and the reference variable is the steam flow rate. Hence, the primary loop (loop 1) involves the control variable and the secondary loop (loop 2) involves the reference variable. The information flow for a two loop cascade control system will typically be as shown in Figure 5. Please note that the user sets the set point for loop 1 while the primary controller sets the set point for loop 2.

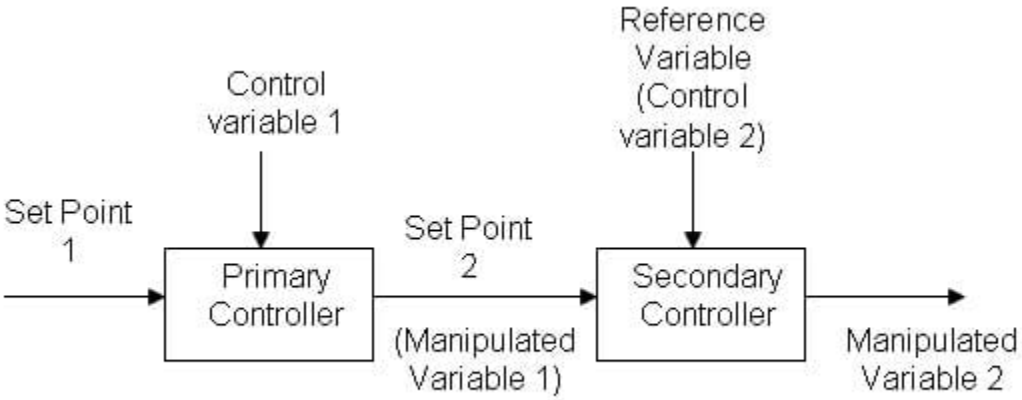
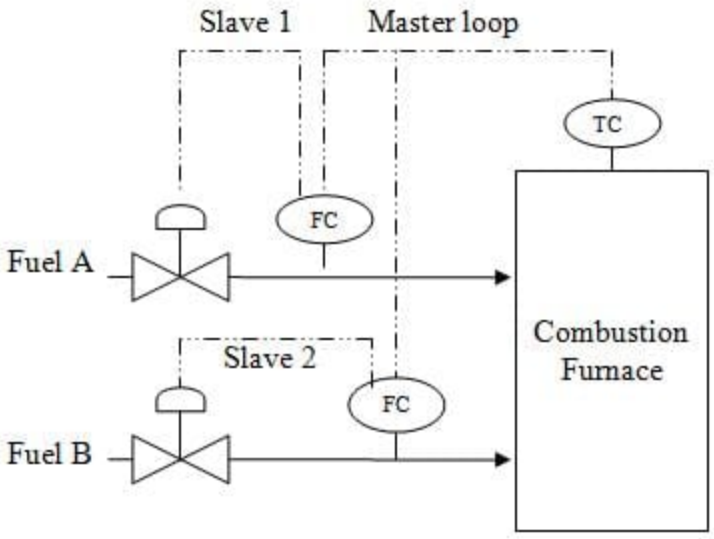


Figure 5: Information flow of a two loop cascade control

In addition to this common architecture, cascade control can have multiple secondary loops; however, there is still one primary loop and a main controlled variable. Unfortunately, with multiple inner loops, tuning the PID becomes even more challenging, making this type of cascade less common. The secondary loops can be either independent of each other, or dependent on each other, in which case each secondary loop affects the set point of the other secondary loop. When tuning such controller, the inner most loop should be tuned first. The loop that manipulates the set point of the inner-most loop should be tuned next and so for. The figure below shows an example of using two secondary loops, independent of each other, in a fuel combustion plant. In this combustion furnace, the master controller controls the temperature in the furnace by changing the set point for the flow of fuels A and B. The secondary loops correspond to the change in the set point for the flow, by opening or closing the valves for each fuel.

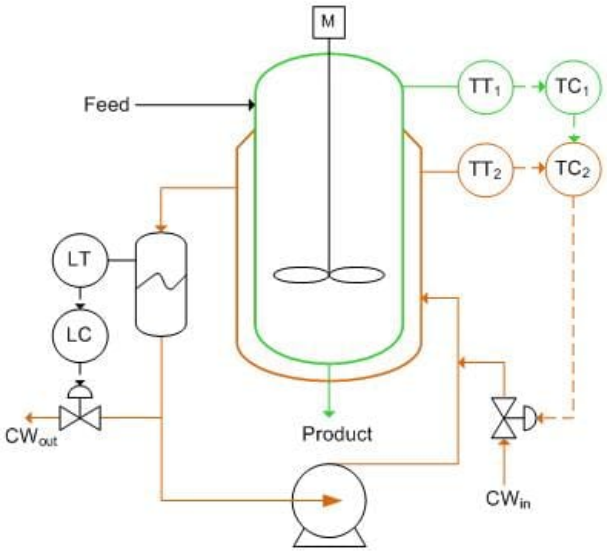


Cascade control is generally useful when

- A system error affects the primary control variable only after a long period of time as it propagates through dead time and lag time.
- A system has long dead times and long lag times.
- Multiple measurements with only one control variable are required for better response to a disturbance of a system.
- Variance occurs in multiple streams

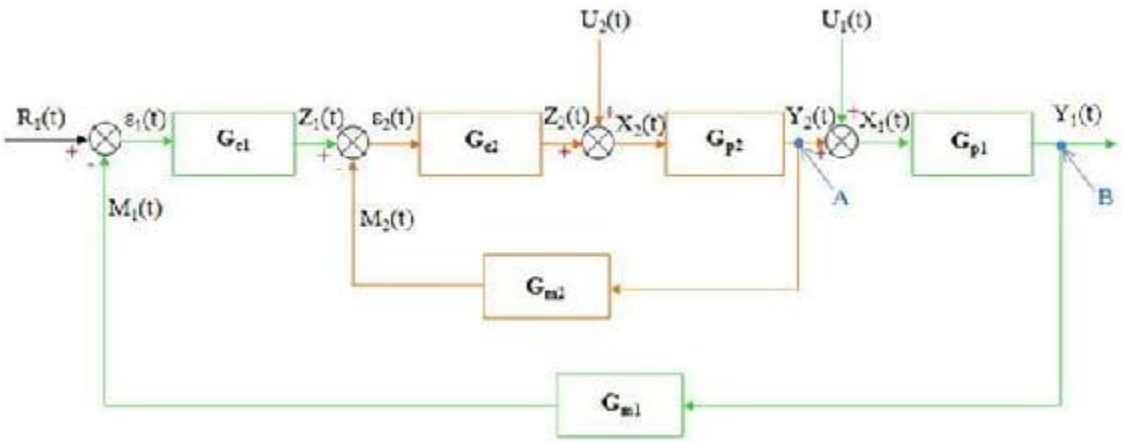
General Cascade Control Schematic

The reactor below needs to be cooled during continuous-feed operation of an exothermic reaction. The reactor has been equipped with a cooling water jacket with the water flow rate being controlled by cold water valve. This valve is controlled by two separate temperature controllers. An “inner-loop” or “slave” (highlighted in orange) temperature transmitter communicates to the slave controller the measurement of the temperature of the jacket. The “outer-loop” or “master” (green) temperature controller uses a master temperature transmitter to measure the temperature of the product within the reactor. The output from the slave controller is fed into the master controller and used to adjust the cold water valve accordingly.



The cascade control loop used to control the reactor temperature can be generalized with the schematic below. We will use this

main diagram to go through the formal derivation of the equations describing the behavior of the system when there are changes in the loads U1 and U2 but with no change in the set point, R1(t). The general equations derived below can be used to model any type of process (first, second, third order differential equations, ect.) and use any type of control mechanism (proportional-only, PI, PD, or PID control). See wiki pages titled “first-order differential equations” and “P, I, D, PI, PD, PID Control” for more details.



Step 1: Write down all the equations for each stage of the control loop

Master Loop

$$Y_1(t) = \hat{G}_{P1} X_1(t)$$

$$X_1(t) = U_1(t) + Y_2(t)$$

$$Z_1(t) = \hat{G}_{C1} \varepsilon_1(t)$$

$$\varepsilon_1(t) = R_1(t) - M_1(t)$$

$$M_1(t) = \hat{G}_{M1} Y_1(t)$$

Slave Loop

$$Y_2(t) = \hat{G}_{P2} X_2(t)$$

$$X_2(t) = U_2(t) + Z_2(t)$$

$$Z_2(t) = \hat{G}_{C2} \varepsilon_2(t)$$

$$\varepsilon_2(t) = Z_1(t) - M_2(t)$$

$$M_2(t) = \hat{G}_{M2} Y_2(t)$$

Gp1 and Gp2 are the process operators and are usually of the form:

$$\hat{G}_p^{-1} = \left(\tau_p \frac{d}{dt} + 1\right)^n$$

Where n is a natural number.

Gc1 and Gc2 are the control operators and depend on the type of controller used. For PID controllers, they would be:

$$\hat{G}_C = K_C\left(1+\frac{1}{\tau_I}\int_0^t dt' + \tau_D\frac{d}{dt}\right)$$

Gm1 and Gm2 are the measurement operators and usually are just equal to 1. Note that there are no equations for the "intersections" A and B shown on the diagram.

Step 2: Simplify the equations for the slave loop

$$Y_2(t) = \hat{G}_{P2}X_2(t)$$

$$Y_2(t) = \hat{G}_{P2}\{U_2(t) - Z_2(t)\}$$

$$Y_2(t) = \hat{G}_{P2}\left[U_2(t) - \hat{G}_{C2}\varepsilon_2(t)\right]$$

$$Y_2(t) = \hat{G}_{P2}\left[U_2(t) - \hat{G}_{C2}\left(Z_1(t) - M_2(t)\right)\right]$$

$$Y_2(t) = \hat{G}_{P2}\left[U_2(t) - \hat{G}_{C2}\left(Z_1(t) - \hat{G}_{M2}Y_2(t)\right)\right]$$

Solve for Y2(t)

$$Y_2(t) = \left(\frac{\hat{G}_{P2}}{1+\hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2}}U_2(t) + \frac{\hat{G}_{P2}\hat{G}_{C2}}{1+\hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2}}Z_1(t)\right)$$

Step 3: Simplify the equations for the master loop

$$Y_1(t) = \hat{G}_{P1}X_1(t)$$

$$Y_1(t) = \hat{G}_{P1}\{U_1(t) + Y_2(t)\}$$

$$Y_1(t) = \hat{G}_{P1}\left[U_1(t) + \frac{\hat{G}_{P2}}{1+\hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2}}U_2(t) + \frac{\hat{G}_{P2}\hat{G}_{C2}}{1+\hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2}}Z_1(t)\right]$$

$$Y_1(t) = \hat{G}_{P1}\left[U_1(t) + \frac{\hat{G}_{P2}}{1+\hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2}}U_2(t) + \frac{\hat{G}_{P2}\hat{G}_{C2}}{1+\hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2}}\hat{G}_{C1}\varepsilon_1(t)\right]$$

$$Y_1(t) = \hat{G}_{P1}\left[U_1(t) + \frac{\hat{G}_{P2}}{1+\hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2}}U_2(t) + \frac{\hat{G}_{P2}\hat{G}_{C2}}{1+\hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2}}\hat{G}_{C1}\left(R_1(t) - M_1(t)\right)\right]$$

$$Y_1(t) = \hat{G}_{P1}\left[U_1(t) + \frac{\hat{G}_{P2}}{1+\hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2}}U_2(t) + \frac{\hat{G}_{P2}\hat{G}_{C2}}{1+\hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2}}\hat{G}_{C1}\left(R_1(t) - \hat{G}_{M1}Y_1(t)\right)\right]$$

Solve for Y1(t)

$$Y_1 = \left(\frac{\left(\hat{G}_{P1} + \hat{G}_{P1}\hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2}\right)U_1(t) + \hat{G}_{P1}\hat{G}_{P2}U_2(t) + \hat{G}_{P1}\hat{G}_{P2}\hat{G}_{C1}\hat{G}_{C2}R_1(t)}{1 + \hat{G}_{P2}\hat{G}_{C2}\hat{G}_{M2} + \hat{G}_{P1}\hat{G}_{P2}\hat{G}_{M1}\hat{G}_{C1}\hat{G}_{C2}} \right)$$

(Note: here the G's are written as operators rather than Laplace transforms, and as such they shouldn't be divided. Thus, expressions in the denominator should be interpreted as inverse operators in the numerator.)

Conditions for Cascade Control

In order to have a smooth flow of information throughout the control system, a hierarchy of information must be maintained. In a double loop cascade system, the action of the secondary loop on the process should be faster than that of the primary loop. This ensures that the changes made by the primary output will be reflected quickly in the process and observed when the primary control variable is next measured. This hierarchy of information can be preserved by applying the following conditions when setting up the cascade controls.

1. There must be a clear relationship between the measured variables of the primary and secondary loops.
2. The secondary loop must have influence over the primary loop.
3. Response period of the primary loop has to be at least 4 times larger than the response period of the secondary loop.
4. The major disturbance to the system should act in the primary loop.
5. The primary loop should be able to have a large gain, Kc.

Cascade control is best when the inner loop is controlling something that happens at fairly high frequency. Cascade control is designed to allow the master controller to respond to slow changes in the system, while the slave controller controls disturbances that happen quickly. If set up in reverse order, there will be a large propagation of error. Hence, it is important to maintain the hierarchy of information. In summary, the master controller responds to SLOW changes in the system, while the slave controller responds to the high frequency, or FAST changes in the system. This also requires that the inner control scheme be tuned TIGHTLY so error is not allowed to build. Commonly, the inner loops controls a flow controller, which will reduce the effect of changes such as fluctuations in steam pressure.

Cascade Control Design Considerations

Open Loop and Closed Loop

Although cascade control generally incorporates information from several secondary loops, the system overall does not automatically become a closed loop or an open loop process. For a cascade control to be open loop, all control loops in the system should be open loop in nature. On the other hand, if any of the control loops in the system is feedback-based, the system overall is considered as closed loop. This is due to the fact that the system is getting some sort of feedback, no matter how little the “fedback” variable influences the system.

Cascade with Feed-Forward and Feedback Controls

Cascade controllers have a distinct advantage over other kinds of controller due to its ability to combine both feedback and feed-forward controls. While feed-forward loops have the potential to adjust the controlled variables to the ideal states, the feedback loops in a mixed system check deviations to make sure the system is on track. The figure below is an example of a mixed cascade control:

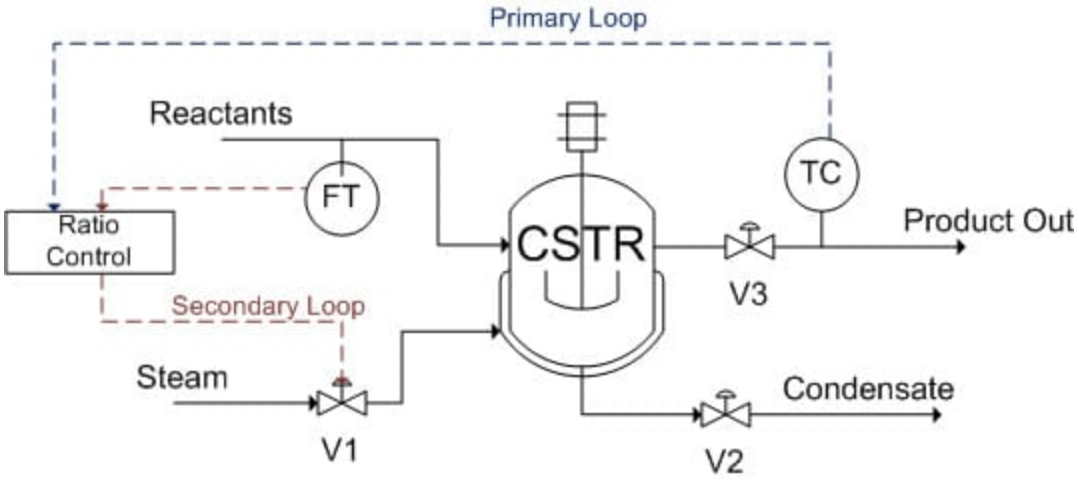


Figure 6. Cascade control with both feed-forward and feedback controls

An endothermic reaction takes place in a jacketed CSTR, in which the reactor is heated by steam. A reactant stream feeds into the CSTR, which serves as the wild stream for the ratio controller to predict a required steam flow rate (feed-forward). On the product stream side, a temperature controller manipulates the ratio setting of the ratio controller in order to minimize the product temperature errors (feedback). The temperature controller is the primary loop, whereas the ratio controller is the secondary loop.

Advantages and Disadvantages of Cascade Control

The following table shows a list of cascade control pros and cons:

Cascade Control Pros	Cascade control Cons
• Accounts for disturbances in the primary vairable more quickly and hence control the primary variable more effectively • Reduces the effects of dead time and phase lag time in the system • Can be combined with feed-forward control or other forms of control • Integrated multiple sensors readings • Improve dynamic performance and provide limits on the secondary variables	• Cascade control makes the system more complex • Cascade control requires more equipment and instrument that will drive up the cost of the process • Tuning cascade controllers is more difficult as the set point changed + more parameters

Starting up a Cascade System

A cascade system needs to be set up properly in order to function. The inner loop should be tuned before the outer loop. The following are the suggested steps for starting up a cascade system (both controllers start in automatic mode):

1. Place the primary controller in manual mode. This will break the cascade and isolate the secondary controller so that it can be tuned.
2. Tune the secondary controller as if it were the only control loop present.
3. Return the secondary controller to the remote set point and/or place the primary controller in automatic mode. This will isolate the primary controller so that it can be tuned.

4. Tune the primary control loop by manipulating the set point to the secondary controller. If the system begins to oscillate when the primary controller is placed in automatic, reduce the primary controller gain.

Cascade control systems can be tuned using conventional methods. Ziegler-Nichols tuning method can be used to tune the secondary controller. Then the parameters (depending on the system, it can be P, PI, or PID) need to be fine tuned by introducing disturbances to the system and changing the parameters accordingly. The secondary controller must be tuned tightly (no oscillations when disturbance is added), otherwise the primary controller will be responding to oscillations due to the secondary controller as well as from disturbances to the system. This will cause the parameters for the primary controller to be inaccurate, which will cause a high controller effort. This is undesirable because it will wear down the controller. The primary controller often utilizes internal model control (IMC), which allows for improved performance by incorporating the process model into the controller setup. Another way to tune the primary controller is by using trial and error. For example, if the system has a PID primary controller, the integral and derivative gains should be set at small values. Then a small proportional gain should be introduced. The proportional gain should be tuned first then the integral and derivative gains can be added. For more information about the trial and error method refer to PID Tuning Classical section.

Startup Example

A simple example of starting up a cascade system is shown using the heat exchange system seen in Figure 3. To begin startup the temperature controller (primary) is set to manual mode. The flow controller (secondary) is then tuned by adjusting the set point of the flow controller. This is shown in Figure 6 below.

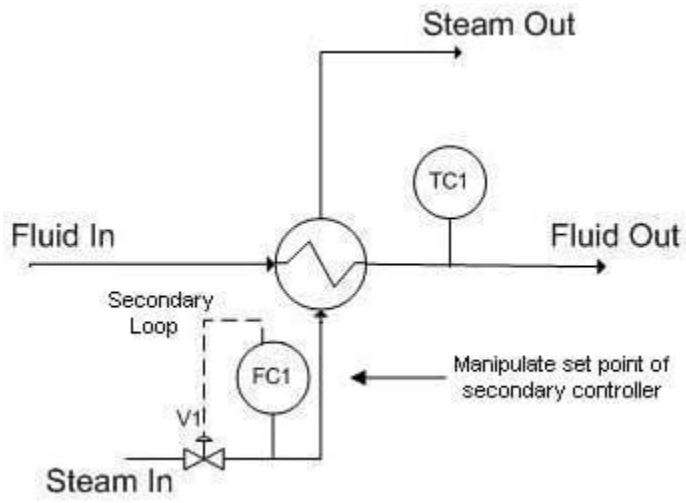


Figure 6: Automatic mode for secondary controller (manual mode for primary controller)

Now the temperature controller is set to automatic mode and the flow controller is set to manual mode. In this system the temperature controller outputs a set point to the flow controller just like it would in regular cascade mode. The temperature controller is then tuned by adjusting the output to the valve. This is shown in Figure 7 below.

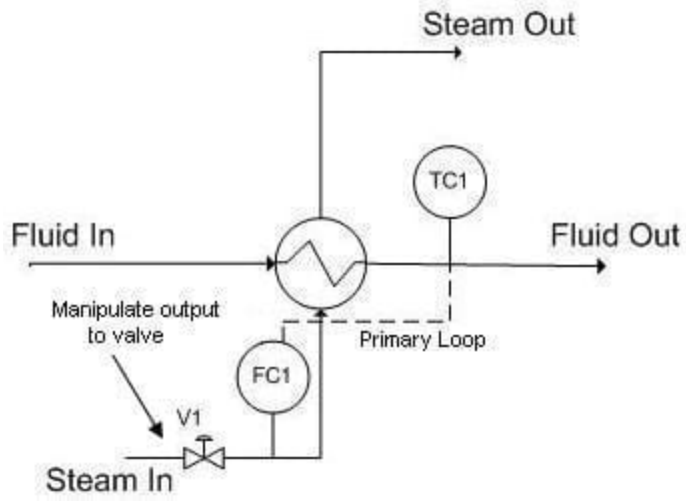


Figure 7: Manual mode for secondary controller (automatic mode for primary controller)

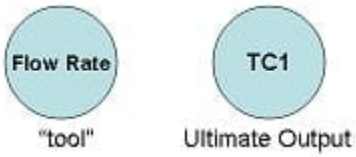
Developing the Structure of a Cascade Algorithm

Below is a step-by-step method through which the basic structure of a Cascade Algorithm can be developed. The examples in each step refer to Figure 3.

1 – **Determine the goal of your algorithm**. That is, which **ultimate output** you would like to ultimately end up changing. Also determine what **"tool"** (or aspect of the system) will most directly and physically allow you to accomplish this desired change.

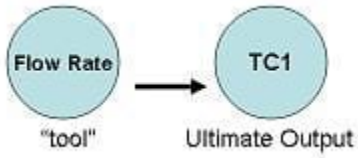
This "tool" can be any physical property of the system: temperature, pressure, surface area, flow regime, flow rate, etc. of any part of any physical component of the system. The "tool" doesn't need to be something read by a sensor, though it can be (as it is in this example, the temperature of the heated fluid leaving the heat exchanger is read by sensor TC1).

For instance, the end goal of the process of Figure 3 is to regulate the temperature of the fluid leaving a heat exchanger. In this case, the only way to physically accomplish this is to change the flow rate of the steam through the heat exchanger (this flow rate just so happens to be read by the sensor FC1).



2 – **Determine how the "tool" physically affects the ultimate output**.

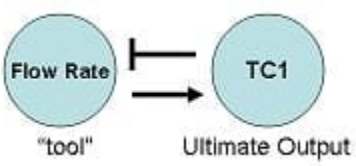
For instance, an **increased** flow rate of **steam** (read by FC1) will physically **increase** the **temperature** of the fluid leaving the heat exchanger, **TC1**.



3 – **Determine how the ultimate output needs to affect the "tool" to achieve the desired ultimate output change**.

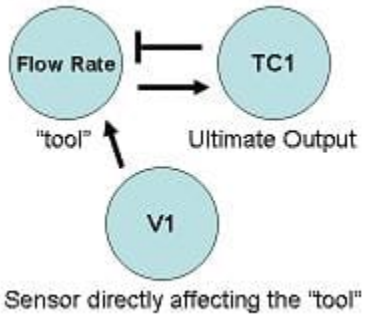
For instance, we presume that we would like to, in general, resist any change in TC1 (i.e. we don't want an increase in TC1 to further increase the value of TC1, that would be bad!). For this reason, we want an increase in TC1 to decrease FC1, which will then

decrease TC1 (negative feedback).



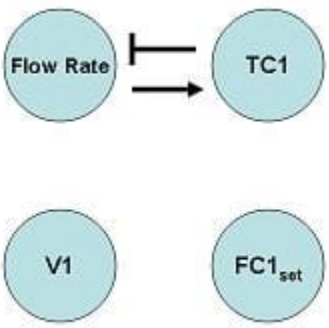
4 – **Determine which valve or sensor's output will** most **directly affect the "tool".** Also, determine how the sensor and "tool" are related**.**

For instance, the setting of valve V1 will most directly affect the flow rate of steam entering the heat exchanger (it is the "closest" controllable component of the process to the "tool", the steam flow rate). An increase in V1 will increase flow rate.

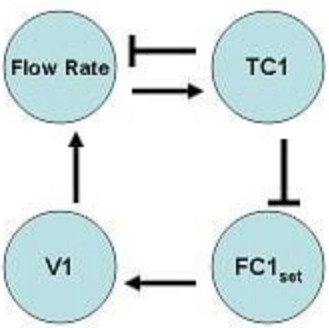


5 – **Determine all valves/sensors in between the ultimate output and valve/sensor that will directly affect the "tool"**.

For instance, between the ultimate output (TC1) and the valve (V1) that will directly affect the "tool", there is only the FC1 sensor.



6 – **Make an incidence graph**. Make sure that you use the **set points** of the valves/sensors that directly affect the "tool", because you cannot control what these sensors read (*you can only control their set points*). Confirm that the relationship between the "tool" and the ultimate output is consistent with the incidence graph!



7 – Use this incidence graph to **construct the algorithm**. The incidence graph is particularly helpful in determining what "sign" changes in one component should result in a the "sign" in another component. If we use incidence diagrams, and the equations for the cascade controller, we can determine a relationship between the components of the system.

For instance, if TC1 were to increase, we would want the steam flow rate to decrease, which ultimately results in a decreased temperature. An increase (a "sign" of +) in TC1 should result in a decrease in $FC1_{set}$ (a "sign of -). This can be seen from equation 2, where if $TC1_{set}$ decreases, the difference between $TC1_{set}$ and TC1 will be smaller. This causes $FC1_{set}$ to decrease relative to its previous set point. A decrease (a "sign" of -) in $FC1_{set}$ should result in a decrease in V1 (a "sign of -). This is shown by equation 1, where if $FC1_{set}$ decreases while FC1 is the same, V1 will decrease relative to its previous position. A decrease (a "sign" of -) in V1 should result in a decreased flow rate (a "sign of -). A decrease (a "sign" of -) in steam flow rate should result in a decreased TC1 (a "sign of -), completing the negative-feedback mechanism initially desired.

$$V_1 - V_{1,Offset} + K_{c,FC1}(FC1_{set} - FC1) + \frac{1}{\tau_{i,FC1}} \int (FC1_{set} - FC1)dt + \tau_{d,FC1} \frac{d(FC1_{set} - FC1)}{dt}$$

(1)

$$FC1_{set} - FC1_{Offset} + K_{c,TC1}(TC1_{set} - TC1) - \frac{1}{\tau_{i,TC1}} \int (TC1_{set} - TC1)dt + \tau_{d,TC1} \frac{d(TC1_{set} - TC1)}{dt}$$

(2)

Failure

A cascade system is not to be confused with a fail-safe system. The sole purpose of implementing a cascade control system is to make the system more responsive, **not** more robust.

In a cascade system, if the master controller fails, the entire system will fail too. Just like for any other design, one should anticipate failure and build in redundancy. An example of redundancy could be having multiple controllers to be the master controller and then using selectors to choose which reading to use.

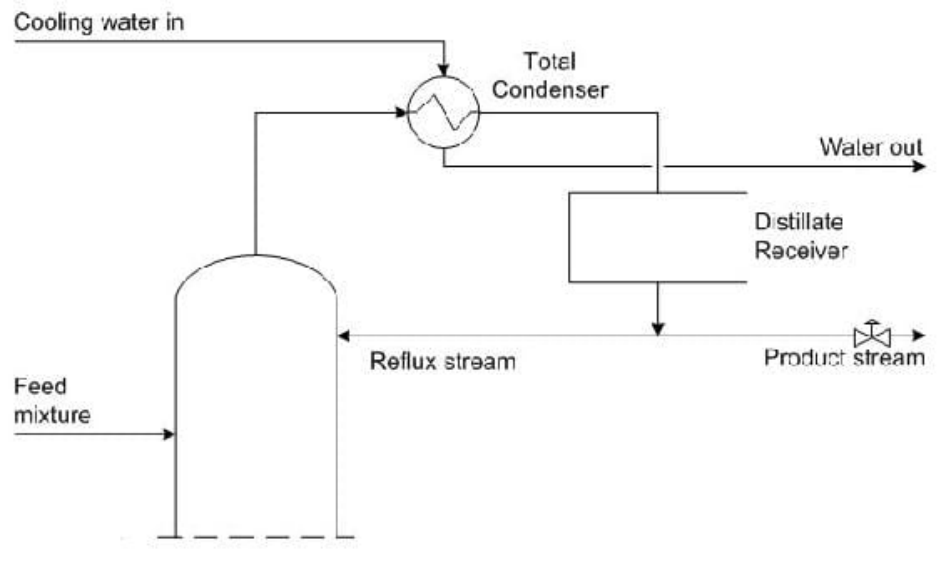
Alternatively, if the cascade system fails and has no built-in redundancy, there are a couple of ways to keep the cascade system running while the controller is being repaired. For example, the controller can be operated manually by an employee, or an average of previous readings can be used to temporarily send a constant signal to the system.

Example 11.3.1

The diagram below shows the top portion of a distillation column used to separate the components of the inlet liquid mixture. Reflux is used for the column to improve the separation efficiency, as the desired product stream must have a mass fraction of component C that is greater than 0.8. At the top of the distillation column the temperature is 200°C. Components A, B, and C, which are gases at that temperature, accumulate and flow out of the top of the distillation column through the total condenser. The total condenser cools the gas stream to 150°C, causing A, B, and C to liquefy. The liquid mixture is collected in a distillate receiver. Finally, the liquid stream exits the receiver and is either collected as product or sent back to the

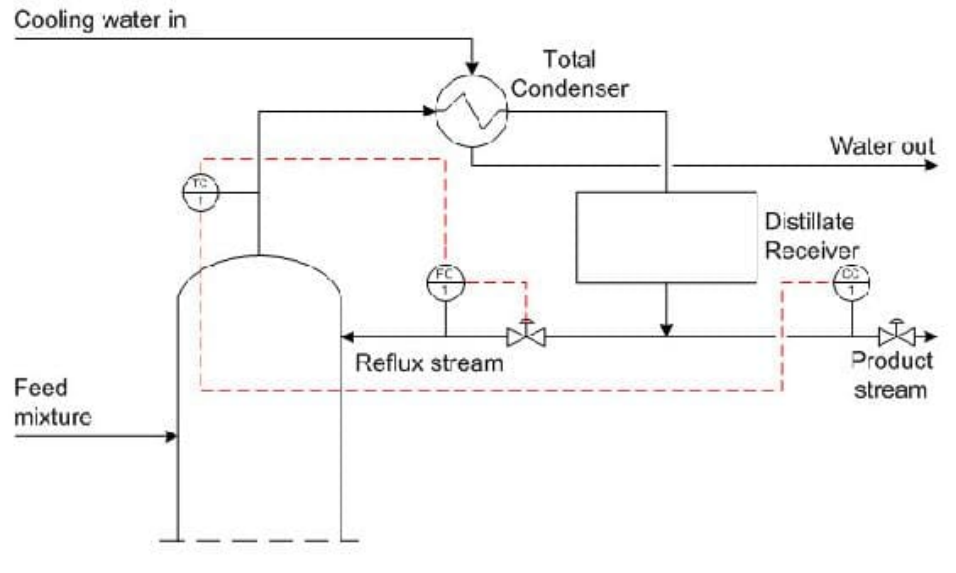
column. A control valve is in place on the product stream to ensure that the necessary purity is attained.

On the diagram, draw the necessary controllers and valves required to implement a cascade control for the exiting stream. Also, describe the cascade control that you have implemented.



Note that this is only one possible solution to the example question. Alternatives to this solution would be to only use a composition controller with a flow controller or a temperature controller but the below solution was chosen to give an example of a multiple cascade control.

Solution



The system can most effectively be controlled with a cascade scheme. A composition controller can be installed to control the composition of the product stream. However, the composition controllers are typically slow to react since it takes time to analyze the samples to obtain compositions. In order to correct for the time lag of the composition controller, a temperature controller, located at the top of the distillation column, measures the temperature of the product stream giving a rough idea of the product composition. Additionally, a flow controller is needed on the reflux stream.

The input from the composition controller measures the composition of the product stream and sends a signal to the temperature controller at the top of the column. The temperature controller compares the set point to the measured temperature of the stream and outputs a signal to the flow controller on the recycle stream. The flow controller compares this set point to the measured flow rate and adjust the control valve on the recycle stream accordingly. This cascade control allows the system to react to disturbances quickly and to more effectively control the reflux process.

Example 11.3.2

You just graduated from the University of Michigan and escaped the cold winter of Ann Arbor for a sweet full-time position in Orange County. Your company manufactures a specialty drug to cure an epidemic in the penguin population in Antarctica.

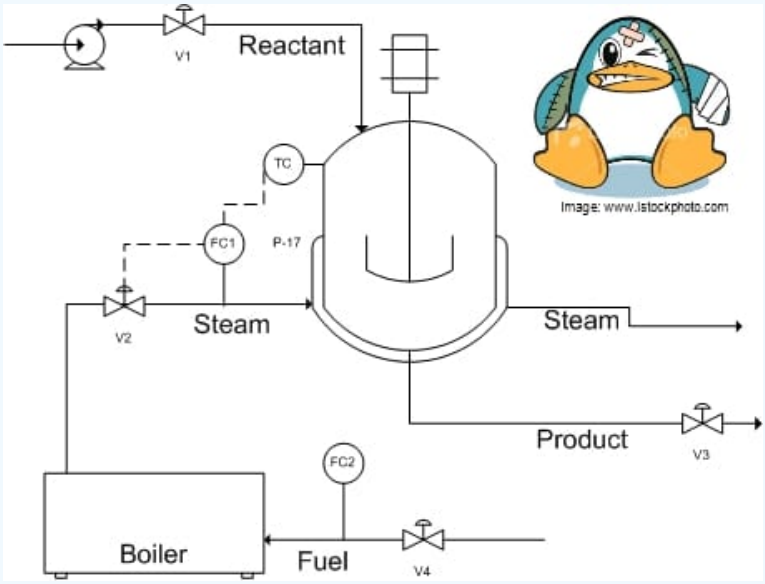
The endothermic reaction A + B --> C is run in a CSTR. A and B are toxic in even dilute concentrations. Therefore, a conversion above 99.95% has to be maintained in order to ensure the safety of the process output.

Steam from a boiler is utilized to maintain the optimum reaction temperature of 100C. A PID cascade control scheme is used to maintain this temperature.

This morning, your plant manager came to you with a problem. He explained that production has been delayed because the conversion dropped below the acceptable level. He noticed that the reactor temperature was well below 100C, even though the steam valve (V2) is open completely.

As can be seen in the diagram below, a cascade control is already present in the system, but it isn't set up properly. Can you fix the problem and save the penguins?

Hint: The boiler also supplies steam to other equipment. You can increase its steam output by adding more fuel to the boiler (assume that water supply is unlimited).



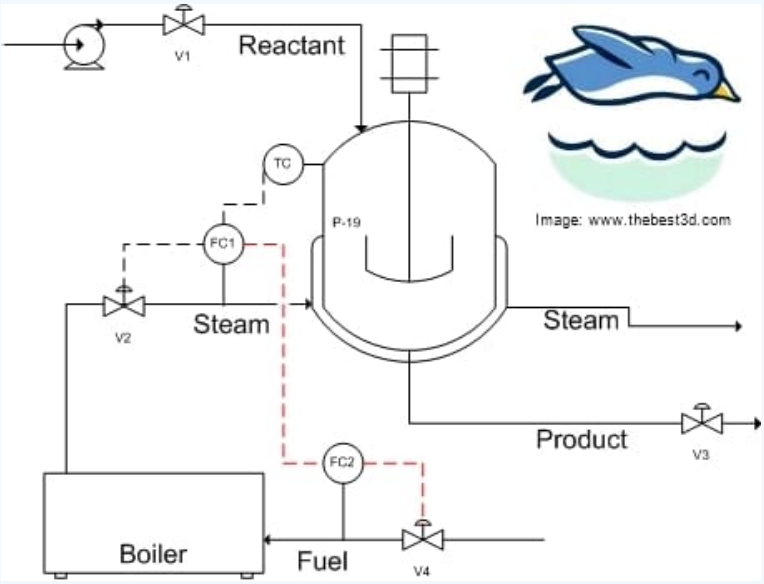
Solution

Before: TC is cascaded with FC1, which controls V2. (TC-->FC1-->V2) Even when V2 is open all the way, there is not enough steam to maintain the reaction at optimum temperature.

In this situation the controller surpasses the saturation limit of the actuator (V2). In other words, the controller is demanding more steam from the boiler than it can supply. A likely cause of this is a disturbance in the steam supply from the boiler because of increased steam demand from other processes. A possible solution is to increase the steam output from the boiler by adding more fuel (controlled by FC2 and V4). A side effect of this control situation would be that the I- part of a PID controller would also windup, preventing the system from responding.

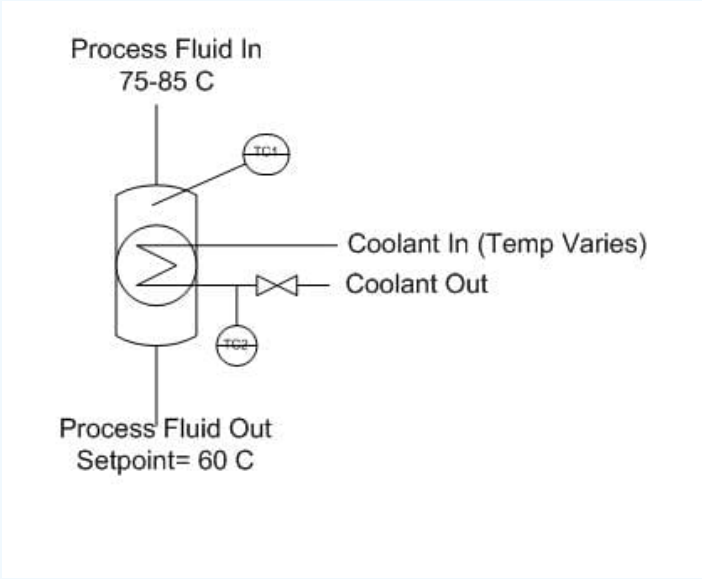
An additional control loop can be added to enable the boiler output to communicate with the fuel input. In addition to controlling V2, FC1 also controls FC2, which in turn will control V4. This additional controller is shown in the figure below.

Congratulations, you just saved the penguins!



Example 11.3.3

You have a holding tank for a warm process fluid that is cooled by a water-filled jacket. The process stream enters at 75-85 C and should be cooled to 60 C. The temperature of the cooling water also varies, since it comes from another process, but its average temperature is about 29 C. You have thermocouples inside the holding tank and on the cooling water jacket exit stream. Unfortunately, the tank is not well mixed, so the thermocouple reading inside the tank lags behind the actual temperature of the outlet stream. When the cooling jacket is first turned on, the tank contents are initially at 70 C. Examine the dynamics of a feedback control stream versus a cascade control scheme for this system.



Solution

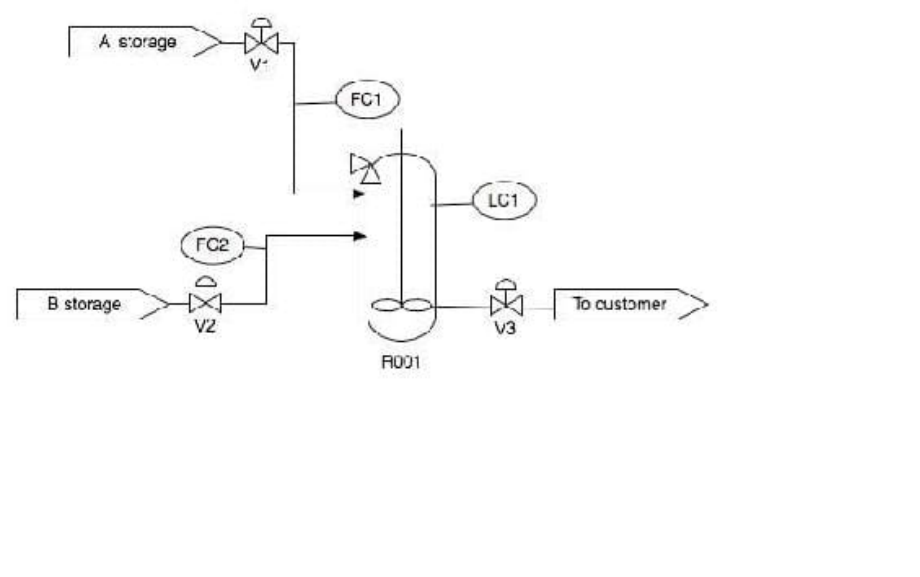
One setup for this system is given in the following Excel file: Media:Cascade.xls

Under the feedback control scheme, the flow rate of coolant is adjusted based on the temperature inside the holding tank. The system overshoots setpoints more significantly, because of the lag time before disturbances are recognized.

In cascade control, the temperature sensor in the process fluid outputs a setpoint for the jacket outlet (based on a knowledge of the process fluid setpoint and the average temperature of the cooling fluid). The signal from the temperature controller on the jacket outlet controls the flow rate of cooling fluid. Since the temperature in the cooling fluid has a shorter time lag, the process remains closer to its setpoint.

Example 11.3.4

We have a P&ID as shown below. Write out the equations necessary to connect LC1 to FC1 to adjust v1. Which equation is the slave and which one is the master?



Solution

The first equation is the slave and the second equation is the master. We have to use the $FC1_{set}$ to determine what V_1 should be set as. And these are both based on the level in the tank.

$$V_1 = V_{1,offset} + K_c2(FC1_{set} - FC1) + \frac{1}{\tau_I 2}\int(FC1_{set} - FC1)dt - \tau_D 1\frac{d(FC1_{set} - FC1)}{dt}$$

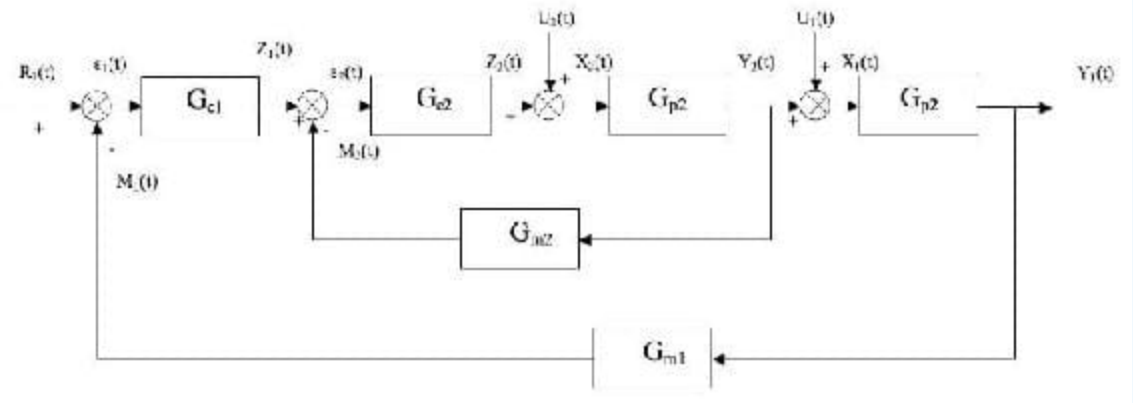
(1)

$$FC1_{set} = FC1_{offset} + K_c1(LC1_{set} - LC1) + \frac{1}{\tau_I 2}\int(LC1_{set} - LC1)dt + \tau_D 2\frac{d(LC1_{set} - LC1)}{dt}$$

(2)

Example 11.3.5

Write the general equation that describes the behavior of the cascade control system shown below. Include changes in U_1 and U_2.



Hint: Start by writing the inner loop and outer loop equations separately.

Solution

For inner cycle:

$\varepsilon_1 = Z_1 + G_{M2}Y_2$

$Z_2 = Z_1G_{C2} - Y_2G_{C2}G_{M2}$

$X_2 = U_2 + Z_2 = U_2 + Z_1G_{C2} - Y_2G_{C2}G_{M2}$

$Y_2 = G_{P2}X_2 = U_2G_{P2} + Z_1G_{P1}G_{C2} - Y_2G_{C2}G_{M2}G_{P2}$

$$Y_2 = \frac{U_2G_{P2}}{1 + G_{C2}G_{M2}G_{P2}} + \frac{G_{C2}G_{P2}}{1 + G_{C2}G_{M2}G_{P2}}Z_1$$

For outer cycle:

$Z_1 = R_1G_{C1} - Y_1G_{M1}G_{C1}$

$Y_1 = U_1G_{P1} + Y_2G_{P1}$

Combining Z_1 from outer cycle with Y_2 from inner cycle:

$$Y_2 = \frac{U_2G_{P2}}{1 + G_{C2}G_{M2}G_{P2}} + \frac{R_1G_{C1}G_{C2}G_{P2}}{1 + G_{C2}G_{M2}G_{P2}} - \frac{Y_1G_{C1}G_{C2}G_{M1}G_{P2}}{1 + G_{C2}G_{M2}G_{P2}}$$

Combining Y_2 above with Y_1 from outer cycle:

$$Y_1 = U_1G_{P1} + \frac{U_2G_{P1}G_{P2}}{1 + G_{C2}G_{M2}G_{P2}} + \frac{R_1G_{C1}G_{C2}G_{P1}G_{P2}}{1 + G_{C2}G_{M2}G_{P2}} - Y_1\frac{G_{C1}G_{C2}G_{M1}G_{P1}G_{P2}}{1 + G_{C2}G_{M2}G_{P2}}$$

Rearranging and solving explicitly for Y_1:

$$Y_1 = \frac{U_1G_{P1} + U_1G_{C2}G_{M2}G_{P1}G_{P2} + U_2G_{P1}G_{P2} + R_1G_{C1}G_{C2}G_{P1}G_{P2}}{1 + G_{C2}G_{M2}G_{P2} + G_{C1}G_{C2}G_{M1}G_{P1}G_{P2}}$$

Practice Quiz

Answer the five questions below with the most appropriate solution. There is one correct answer for each question. Correct answers are awarded 4 points, whereas a loss of 2 points occurs for incorrect answers. Good luck!

1. Which of the following is never true in a process controlled by a cascade control scheme?

a) Changes in the secondary control variable affect the primary measured variable.
b) Disturbances in the secondary loops propagate at least four times faster than disturbances in the primary loop.
c) One control variable is used to influence the primary and secondary measured variables.
d) Major disturbances in the process are detected in the primary loop.

2. Which of the following in not an advantage of cascade control?

a) The system reacts to disturbances more quickly.
b) More efficient control of the primary variable.
c) It is a cheap control scheme to implement.
d) The effects of dead time and phase lag time are reduced in the system.

3. True or False: During start-up of a cascade control system, after the primary and secondary controllers are turned and placed in automatic mode, the primary controller gain should be increased to reduce system oscillations.

4. The primary loop, or outer loop, involves which of the following variables?

a) Control variable
b) Temperature variable
c) Reference variable
d) I love penguins (variable)

5. True or False: The primary loop is generally a function of the secondary loop.

Answers

Based on your selections, calculate your score on this quiz.

1. C- Cascade control schemes use multiple control variables for one manipulated variable.

2. C- Increased amounts of equipment and instrumentation may drive up the cost of the process.

3. False- The controller gain (Kc) should be reduced to reduce system oscillations

4. A- The best answer is (a) since the control variable may or may not be temperature. The reference variable is involved in the secondary loop. Penguins are really cute.

5. True- The output of the primary controller is the set point for the secondary controller. The output of the secondary controller adjusts the control variable.

Scoring

If you got...
20 points: CONGRATULATIONS! You are a cascade genius!
13-19 points: Make a list of confusing topics and bring your questions to a cascade control sage.
0-12 points: You may want to re-read this article and try again...
-10 points: Choose a new major. You may be better off in LS&A.

References

- Astrom, Karl J. and Hagglund, Torr. "Advanced PID Control" The Instrumentation, Systems and Automation Society.
- Bequette, B. Wayne *Process Dynamics: Modeling, Analysis, and Simulation* Prentice- Hall PTR, Upper Saddle River, NJ 07458 (c) 2006.
- Mahoney, D.P., Svrcek, W.Y., and Young, B.R. (2006) *A Real-Time Approach to Process Control*, 2nd Edition. John Wiley & Sons, Ltd. ISBN 0470025336
- Industrial Controller: basic theory. Retrieved from http://www.sapiensman.com/control/index.htm. Last accessed on 12/12/2008

Contributors and Attributions

- **Stewards:** Irene Brockman, Robert Chockley, Scott Dombrowski, Chloe Sweet

This page titled 11.3: Cascade Control is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

11.4: Ratio Control

Introduction

Ratio control architecture is used to maintain the relationship between two variables to control a third variable. Perhaps a more direct description in the context of this class is this: Ratio control architecture is used to maintain the flow rate of one (dependent *controlled feed*) stream in a process at a defined or specified proportion relative to that of another (independent *wild feed stream*) (3) in order to control the composition of a resultant mixture.

As hinted in the definition above, ratio control architectures are most commonly used to combine two feed streams to produce a mixed flow of a desired composition or physical property downstream (3). Ratio controllers can also control more than two streams. Theoretically, an infinite number of streams can be controlled by the ratio controller, as long as there is one controlled feed stream. In this way, the ratio control architecture is feedfoward in nature. In this context, the ratio control architecture involves the use of an independent *wild feed stream* and a dependent stream called the *controlled feed*.

Ratio Control is the most elementary form of feed forward control. These control systems are almost exclusively applied to flow rate controls. There are many common usages of ratio controls in the context of chemical engineering. They are frequently used to control the flows on chemical reactors. In these cases, they keep the ratio of reactants entering a reaction vessel in correct proportions in order to keep the reaction conditions ideal. They are also frequently used for large scale dilutions.

Ratio Control based upon Error of a Variable Ratio

The first type of ratio control architecture uses the error of a variable ratio (R_{actual}) from a set ratio (R_{set}) to manipulate y (the *controlled stream*). R_{actual} is a ratio of the two variables *wild stream* and *controlled stream*. The controller adjusts the flow rate of stream y (controlled stream) in a manner appropriate for the error ($R_{actual} - R_{set}$).

The error in this system would be represented using the following equations:

$$\frac{y}{y_w} = R_{\text{actual}}$$

or

$$\frac{y_w}{y} = R_{\text{actual}}$$

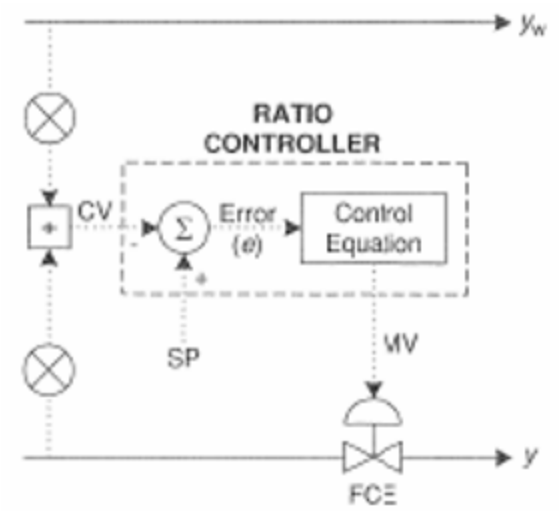
with

$$\text{Error} = R_{\text{actual}} - R_{\text{set}}$$

This error would be input to your general equation for your P, PI, or PID controller as shown below.

$$V_y = Offset + K_c(Error) + \frac{1}{\tau_I}\int(\text{Error})dt + \tau_D\frac{d(\text{Error})}{dt}$$

NOTE: Vy is used instead of y, because y is not directly adjustable. The only way to adjust y is to adjust the valve (V) that affects y.



4.2.1 Diagram of Ratio Dependant System. Image taken from Svreck et al.

Ratio Control based upon Error of the Controlled Stream

The second type of ratio control architecture uses the error of a setpoint (yset, setpoint of the control variable) from y (controlled variable) to control, once again, y. The controller adjusts the flow rate of stream y in a manner appropriate for the error (y-yset).

The error in this system would be represented using the following equation:

$$\boldsymbol{y}_w \boldsymbol{R}_{set} = \boldsymbol{y}_{set}$$

or

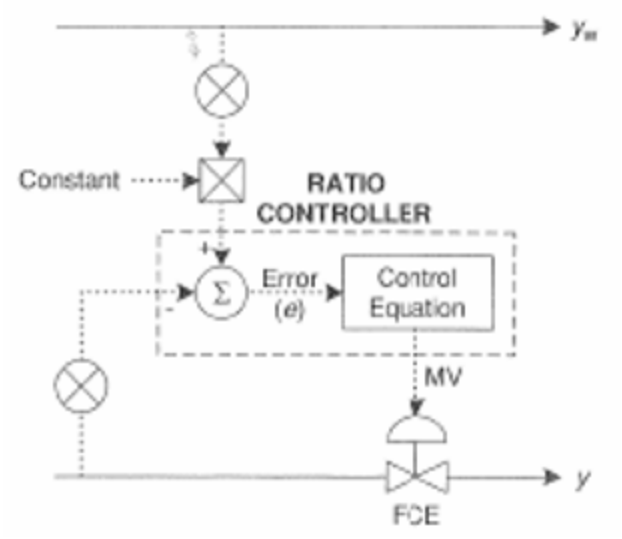
$$\frac{\boldsymbol{y}_w}{\boldsymbol{R}_{\text{set}}} = \boldsymbol{y}_{\text{set}}$$

with

$$\text{Error} = y - y_{\text{set}}$$

This error would be input to your general equation for your P, PI, or PID controller as shown below.

$$V_y = \text{Offset} + K_c(Error) + \frac{1}{\tau_I}\int (Error)dt + \tau_D \frac{d(Error)}{dt}$$



4.3.1 Diagram of Flowrate Dependant System. Image taken from Svreck et al.

Comparing the Two Types of Ratio Control

The main difference between the two aforementioned types of ratio control architecture is that they respond to changes in the monitored variable (Ractual or y - the variables through which the error is determined) differently.

The first method mentioned that defines error as (R_{actual}-R_{set}\) (monitored variable of R_{actual}) responds sluggishly when y is relatively large and too quickly when y is relatively small. This is best explained by examining the equations below.

$$R = \frac{y_w}{y} \longrightarrow \frac{\partial(R)}{\partial\left(y_w\right)} = \frac{1}{y} \tag{11.4.1}$$

$$R = \frac{y}{y_w} \longrightarrow \frac{\partial(R)}{\partial\left(y_w\right)} = -\frac{y}{y_w^2} \tag{11.4.2}$$

Unlike the first method, the second method mentioned that defines error as $y - y_{set}$ (monitored variable of y) does not respond differently depending upon the relative amounts of y (or anything else, for that matter). This is best explained by examining the equations below.

$$y = y_w R_{ytt} \longrightarrow \frac{\partial(y)}{\partial\left(y_w\right)} = R_{ytt} \tag{11.4.3}$$

$$y = \frac{y_w}{R_{set}} \longrightarrow \frac{\partial(y)}{\partial\left(y_w\right)} = \frac{1}{R_{yet}} \tag{11.4.4}$$

Difficulties with Ratio Controllers

Steady State Issues

A common difficulty encountered with ratio controllers occurs when the system is not at steady state. Unsteady state conditions can cause a delay in the adjustment of the manipulated flow rate and thus the desired state of the system cannot be met.

Blend Stations

To overcome the difficulty with steady state, a Blend station can be used, which takes into account both the set ratio and the wild stream. As can be seen by the diagram below, the Blend station takes in the wild stream flow rate, the set ratio, and the initial set point for the wild stream (r1) as inputs. Using a weighting factor, called gain, the Blend station determines a new set point for the manipulated stream (r2). The relationship that determines this new set point can be seen below.

$$r_2 = \text{set } ratio\,[\gamma^* r_1 + (1-\gamma) y_w]$$

where γ is the gain factor.

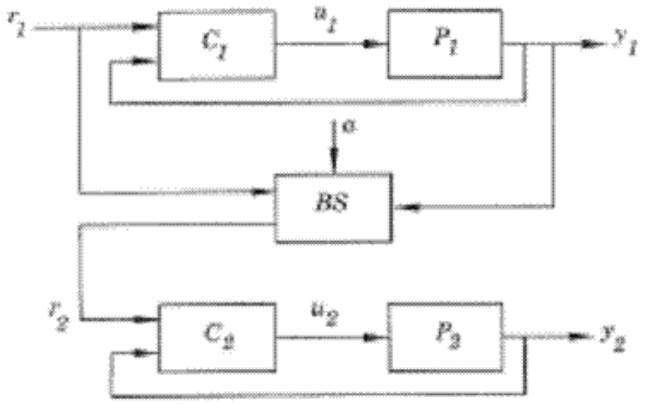


Figure 12.18 Ratio control using the Blend station *(BS)*.

Image taken from Astrom and Hagglund article on "Advanced PID Control."

Accuracy Issues

Another problem (which is an issue with all "feed-forward" controllers) is that the variable under control (mix ratio) is not directly measured. This requires a highly accurate characterization of the controlled stream's valves so that the desired flow rate is actually matched.

One way to address this problem is to use two levels of PID control (Cascade Control). The first level monitors the controlled streams flow rate and adjusts it to the desired set point with a valve. The outer level of control monitors the wild stream's flow rate which adjusts the set point of the controlled stream by multiplying by the desired ratio. For example, if stream A's flow rate is measured as 7 gpm, and the desired ratio of A:B is 2, then the outer level of control will adjust B's set point to 3.5. Then the inner level of control will monitor B's flow rate until it achieves a flow rate of 3.5.

Ratio Control Schemes

The following subsections introduce three ratio controller schematics. Each schematic illustrates different ways that the controlled feed stream can be manipulated to account for the varying wild feed stream.

Ratio Relay Controller

The wild feed flow rate is received by the ratio relay and then multiplied by the desired mix ratio. The ratio relay outputs the calculated controlled feed flow rate which is compared to the actually flow rate of the controlled feed stream. The flow controller then adjusts the controlled feed flow rate so that it matches the set point (3).

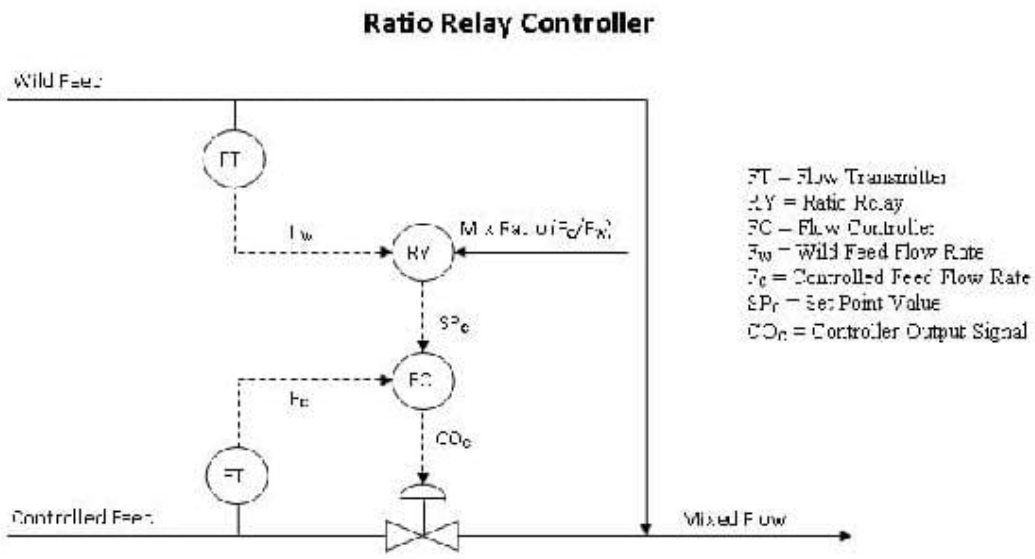


Image adapted from Houtz, Allen and Cooper, Doug "The Ratio Control Architecture"

The mix ratio (Fc/Fw) is not easy to access, so it requires a high level of authorization to change. This higher level of security may be an advantage so that only permitted people can change the mix ratio and decrease the chance that an accidental error occurs. A disadvantage is that if the mix ratio does need to be changed quickly, operation may be shut down while waiting for the appropriate person to change it.

Another disadvantage is that linear flow signals are required. The output signals from the flow transmitters, Fw and Fc, must change linearly with a change in flow rate. Turbine flow meters provide signals that change linearly with flow rates. However, some flow meters like orifice meters require additional computations to achieve a linear relationship between the flow rate and signal.

Flow Fraction Controller

Flow fraction and ratio relay controllers are very similar except that the flow fraction controller has the advantage of being a single-input single-output controller. A flow fraction controller receives the wild feed and controlled feed flow rates directly. The desired ratio of the controlled feed to wild feed is a preconfigured option in modern computer control systems (3).

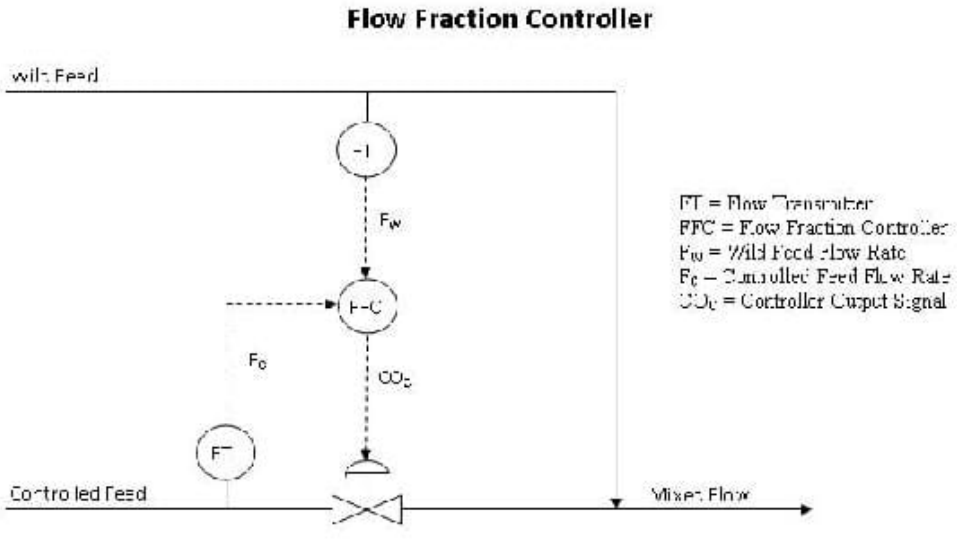


Image adapted from Houtz, Allen and Cooper, Doug "The Ratio Control Architecture"

Ratio Relay with Remote Input

The ratio relay with remote input model of a ratio controller is similar to the cascade control model in that it is part of a larger control strategy. The purpose of this model is to account for any unexpected or unmeasured changes that have occurred in the wild and controller feed streams. The diagram below illustrates how the composition of the mixed flow is used to change the relay ratio (3).

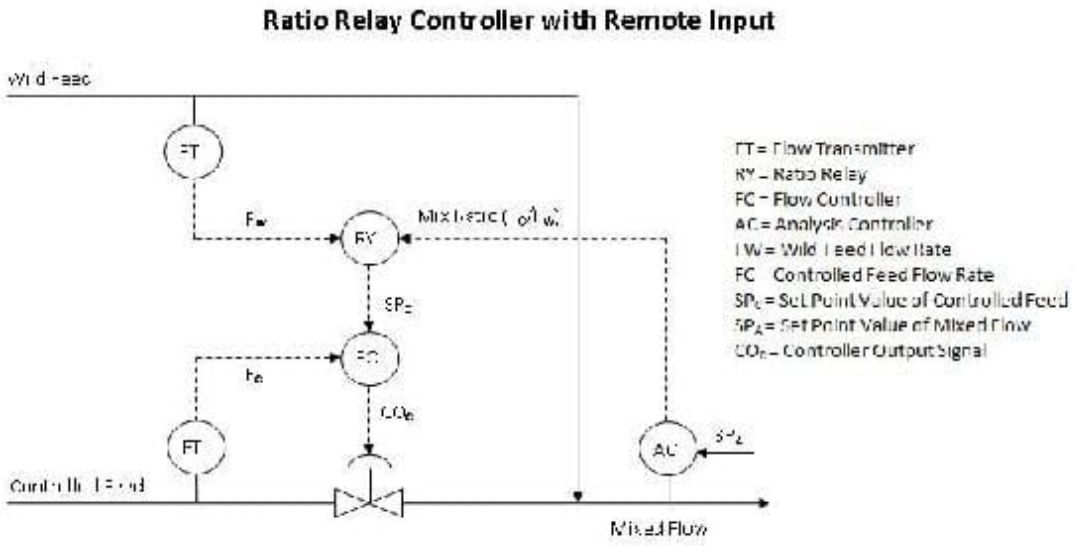


Image adapted from Houtz, Allen and Cooper, Doug "The Ratio Control Architecture"

The main advantage of having remote input is that the mix ratio is constantly being updated. A disadvantage of using an additional analyzer is that the analysis of the mixed stream may take a long time and decrease the control performance. The lag time depends on the type of sensor being used. For example, a pH sensor will most likely returned fast, reliable feedback while a spectrometer would require more time to analyze the sample.

Advantages and Disadvantages

There are some pros and cons to using ratio control where there is a ratio being maintained between two flow rates.

Advantages

1. Allows user to link two streams to produce and maintain a defined ratio between the streams
2. Simple to use
3. Does not require a complex model

Disadvantages

1. Often one of the flow rates is not measured directly and the controller assumes that the flows have the correct ratio through the control of only the measured flow rate
2. Requires a ratio relationship between variables that needs to be maintained
3. Not as useful for variables other than flow rates

Select Elements in Ratio Control

A select element enables further control sophistication by adding decision-making logic to the ratio control system. By doing so, a select variable can be controlled to a maximum or minimum constraint. The figure below depicts the basic action of both a low

select and high select controller.

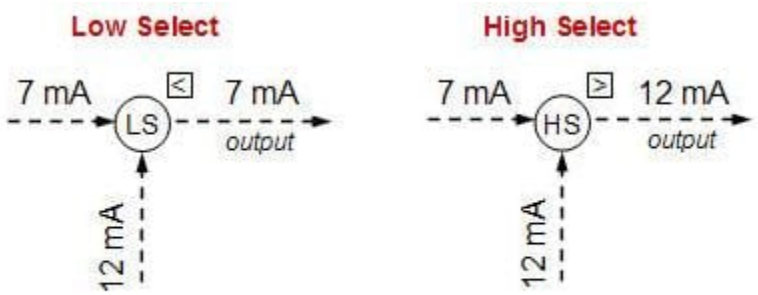
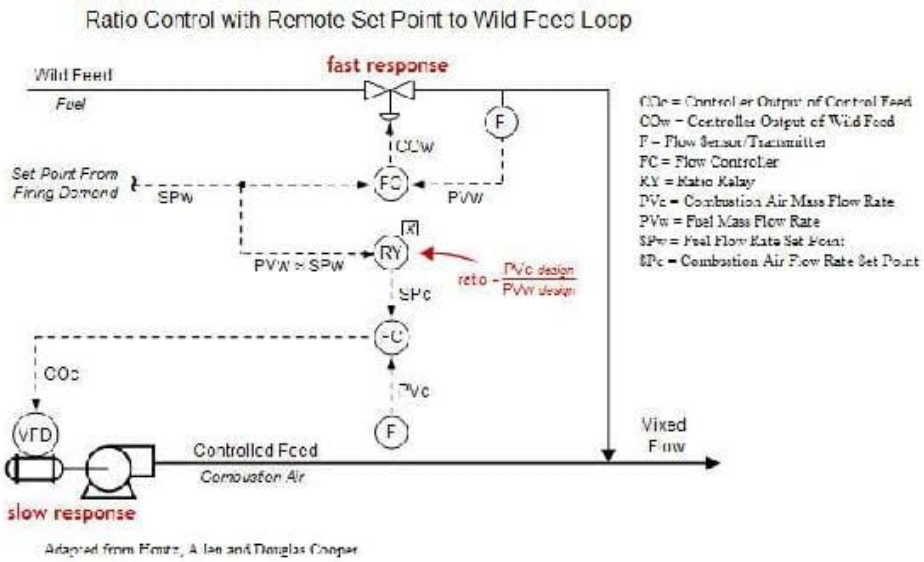


Image adapted from Houtz, Allen and Douglas Cooper

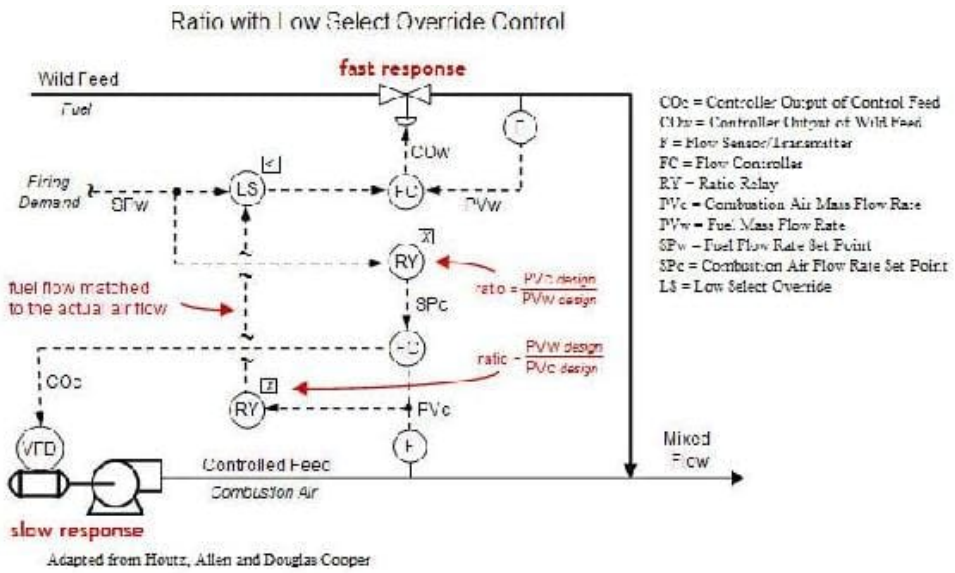
Single Select Override Control

Often times, these select elements are used as an override element in ratio control. Take, for example, ratio control system (with remote input) of a metered-air combustion process in the figure below.



Adapted from Houtz, Allen and Douglas Cooper

In this design, the fuel set point, SPw, comes in as firing demand from a different part of the plant, so fuel flow rate cannot be adjusted freely. As the fuel set point (and therefore fuel mass flow rate, PVw) fluctuates, a ratio relay is employed to compute the combustion air set point, SPc. If the flow command outputs (COw and COc) are able to respond quickly, then the system architecture should maintain the desired air/fuel ratio despite the demand set point varying rapidly and often. However, sometimes the final control element (such as the valve on the fuel feed stream and blower on the combustion air feed stream in the diagram) can have a slow response time. Ideally, the flow rates would fluctuate in unison to maintain a desired ratio, but the presence of a slow final control element may not allow the feed streams to be matched at that desired ratio for a significant period of time. Valves often have quick response times, however, blowers like the one controlling the combustion air feed stream can have slow response times. A solution to this problem is to add a “low select” override to the control system, as shown in the figure below.



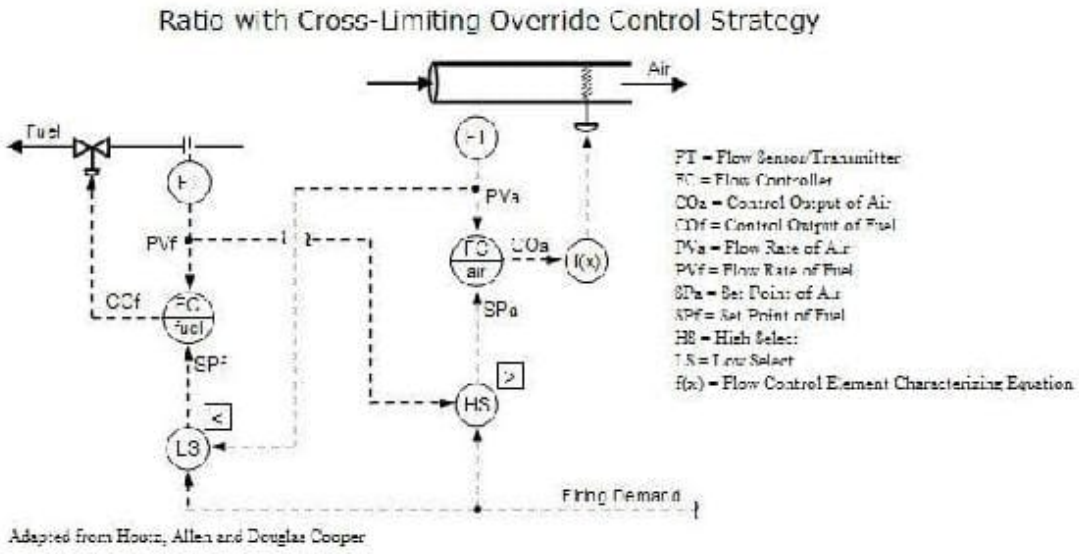
Adapted from Houtz, Allen and Douglas Cooper

The second ratio controller will receive the actual measured combustion air mass flow rate and compute a matching fuel flow rate based on the ideal design air/fuel mixture. This value is then transmitted to the low select controller, which also receives fuel flow rate set point based from the firing demand. The low select controller then has the power to pass the lower of the two input signals forward. So, if the fuel flow rate firing demand exceeds combustion air availability required to burn the fuel, then the low select controller can override the firing demand and pass along the signal of the fuel flow rate calculated based on the actual air flow (from the second ratio relay).

This low select override strategy ensures that the proper air/fuel ratio will be maintained when the firing demand rapidly increases, but does not have an effect when the firing demand is rapidly decreasing. While the low select override can help eliminate pollution when firing demand rates increase quickly, a rapid decrease in firing demand can cause incomplete combustion (as well as increased temperature) and loss of profit.

Cross-Limiting Override Control

Cross-limiting override control utilizes the benefit of multiple select controls. Most commonly, a control system using the cross-limiting override strategy will implement both a high select and low select control. As a result, the system can account for quickly increasing firing demand as well as quickly decreasing firing demand. The figure below depicts a such a scenario using the inlet combustion air and fuel for a boiler as the example.

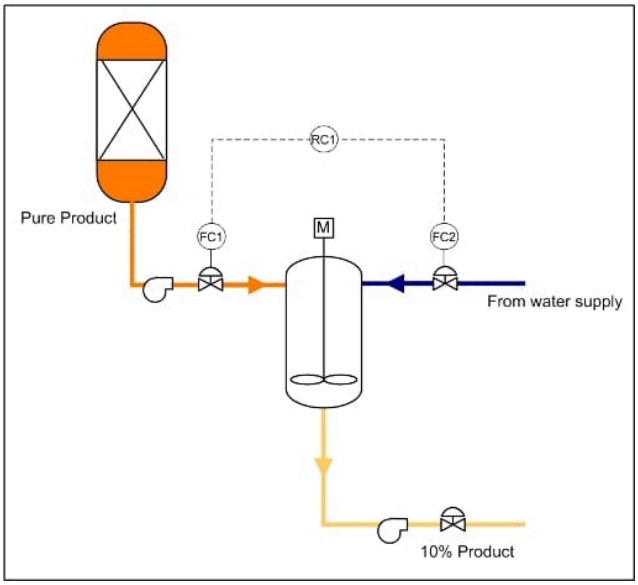


Adapted from Houtz, Allen and Douglas Cooper

It is assumed that same firing demand enters the high select and low select control and the flow transmitters have been calibrated so that the ideal air/fuel ratio is achieved when both signal match. As a result, the set point of air will always be the greater value of the firing demand and current fuel flow signals. Likewise, the set point of fuel will always be the lower value of the firing demand and current air flow signals. So, as the firing demand increases, the set point of air will increase. At the same time, the low select will keep the set point of fuel to the signal set by the present flow or air. Therefore, the set point of fuel will not match the increasing firing demand, but will follow the increasing air flow rate as it is responding upwards. Similarly, if the firing demand decreases, the low select control will listen to the firing demand and the high select controller will not. As a result, the firing demand will directly cause the set point of fuel to decrease while the set point of air will follow the decreasing fuel rate as it responds downwards. The cross-limiting override strategy allows for greater balance in a ratio control system.

Worked out Example 1

In the following process shown below, a concentrated solution of product is diluted continuously to be sold as a final 10% solution. Flowrates coming from the unit feeding the pure product to this mixing tank are not constant (the wild stream), and therefore a ratio controller is used to properly dilute the solution.



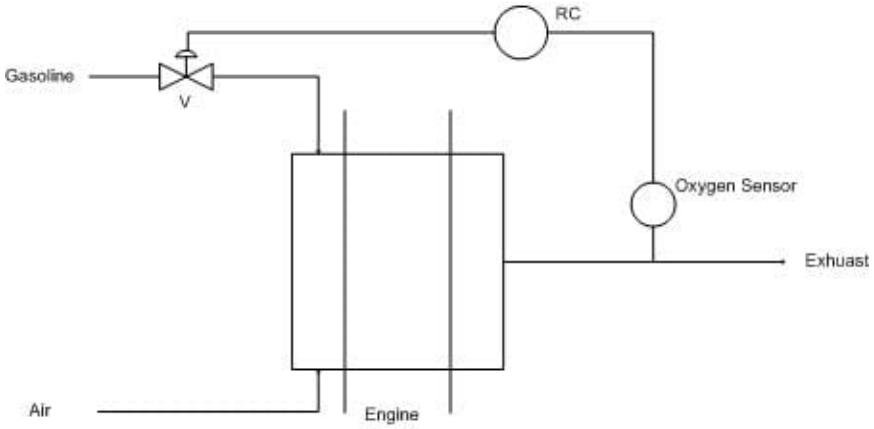
In this example, the ratio controller would be set to a value of FC1/FC2 = RC1 = 9. The ratio contoller in this case would then work by the following logic:

For an additional challenge, determine an error equation that can is related to this controller.

Worked out Example 2

One of the most common uses of a ratio controller is used for correcting the air-fuel ratio in automobiles. A specified amount of air is always taken up by the engine and the amount of fuel injected is controlled by a computer through the use of a ratio controller. The stoichiometric ratio for air to gasoline by mass is 14.7:1. There is an oxygen sensor in the exhaust manifold as well as sensors for all of the exhaust components. This sensor is connected to the valve for the fuel injection. If there is too much oxygen measured in the exhaust, then the mixture is said to be lean. Is the ratio of air to gas in the original mixture greater than or less than 14.7:1 and what should the valve controlling the fuel injection do?

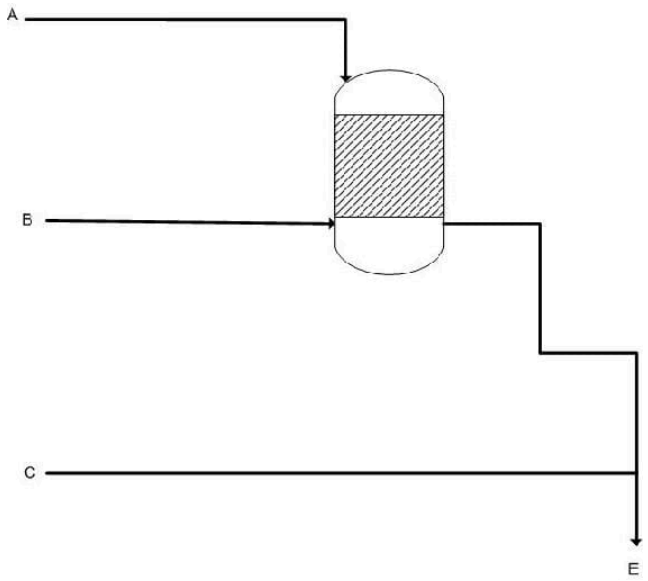
Note: This is an indirect form of a ratio controller as it measuring a physical property of downstream in order to determine if the ratio is too high or low. Further information may be found here about Air-fuel Ratio and Oxygen Sensors.



Answer: Since there is leftover oxygen in the exhaust, there must have been too much in the original mixture. The ratio of air to gas in the original mixture is greater than 14.7:1 and thus the computer should open the valve more for the fuel injection.

Worked out Example 3

Many processes in industry require multiple process streams entering at different locations in the system. A senior engineer gives you the project to design the control scheme for a new process to be implemented at the company. The process involves 2 reactions and 3 different feed streams.



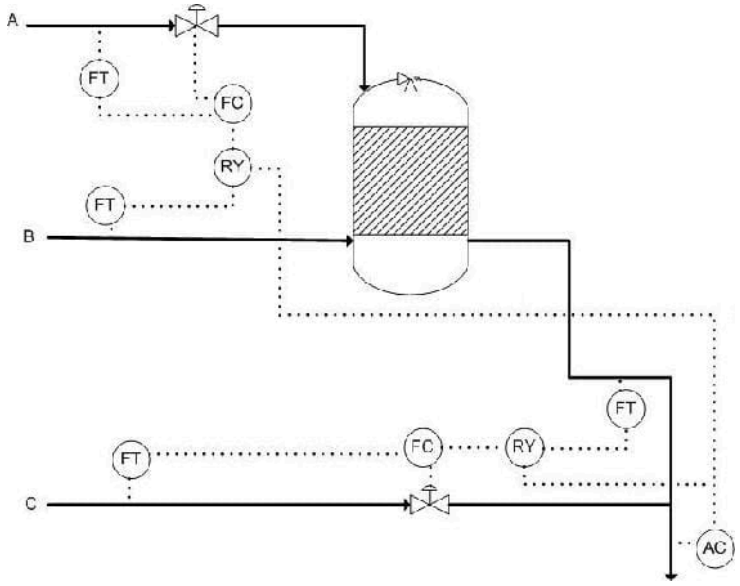
Reactions: 2A + B --> D

$$\mathrm{C} + 3\,\mathrm{D} \longrightarrow \mathrm{E}$$

Assume the reactions go to 100% completion (you only need stoichiometric amounts).

Complete the P&ID with the control scheme and write the ratios (with respect to B) that the other molar flow rates will be set to. Also explain how the system will take action when it detects excess C or D in the product stream.

Solution:



A: 2 B: 1 C: 1 D: 1/3 E: 1/3

If the AC detects greater or less amounts of C expected, it will adjust the ratio setting in RY connected to C's flow controller, reducing or decreasing the ratio of C to B, respectively.

Exercise 11.4.1

Why is a Blend Station used?

a. to create the set point for the ratio controller
b. to reduce lag time in adjusting flow rates
c. to measure flow rates
d. to multiply one flow rate by the desired ratio

Answer available at the end of this section.

Exercise 11.4.2

How is the set ratio used in the first type of ratio controller described above?

a. Outputs how much each stream should be flowing.
b. It is compared against two flow rates to adjust one if needed.
c. Tells the system to shut down.
d. It adjusts for lag time when the system is not at steady-state.

Answer available at the end of this section.

References

- Astrom, Karl J. and Hagglund, Torr. "Advanced PID Control" The Instrumentation, Systems and Automation Society.
- Mahoney, Donald P., Svrcek, William Y. and Young, Brent R. "A Real-Time Approach to Process Control", Second Edition. John Wiley & Sons, Ltd.
- Houtz, Allen and Cooper, Doug "The Ratio Control Architecture" http://www.controlguru.com (http://www.controlguru.com/

2007/120207.html)

Answers

Exercise 11.4.1

TBA

Exercise 11.4.2

Add texts here. Do not delete this text first.

This page titled 11.4: Ratio Control is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Benjamin Terrien, John Zalewski, Michael Oom, & Stephanie Fraley via source content that was edited to the style and standards of the LibreTexts platform.

11.5: Common control loops and model for liquid pressure and liquid level

This page titled 11.5: Common control loops and model for liquid pressure and liquid level is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

11.6: Common Control Loops and Model for Temperature Control

Introduction

Temperature, pressure, flow, and level are the four most common process variables. Temperature is important because it provides a critical condition for processes such as combustion, chemical reactions, fermentation, drying, distillation, concentration, extrusion, crystallization, and air conditioning. Poor temperature control can cause major safety, quality, and productivity problems. Although highly desirable, it is often difficult to control the temperature because its measurement must be within a specified range of accuracy and have a specified degree of speed of response, sensitivity, and dependability. Additionally, temperature measurements must also be representative of true operating conditions in order to achieve successful automated control. The instrument selected, installation design, and location of the measuring points determine these specifications.

This article will serve as a guide to select the best location of measuring points to achieve the best automatic control. It will consider temperature control for three common process types: a CSTR, distillation column, and heat exchanger.

Temperature Control Loops

Before temperature control loops for specific processes are explained, we must discuss the general considerations common for all temperature control loops.

Temperature control loops can either be endothermic (requiring heat energy) or exothermic (generating heat energy). Both types are similar in that they both result in a response representing a process with a dominant capacitance plus a dead time. For both types of processes, one of the following devices is used to measure temperature:

- thermocouple
- filled thermal well system
- Resistance temperature detector (RTD)

The measurement device, or thermal well, should be selected so that is minimizes additional lag to the overall process lag. Minimizing temperature measurement lag in the temperature control loop is important in both slow and fast loops. Some general rules of thumb for reducing temperature measurement lag are:

1. Use a small-diameter bulb or thermal well to minimize the thermal resistance and thermal capacity of the measuring element.
2. Use a thermal well made from a material that minimizes thermal resistance and thermal capacity of the measuring element.
3. Use a small pipe or orfice near the measuring device to increase velocity of the passing flow. Increasing flow will increase the rate of heat transfer between the process fluid and the measuring device.
4. Place the measuring element in the liquid phase when measuring temperature in a two-phase system since thermal resistance is smaller in the liquid phase than in the vapor.
5. Use a transmitter with derivative action to cancel out some of the lag in the measuring element. Compensate for this added derivative gain in the transmitter by reducing the derivative gain in the controller.

CSTR Temperature Control

Endothermic Reactor Temperature Control Loops

Endothermic CSTR reactors are generally easier to control than exothermic CSTR reactors because when the temperature reaches a critical minimum, the reaction does not proceed until there is adequate heat. In this sense, endothermic CSTR reactors are self-regulating. A good way to think of an endothermic CSTR's controls is that of a heat exchanger being used to heat a reaction solution. This heat exchanger's response is controlled by the dead time, so typically a PI or PID controller is used.

There are two main types of temperature control methods for endothermic CSTRs: control via steam flow rate and control via

steam pressure. For the steam flow rate control case, the temperature control sends a signal to the flow control, which then controls the steam inlet valve. For the steam pressure control case, the temperature control sends a signal to the pressure control, which controls the steam inlet valve. Examples of an endothermic CSTR temperature control loop can be seen below in Figure 1 a-b.

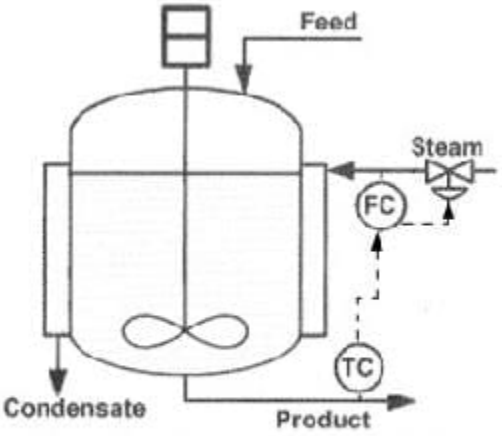


Figure 1.a. Temperature Control of an Endothermic CSTR via Steam Flow Rate

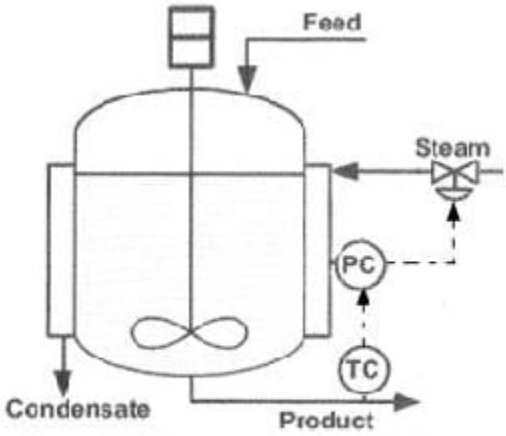


Figure 1.b. Temperature Control of an Endothermic CSTR via Steam Pressure

Image Modified From: Riggs, James B., Karim, Nazmul M.. Chemical and Bio-Process Control. Third Edition Chapter 18. Ferret Publishing.

Exothermic Reactor Temperature Control Loops

In an exothermic reaction, energy is released in the form of heat. In some cases, a cooling system is require to bring the temperature back to a set point and also to ensure that the temperature does not rapidly increase, or runaway. Some steps can be taken to prevent runaway temperatures, such as reducing the feed rates or concentrations of reactants. Also, the ratio of the heat transfer area to the reactor volume can be increased to help increase the controllability of the CSTR. Typically, the temperature is controlled using a PID controller, which is described in the previous section PID Control.

Exothermic CSTRs are very difficult to control because they are very unstable and the temperature can easily runaway. The relationship between heat generation and temperature is non-linear, while the relationship between heat removal and temperature is linear. This relationship is illustrated in Figure 2. The stability of the temperature control loops depend on the rate at which heat can be removed from the system. For example, an exothermic CSTR that generates heat at a slow steady rate is more stable than a reactor that rapidly produces heat. In this example, the rate at which the heat can be removed from the system depends on the rate at which the temperature can be changed in the cooling jacket surrounding the CSTR.

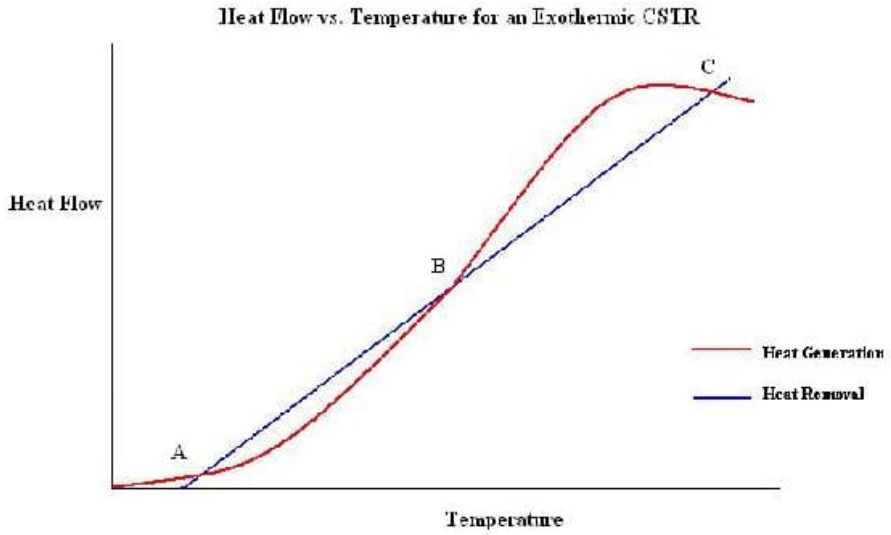


Figure 2. Heat generation or removal as a function of temperature in an exothermic CSTR

The intersections of the two curves, labeled A, B and C, represent steady state operating points. A and C are stable operation points while B is unstable. A and C are stable because as the temperature increases from this point, the rate of heat generation is less than the rate of heat removal. This means that as the temperature increases, heat is removed faster than it is generated so the temperature will be brought back down to that operating point. The same will happen if the temperature decreases from that point. As the temperature decreases, the heat generation rate is greater than the heat removal rate so the temperature will be brought back up to the operating point.

Point B is unstable because when the temperature increases from that point, the heat generation is greater than the heat removal. This means that as the temperature increases, heat is continuously added to the system and the temperature will rise until it reaches one of the stable operating points, in this case C. If the temperature decreases from point B, the heat generation rate is less than the heat removal rate so the temperature will continue to decrease until it reaches the lower, stable operating point, A.

This graph would be similar for an endothermic CSTR but there would be only one stable steady state operating point.

Very rapid exothermic reactions are the most difficult to control and they are sometimes carried out in a semi-batch reactor so the addition of reactants can be carefully controlled and runaway temperatures can be avoided.

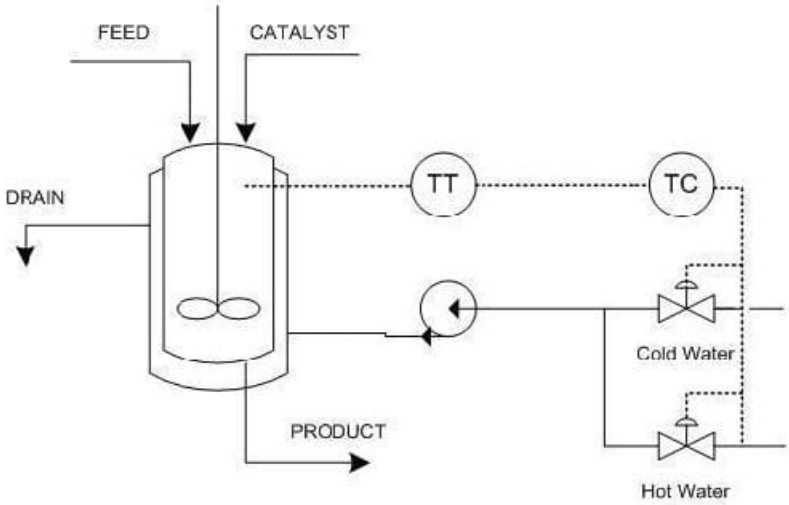


Figure 3.Temperature control loops for an exothermic CSTR

Figure 3 illustrates a common control loop for an exothermic CSTR with a cooling jacket. A temperature transmitter (TT) sends a signal to the temperature controller (TC), which controls the hot and cold water valves on the jacket. The cold and hot water are pumped into the jacket which controls the reactor at a set temperature.

Temperature Control in Distillation

In a distillation column, temperature control is used as a means to control composition because temperature sensors are cheaper, more reliable, provide continuous measurements, and respond quicker than composition analyzers. Through equilibrium relationships, temperature measurements can be used to infer composition of the product. Inferential temperature control is only effective when the relative volatility of the components is high (greater than 2.0). Temperature controllers are used as feedback composition controllers to adjust column operation to meet production requirements. They must be able to satisfy the constraints defined by production requirements at all times, even in the face of disturbances.

Inferential Temperature Control

To control composition via temperature, a correlation must be made between temperature in the tray and composition of the key components. Determining a correlation is challenging because the temperature-composition relationship is affected by process nonlinearity and disturbances in feed composition, flowrates, and occurrence of entrainment or fouling. For multi-component separations, temperature does not determine a unique composition, so either an online composition analyzer or periodic lab tests must be utilized to verify composition at the temperature setpoint. Additionally, the column pressure greatly affects the measurement of tray temperature. For most systems, the following linear equation can be used to correct for variations in pressure:

$$T_{pc} = T_{meas} - K_{pr}\left(P - P_0\right) \tag{11.6.1}$$

Where,

- T_{pc} =corrected temperature
- T_{meas} =measured tray temperature
- K_{pr} =pressure correction factor
- P =column pressure
- P_0=reference pressure

The pressure correction factor, K_{pr}, can be estimated using a steady-state column simulator for two different operating pressures and the equation:

$$K_{pr} = \frac{T_i\left(P_1\right) - T_i\left(P_2\right)}{P_1 - P_2} \tag{11.6.2}$$

where T_i is temperature of tray i predicted by the column simulator.

Successful temperature control in the column depends on the dynamic response of measuring the tray temperature with respect to the manipulated energy source used to actuate temperature. The energy source is either the reboiler or the reflux. To have tight process control means that the equivalent dead time in the loop is small compared to the shortest time constant of a disturbance with significant amplitude. According to Svreck, et al, the following observations from experimental tests are cited:

- Temperature control is made less stable by measurement lag or response times.
- The speed of response and control stability of tray temperature, when controlled by reboil heat, is the same for all tray locations.
- The speed of response and control stability of the tray temperature, when controlled by reflux, decrease in direct relation with the number of trays below the reflux tray.
- When pressure is controlled at the temperature control tray, the speed of response of the temperature control instrument can vary considerably with tray location, and is normally slower.

To achieve the best composition control, you need to determine the tray(s) whose temperature(s) show the strongest correlation with product composition. The following procedure using a steady-state column model can be used:

1. Run the column at the base conditions (x^{BC} and y^{BC} at steady-state and record the temperature, T_i^{BC}, of each tray.
2. Increase the impurity level in the bottoms product ($x^{BC} + \Delta x$ and y^{BC}), so that Δx is about 25-50% of the impurity level for the base case. Record the temperature, $T_i^{\Delta x}$, of each tray.
3. Increase the impurity level in the overhead product $(x^{BC}$ and $y^{BC} + \Delta y)$, so that Δy is about 25-50% of the impurity level for the base case. Record the temperature, $T_i^{\Delta y}$, of each tray.
4. The best tray for inferential temperature control of the stripping section will be the least sensitive to variations in the composition of the overhead product. This tray is the one that maximizes:

$$\Delta T_i^{\text{net}} = \left(T_i^{\Delta x} - T_i^{BC}\right) - \left(T_i^{\Delta y} - T_i^{BC}\right) \tag{11.6.3}$$

5. The best tray for inferential temperature control of the rectifying section will be the least sensitive to variations in the composition of the bottoms product. This tray is the one that maximizes:

$$\Delta T_i^{net} = \left(T_i^{\Delta y} - T_i^{BC}\right) - \left(T_i^{\Delta x} - T_i^{BC}\right) \tag{11.6.4}$$

6. Repeat this procedure for a range of feed compositions. Then, select the tray that maximizes the ΔT_i^{net} equations most often for the range of feed compositions.

Single Composition Control

Once you have determined the best tray to use for inferential temperature control, you must choose a manipulated variable (MV) that will have a significant gain on the tray temperature, and thus on the composition. Usually in distillation, the composition of only one product stream is controlled while the composition of the other product stream is allowed to drift. This is called single composition control. The chemical industry primarilly uses sing composition control.

When you are interested in controlling the composition of the overhead product, the reflux is used to control the purity, while the reboiler duty is held constant. The bottom composition is not directly controlled and will vary as the feed composition varies. The reflux, L, is the manipulated variable that will result in the tightest control, rather than either the distillate flowrate, D, or the reflux ratio, L/D. This is because L is the fastest-acting MV for the overhead and is the least sensitive to changes in the feed composition This scenario is depicted below in Figure 4a.

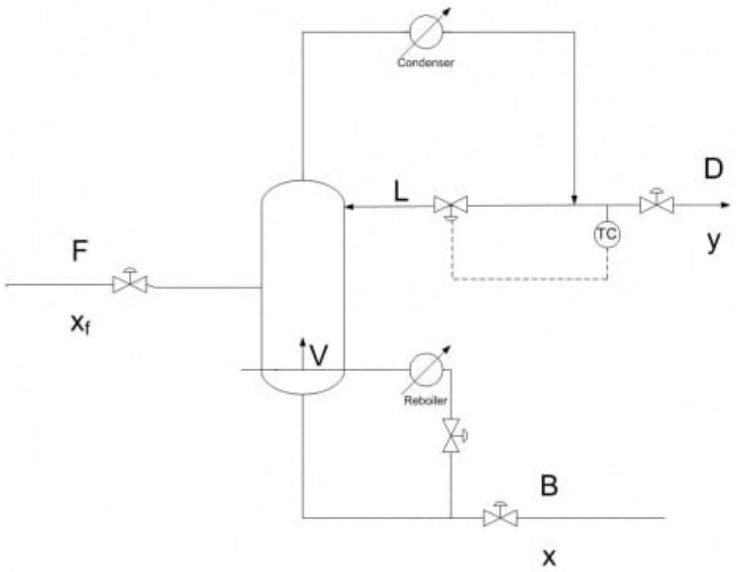


Figure 4a. Control diagram for inferential temperature-single composition control of overhead product stream by L.

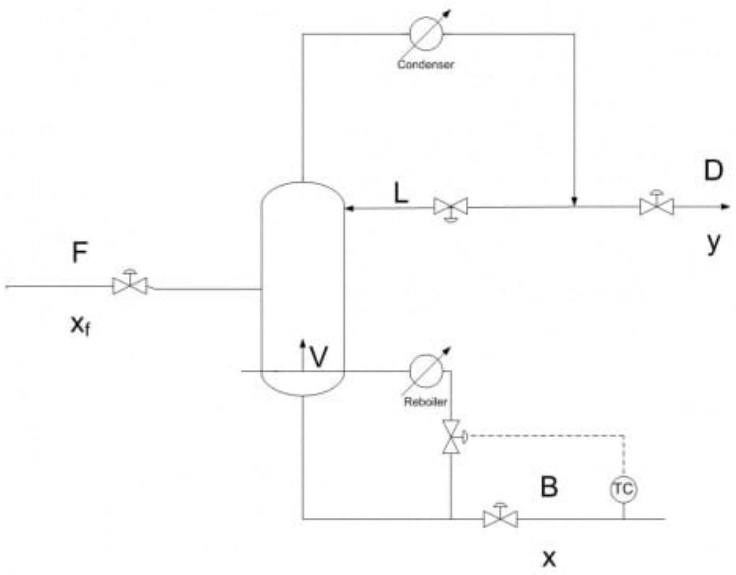


Figure 4b. Control diagram for inferential temperature-single composition control of bottom product stream by V.

When you are interested in controlling the composition of the bottom product, the boilup rate, V, is used to control the purity, while the reflux rate is held constant. The overhead composition is not directly controlled and will vary as the feed composition varies. V is the manipulated variable that will result in the tightest control, rather than either the bottoms product flowrate, B, or the boilup ratio, V/B. This is because V is the fastest-acting MV for the bottoms and is the least sensitive to changes in the feed composition This scenario is depicted above in 4b.

Dual Composition Control

Although single composition control is most commonly used, there are some industries that need to simultaneously control the composition of both the overhead and bottom product streams. This is called dual composition control and is harder to implement, tune, and maintain than single composition control. Dual composition control increases product recovery and reduces utility costs and is used for refinery columns that generally have high associated energy usage. Industries that have refineries, such as the oil industry, use dual composition control because they are interested in using both the overhead and bottom streams as product.

In dual composition control, there are many possible control configurations. The control configuration is described by the control objective and by the MV. There are four possible control objectives: bottom product x, overhead product y, reboiler level, and accumulator level. The chosen control objectives are then paired with a MV. There are a variety of choices for the MV. These include: L, D, L/D, V, B, V/B, B/L, and D/V. Therefore, in dual composotion control, there are many possible configurations to consider, although most are not practical.

The most commonly used configuration for dual composition control is (L, V) which uses L to control y (overhead composition) and V to control x (bottom composition), because it provides good dynamic response, is the least sensitive to changes in feed composition, and is the easiest to implement. However, it is highly susceptible to coupling. In this configuration, depicted below, the setpoint for the reflux flow controller is set by the overhead composition controller, and the setpoint for the flow controller on the reboiler duty is set by the bottom composition controller. D is then used to control the accumulator level, and B is used to control the reboiler level.

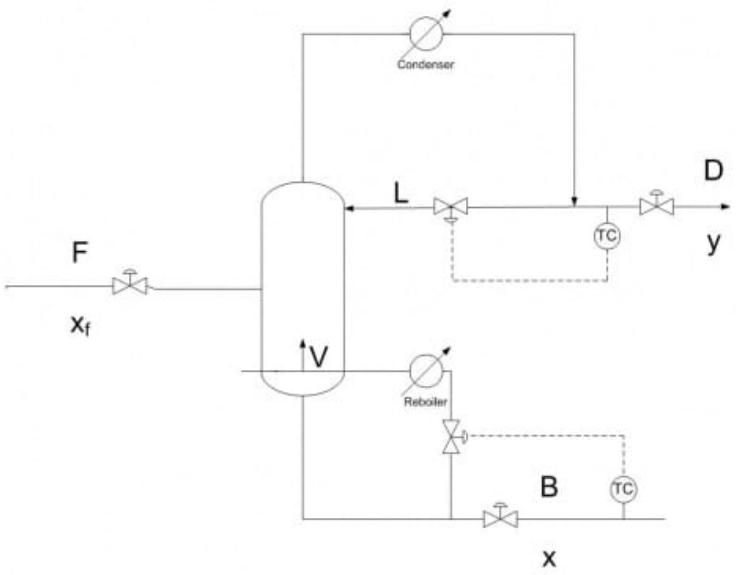


Figure 5. Control diagram for inferential temperature-dual composition configuration (L,V), the control of overhead product stream by L and bottom product stream by V.

Since there are so many possibilities for configurations, there is no clear best choice for a configuration of MVs in dual composition control in distillation columns. It is impossible to know theoretically which configuration is optimum for a particular process, but there are some rules of thumb to follow to increase the possibility of choosing a good configuration. These rules are summarized in the chart below.

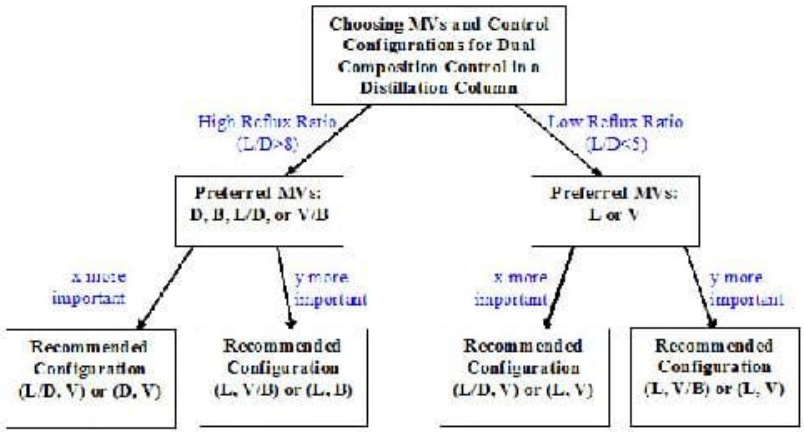


Figure 6. Chart to determine a good configuration for your process when using dual composition inferential temperature control

For the gray area where 5

Controller Tuning and Constraints

When temperature is used for inferential composition control, a PID controller is usually used because of the significant sensor lag. Common constraints on the extent of temperature control in a distillation column include:

- Capacity of reboiler power
- Capacity of overhead condenser
- Flooding or entrainment

- Maximum (undesired) column temperature

In general, temperature control can be used to control composition in a distillation column. The temperature measuring device should be chosen to minimize the lag time. A steady-state model can be used to correct for pressure changes in the column and to determine the best tray at which to measure temperature. The choice of the MV depends on if you are operating under single or dual composition control and which product composition you want to control.

Heat Exchanger Control

In heat exchanger control, the temperature of the process exit stream is the controlled variable (CV) and can be adjusted by one of four possible manipulated variables: cool side entry stream, cool side exit stream, hot side entry stream, or hot side exit stream. The selection of where control will be implemented is a combined result of three factors:

1. Whether we want to heat or cool the process stream
2. The response time of the controller
3. The capital cost

To illustrate the selection process, first it is shown that the hot stream preferred for temperature control, then two case studies are discussed, one for heating up the process stream and the other for cooling the process stream.

Controlling the Cool Side Stream

It is not advisable to place temperature control on the cool side stream. This is explained below.

The particular MV in each stream is the flow rate. It is known that the temperature of the hot stream will change once the flow rate of the cool stream is adjusted.

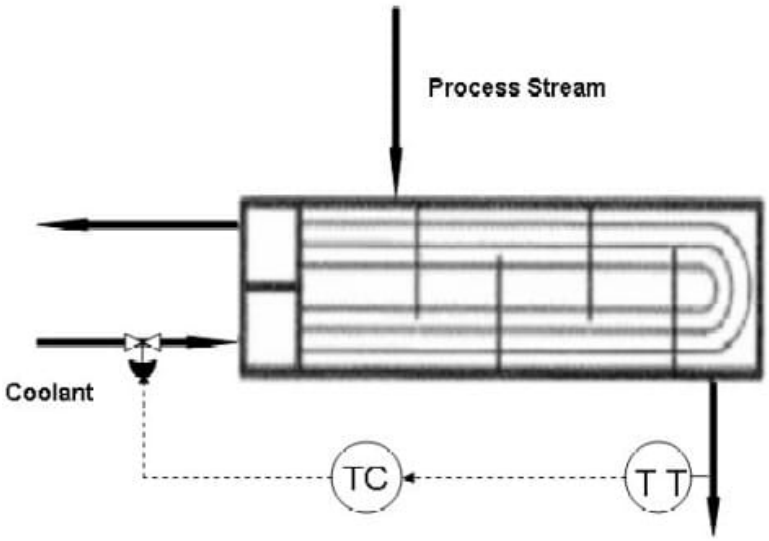


Figure 7. Image modified from Riggs & Karim, shows coolant temperature control loop for a Liquid/Liquid Heat Exchanger.

It is easier to think of the temperature change as the magnitude of the process gain ($|T_{t+1} - T_t|$). This just is the magnitude of the temperature change over a specified interval. This way the discussion can be general for both heating or cooling the process stream.

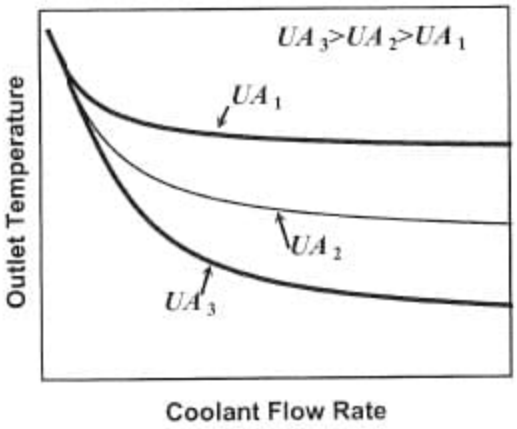


Figure 8. Taken from Riggs & Karim, demonstrates the limit that increasing the flow rate of the coolant stream has on the exit temperature of the process stream for three heat exchanger areas

As the rate of the cool stream is increased from zero, the magnitude of the process gain continuously decreases. Above a certain rate, the gain will not be a measurable change in the process stream temperature. Because of this, *making the coolant stream the MV will make the process uncontrollable at a critical point and is not a good idea.*

Controlling the Hot Side Stream

As indicated above it is never recommended to place the temperature control on the cool side stream. This reduces controls consideration to only the hot stream. There are two cases where the control can be on the hot stream; when heating up and when cooling down the process stream.

Heating Up The Process Stream

CASE STUDY 1: Steam-Heat Heat Exchanger

In this case the process stream is the cooler of the two entering streams, the previous section states that the MV would have to on the stream used to heat up the process stream.

The temperature control loop can either be the direct steam flow or the steam pressure. Both controls the amount of steam flowing into the heat exchanger by adjusting a valve, however, the preferred of the two is the steam pressure as the MV in the temperature control loop. This is because a change in the steam supply can quickly cause the internal pressure of the heat exchanger to change hence affecting the temperature exchange. The steam pressure control responds and corrects this faster than the steam flow control. This becomes especially useful for heavy duty requirements when the change in response to change in flow rate of the process stream. The two figures below show two preferred configurations;

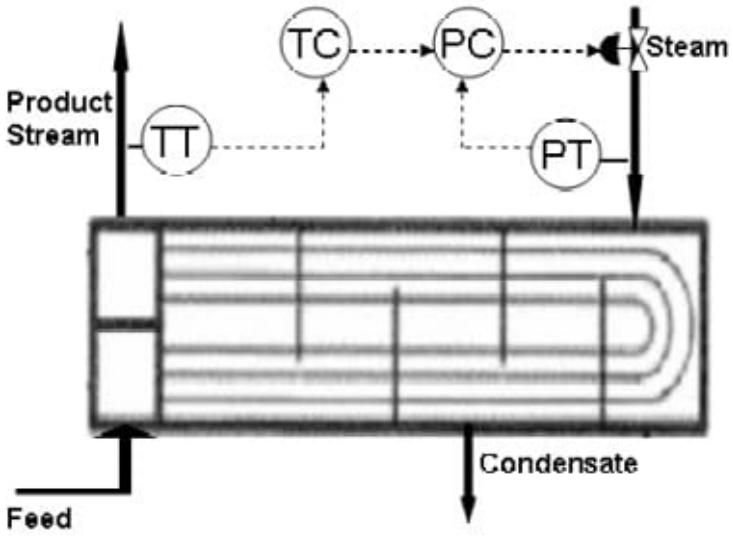


Figure 9. Image modified from Riggs & Karim, shows one arrangement for temperature control for heating up a process stream

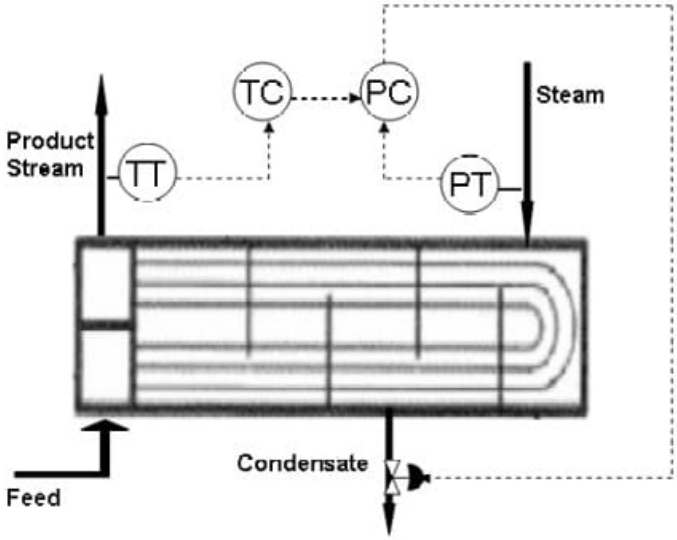


Figure 10. Image modified from Riggs & Karim, shows an alternative arrangement for temperature control for heating up a process stream

Notice that Figure 10 shows the control valve on the condensate stream. This is an acceptable control placement, however, in most practical application the first of the configurations in Figure 9 is preferred. In addition there are some ramifications to making the condensate flow the MV. The temperature loop in this case would not be as responsive as in the previous case. This is because the level responds slower than the pressure to the changes in the respective valves. The capitol cost involved in placing the valve in the condensate stream is much lower that that in the steam stream because the steam stream is generally larger and requires a larger, more expensive valve. In addition a steam trap is required down stream of the condensate valve, for that configuration.

Cooling Down The Process Stream

CASE STUDY 2: Liquid/Liquid Heat Exchanger

In this case the process stream is the hotter of the two entering stream so the controls will have to be on this stream. As in case one we are setting up the temperature control loop for varying flows in the process stream because in a plant the flow rate of this stream is dependent on a process before the heat exchanger. Hence when creating control loop for cooling down the process stream the coolant flow is kept constant (as stipulated in the section "Coolant Stream Controls") and maintain the CV at the set point. The best choice for the MV in this temperature control loop is having a feed bypass stream, as illustrated in Figure 11 below.

The advantages of having this bypass stream are that the process dynamics are considerably faster with low levels of dead time and this condition is more linear, as opposed to varying the coolant flow rate. With this configuration the coolant flow rate can be maintained very high which reduced the tendency for fouling on the process fluid side of the heat-transfer surface.

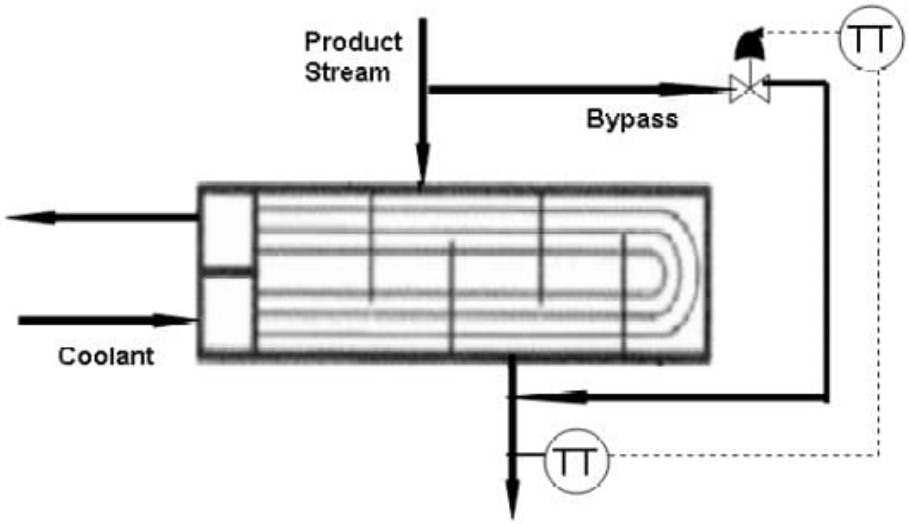


Figure 11. Image modified from Riggs & Karim, shows the common arrangement for temperature control for cooling down a process stream

Example 11.6.1

a. Is an exothermic CSTR easier or harder to control than an endothermic CSTR? What are some reasons for this, and what are a few steps that can be taken to increase control?
b. In Graph 1, why are points A and C stable operating points, while point B is an unstable operating point?

Solution:

(a) An exothermic CSTR is harder to control. The relationship between heat generation and temperature is non-linear for an exothermic system (Graph 1). If the rate of heat generation exceeds the maximum rate of heat removal then the temperature will runaway.

An endothermic CSTR is much easier to control because these reactors are self-regulating and there is no concern for runaway temperatures.

The control of an exothermic CSTR can be increased by reducing the feed flow of the reactants, reducing the concentration of reactants and increasing the ratio of the heat transfer surface area to the reactor volume.

(b)

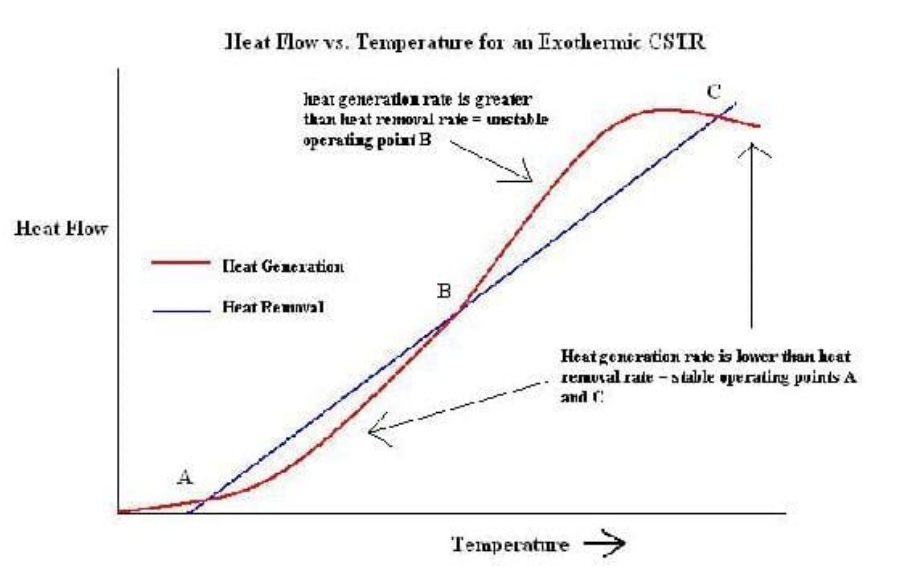


Figure E1-1. Relationship between heat generation or heat removal and temperature for an exothermic system.

Points A and C are stable because if you increase the temperature from this point, the heat generation rate is lower than the heat removal rate. If you increase the temperature from point B, the heat generation rate is higher than that heat removal rate, shich is not desired. If the rate of heat generation is greater than the heat removal then the temperature is rising and could possible runaway. This is an unstable state for the reactor.

Example 11.6.2

Refer to Heat Exchanger Modeling Worked Example 1 problem statement.

After you successfully created your ODE model, you are assigned a partner who will work along with you to put temperature controls on the heat exchanger. Your partner, a new college graduate, seems to remember that in his process controls class, the best way to control a temperature loop for a heat exchanger is to make the flow rate of the coolant stream the manipulated variable. Since you are such a process guru, you know that this is not right. However, the hot shot college graduate just feels that you are over the hill and he knows his stuff.

Vary the coolant flow rate in the Euler's Method heat exchanger to demonstrate to your new partner why it is not a good idea to place the controls on the coolant side flow of a stream. Produce a process gain plot similar to that in the article to show this. Also, find the magnitude of the process gain for three different points in the plot: low, medium and high coolant flow rate, to show process gain goes to zero as the coolant flow rate increases.

Solution

Note that in this case we are dealing with a liquid/liquid heat exchanger and the hot stream is the process stream. The table below summarizes the results.

Controller Flow Rate	Process Stream Exit Temperature
m^3/s	K
0	330
0.05	288.156
0.075	279.61
0.1	274.506
0.2	266.171
0.3	263.374
0.4	262.005
0.5	261.198
0.6	260.667
0.7	260.291
0.8	260.291
0.9	259.798
1	259.626

Table E2-1. Process stream exit temperature as a function of coolant flow rate.

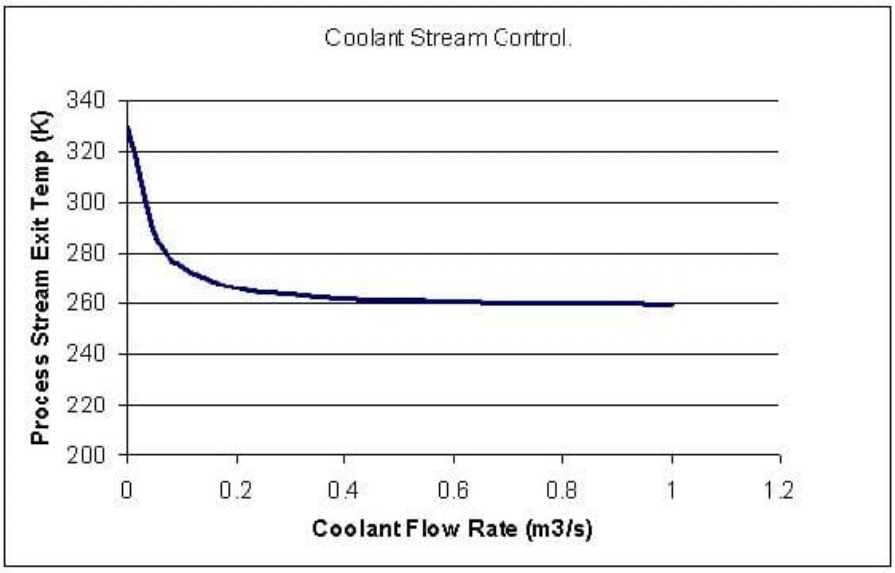


Figure E2-1. A continuous picture of how the magnitude of process gain decreases to zero as the coolant flow rate is increased.

From the graph or table above one can show that the magnitude of process gain decreases as the coolant flow rate increases. This is done as follows;

Magnitude of process gain for;

Low flow rates, choose from 0 to 0.1 m^3 / s.

$$= |330 - 274.50| = 55.5K$$

Medium flow rates, choose (same interval size) from 0.4 to 0.5 m^3 / s.

$$= |262.01 - 261.20| = 0.81K$$

High flow rates, choose (again same interval size) from 0.9 to 1 m^3 / s.

= |259.798 - 259.626| = 0.172 K

Clearly, as the flow rate increases, the magnitude of process gain decreases towards zero. This puts a limit on the amount of control of the heat exchanger if the coolant flow rate is chosen as the MV in the temperature control loop.

Multiple Choice Question 1

What two types of endothermic CSTR temperature control methods are there?

a. steam flow rate and reactor level

b. reactor composition and reactor level

c. steam flow rate and steam pressure

d. reactant flow rate and reactor level

Multiple Choice Question 2

What are temperature controllers commonly used for in distillation columns?

a. To control composition

b. To control reflux ratio

c. To control pressure

d. None of the above

References

- Svrcek, W., Mahoney, D., Young, B.. *A Real-Time Approach to Process Control*. Second Edition. John Wiley and Sons.
- Riggs, James B., Karim, Nazmul M.. *Chemical and Bio-Process Control*. Third Edition Chapter 18. Ferret Publishing.

Contributors and Attributions

- Authors: Danesh Deonarain, Carly Ehrenberger, Georgina Mang, Teresa Misiti

This page titled 11.6: Common Control Loops and Model for Temperature Control is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

11.7: Common control architectures and model for reactors

Reactors are the central focus of many chemical plants. Many parameters must be controlled in a reactor for proper operation. Temperature is of great importance because it affects reaction rates and equilibrium relationships. A major challenge for temperature control is handling the nonlinear nature of temperature inside most reactors. Therefore, it is important to design an effective control architecture in order to ensure optimal operation of the reactor.

This article discusses the common control architectures and topologies in CSTRs. The control architectures are designed based on whether the reactor is endothermic or exothermic. The same concepts introduced in the wiki can be applied to other reactors as well. However, only CSTRs will be discussed for simplicity.

Common Topologies

Here we will introduce a few of the most common control topologies which will be examined in applications to endothermic and exothermic CSTRs below. Please refer to the specific wiki pages for each type for detail.

Feedback and Feed-Forward

Feedback and feed-forward refer to the direction in which the sensor information is transfered to an actuator valve. Feedback control dictates that sensor information is "fed back" to a previous part of the process. For example, the reading from a level sensor of a filling tank can be "fed back" to the valve controlling the input to the tank. Feed-forward control means the sensor information is used to control something downstream from where the reading was taken. For example, the measured flow rate of water going into an evaporator can be used to control the heating coil inside of the evaporator. For detail about these two control topologies, please see the respective pages: Feedback Control Feed-Forward Control.

Ratio Control

Ratio Control is used when the ratio between two measured process variables has an optimal value. In the context of two input streams with optimal flow ratio going into a reactor, one stream is designated as the control stream, and one stream is designated as the wild stream. The wild stream fluctuates, and a valve on the control stream is opened or closed to maintain the ratio between the two stream flows. For detail about ratio control, see Ratio Control.

Cascade Control

Cascade control simply means that instead of one control loop found in simple control topologies where a sensor's measurement directly controls an actuator valve, multiple loops are used so that sensors' measurements can control set points for other controllers. For example, the temperature sensor measurement of process fluid exiting a reactor can be used to modify the set point of the flow controller of steam feeding the heating jacket, which then sets the steam valve. This multiple-loop system eliminates some problems caused by variable-pressure feeds, for instance. For detail about cascade control, see Cascade Control.

Disturbances to CSTRs

There a few very common disturbances that CSTR may be subjected to. When designing a control architecture for a CSTR, you must invertigate the possibility of all of these disturbances, determine the magnitude of each possible disturbance, and address how each will be handled.

- Changes in feed properties
- flow rate
- composition
- temperature
- Changes in enthalpy of heat exchange medium
- Change in heat transfer properties (ex: fouling)

Disturbances to PFRs

Plug Flow Reactors (PFRs) behave differently than CSTRs and will have different properties to consider when designing a control architecture for them. The biggest difference is that the temperature, flows, and compositions will all vary along the reactor. There is danger of exceeding a design limitation in temperature or flow in certain parts of the PFR, so control of the reactor is important along its length. To design a control architecture for a PFR, the following disturbances and changes to the system must be addressed:

- Temperature control in multiple places along reactor (hot spots can easily occur in PFRs)
- Flow control in multiple places along the reactor
- Inlet/outlet pressures
- Feed property disturbances:

- flow rate
- composition
- temperature

- Changes in enthalpy of heat exchange medium
- Change in heat transfer properties (ex: fouling)

Endothermic Reactors

DISCLAIMER: All reactors shown posess two controllers. We realize that in the real-world a reactor would require more than two temperature controllers but we are concerned primarily with the common placement of controllers outside of the reactor in this wiki. Thank you and continue reading.

Endothermic reactors tend to be easier to control than exothermic ones because they are much less prone to runaway. There are two commonly used methods to control an endothermic CSTR. These two methods are differentiated by the variables they manipulate. In the first method, the steam pressure in the reactor jacket is in the manipulated variable, whereas in the second method, the steam flowrate is the manipulated variable. In endothermic reactors, as well as exothermic ones as we will later see, feedback is the most common type of control system used. This is done to ensure that changes are being based on what is actually happening in the reactor and not what is predicted to happen. This is important in the endothermic case to ensure that the proper amount of heat is being added to the reactor to obtain a desired conversion. Feedback is more useful for control since the amount of heat needed can change quickly based on the amount or concentration of reactants being added. Generally the temperature control of the reactor is independent of the reactant feed, so the system needs a way to adjust to changes within reactant feed, and this is accomplished by using feedback control in the streams that control the temperature (i.e. the steam feed stream).

Controlled by Steam Pressure

Heat duty can be more effectively removed with the use of steam pressure as the manipulated variable because changes in the heat duty of the reactor will cause an immediate change in the steam pressure. By using the steam pressure, it also linearizes the temperature control system which is not the case when the steam flowrate is the manipulated variable. Using the steam pressure as the manipulated variable does not provide a direct measurement of the heat load, the amount of heat needed by the reactor. If a direct measurement is required, a steam flow indicator can be installed which would allow for the direct measurement of the applied heat load. An added benefit, which those of you who have worked with the temperature controller in ChE 460 might have noticed, is that often times the pressure of steam being provided by the utilities plant is variable. By using the steam pressure as the manipulated variable the control system will automatically adjust to this type of disturbance from the source. If a large change in steam pressure were to occur, the pressure controller would measure this change, and the control scheme would intake the change in steam pressure and translate this into an appropriate change in the valve setting.

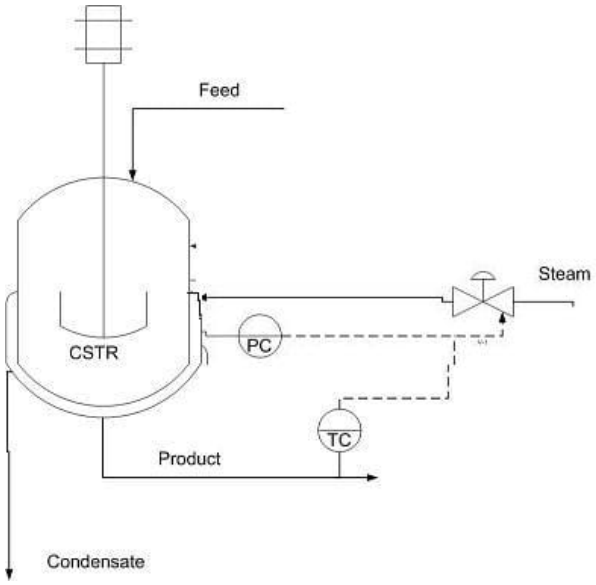


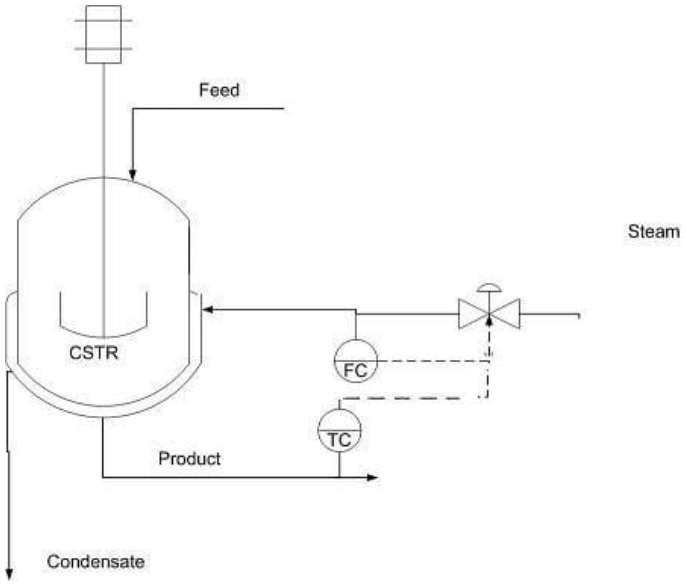
Figure: Image Modified From: Riggs, James B., Karim, Nazmul M.. Chemical and Bio-Process Control. Third Edition Chapter 18. Ferret Publishing.

As can be seen in the figure above, when steam pressure is used as the manipulated variable the control system is run in a feedback mode. One of two things can control the amount of steam fed into the jacket. The first is the temperature of the product, and the second is the steam pressure in the jacket. As mentioned above, a change in the heat duty required by the reactor will quickly change the steam pressure which is why it is commonly used over the steam flowrate.

Notice that there are two controllers that are responsible for adjusting the steam valve. The pressure controller is sensitive to changes in the heat duty required by the reactor and is used to adjust the steam according to the needs of the temperature controller on the product stream. This is a prime example of cascade control. The temperature of the product stream would output a setpoint to the pressure controller for the amount of steam needed to attain the desired temperature setpoint. The pressure controller would then communicate to the valve what needs to be done in order to achieve this temperature setpoint based on the steam pressure.

Controlled by Steam Flowrate

The second method for controlling a endothermic CSTR is by manipulating the steam flow rate. Using flow rate as the manipulated variable makes the control system prone to changes in heat load and changes in the enthalpy of supplied steam. These changes require direct action of the temperature control system. One advantage of using flowrate as the manipulated variable is that heat load is directly measured, and thus, conversion is directly measured.



As can be seen in the figure above the system works in a feedback mode when the steam flowrate is used as the manipulated variable. The temperature of the product stream is the primary factor in adjusting the amount of steam fed into the system, and therefore this system is less responsive to changes in the amount of heat duty required for the process.

Note here that there are again two controllers used to adjust the steam valve. This setup is similar to the previous case, but now the flow controller is the "slave" controller to the temperature controller. The cascade control scheme is again at work!

Control Architecture for Endothermic Reactor		
Manipulated Variable	**Advantage**	**Disadvantage**
Steam pressure	More responsive to changes in heat duty	No direct measurement of heat load
	Linearizes temperature control system	
Steam flowrate	Heat load and conversion directly measured	Less responsive to changes in heat duty

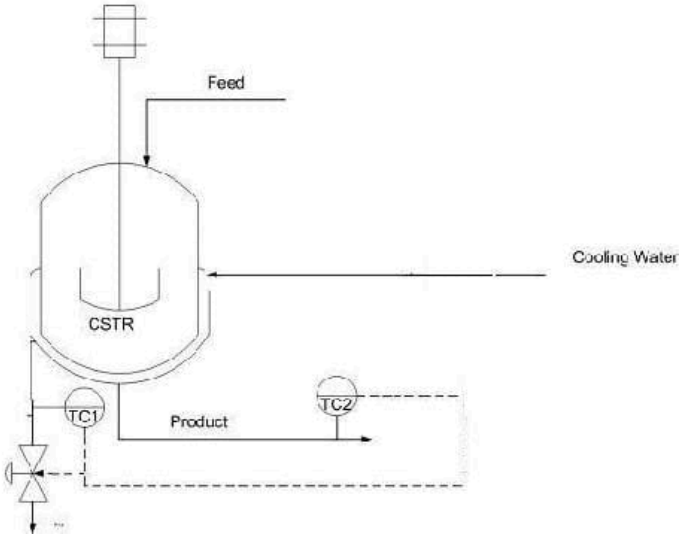
Exothermic Reactors

Exothermic reactors are harder to control because safety is dependent on heat removal. There are two common control architectures for an exothermic CSTR based on the temperature of the coolant. The first method uses the outlet temperature of the coolant as the manipulated variable, while the second method uses the inlet temperature of the coolant. Just like the endothermic reactors, feedback control is very commonly used when controlling the temperature of exothermic reactors. The reasoning for feedback control in the exothermic case is slightly different but similar to that of endothermic reactors. In endothermic reactors we wanted to ensure that the proper amount of heat was being supplied to our reactor, and if the amount needed changed; we needed a way to adjust our temperature accordingly. In the exothermic case we need to ensure that we are removing the proper amount of heat, not only to ensure an optimal reaction temperature, but also to prevent a runaway reaction. Therefore, we need to know what is actually going on within the reactor instead of forecasting possible temperature changes within it. Cascade control is also used. In this case the temperature of the product stream outputs a setpoint to the temperature controller whether it be located on the outlet or inlet coolant feed. The "slave" controller can then adjust the valve position based on the temperature of the coolant water.

Controlled by Outlet Coolant Temperature

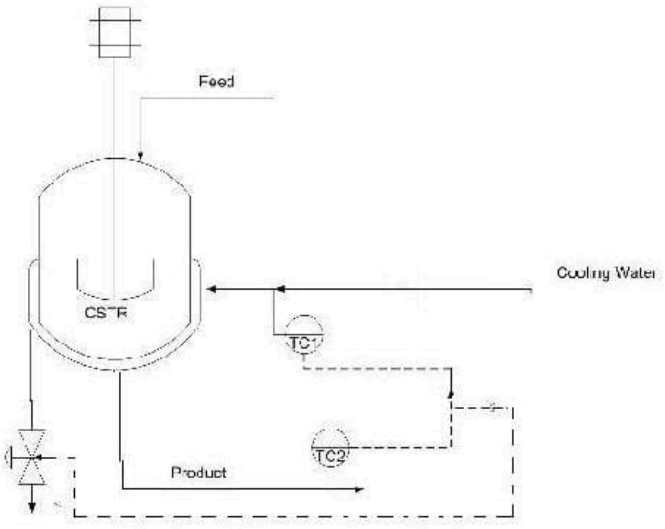
The figure above shows the control architecture for CSTR reactor temperature controlled by outlet coolant temperature in a feedback mechanism. There is a temperature control on both the product stream and outlet coolant temperature. Both of the

temperature controls are used to control the valve on the inlet coolant stream. The advantage of this setup is that it responds faster to fouling on heat-transfer surfaces than the setup with the temperature controller on the inlet. This is because in order for the inlet temperature to adjust to fouling, the fouling must first affect the temperature of the product stream. Therefore the fouling has less direct effect on the controller, but in this case fouling will immediately affect the temperature of the exiting cooling water. However, this setup has the disadvantage of responding slower to changes in the inlet coolant temperature.



As can be seen in the above diagram you will notice that again cascade control is utilized in this particular system. The temperature sensor on the product stream provides information on whether the stream needs to be cooled more or less and outputs a setpoint to the temperature controller on the recycle stream. This controller can then take into account the temperature of the recycled coolant water and make an adjustment to the amount of fresh coolant water that is added to the system.

Controlled by Inlet Coolant Temperature



As seen in the figure above, the control configuration for CSTR reactor temperature is controlled by inlet coolant temperature in a feedforward mechansim. There are temperature controls on both the product stream and inlet coolant temperature; both of the controls are used to control the valve on the inlet coolant stream. One advantage of this setup is that it responds faster to changes in inlet coolant temperature. A disadvantage is that it responds slower to fouling on heat-transfer surfaces.

More on Exothermic Reactors

How do you decide which exothermic reactor control architecture to use? It depends on whether a faster response to fouling on heat-transfer surfaces or a faster response to changes in inlet coolant temperature is more important.

In addition to stabilizing the temperature within an exothermic reactor, it is often a goal to maximize production rate. This can be achieved by placing a valve on the feed stream. As the temperature control adjusts the cold water valve, the cold water valve position is then communicated to the feed valve and adjusts it.

Cascade control is again utilized. In this case the secondary or "slave" controller is located in a different location, which is the only difference between the two scenarios.

Control Architecture for Exothermic Reactor		
Manipulated Variable	**Advantage**	**Disadvantage**
Outlet coolant temperature	Faster response to fouling on heat-transfer surfaces.	Slower response to changes in inlet coolant temperature
Inlet coolant temperature	Faster response to changes in inlet coolant temperature	Slower response to fouling on heat-transfer surfaces

Example 11.7.1

As part of a community outreach program, your company is sponsoring a haunted house. The haunted house is being constructed on site in an old, dilapidated warehouse located on the plant property. You have been put in charge of designing a scene in which a witch is creating a witch's brew. It has been proposed that you use a reaction between "Witch's Brew Deluxe", a commercial compound used in the haunted house industry, dry ice, and water.

Dry Ice + "Witch's Brew Deluxe" + Water --> Bubbles + Mist

The manufacturer of "Witch's Brew Deluxe" supplies optimal conditions for real life cauldron imitation at 30°C water temperature. Temperatures in excess of this result in a surplus of cauldron bubbles and mist which impair the visual experience. Temperatures below this do not allow for large enough reaction rates and do not produce any steam or bubbles.

The reaction is highly endothermic, so a supply of steam has been diverted from the main plant to supply heat to the reactor. However, due to the great distance this steam is traveling, the supply steam pressure is highly variable. An uncovered, jacketed CSTR is available for use as a cauldron. A mixture of water and "Witch's Brew Deluxe" is constantly fed into the reactor, and solid dry ice is manually feed by the operator.

What control scheme is most suitable for temperature control and why? Specifically, what should you use as the manipulated variable?

Solution

Because the steam supply pressure will vary greatly, a steam pressure control architecture is preferred. By monitoring the jacket pressure, changes in steam supply pressure will immediately be observed and corrected for.

Example 11.7.2

You have just recently taken a job offer from a haunted engineering firm run by a group of witches located in Salem, Massachusetts. As the new guy on the job they would like your input on which variable they should treat as the manipulated variable for the reactor temperature controller. Their reactor is an exothermic CSTR. The following reaction is what takes place within this CSTR.

3 frog's legs + 2 cow hooves --> 4 moles of magic elixir

The magic elixir is generally used as a wart remover but also shows promise as an antacid.

The head witch at the plant tells you that significant fouling occurs on the heat transfer surfaces because the city only allows them access to dirty water with high mineral content because they are witches and generally considered substandard citizens. She (the head witch) also points out that they have recently developed a spell to maintain the inlet temperature and pressure of their coolant water.

Which variable should be the manipulated one and why?'

Solution

The correct manipulated variable for the reactor temperature controller would be the outlet coolant temperature. This allows for a quicker response to fouled heat-transfer surfaces. Since the head witch told you that inlet temperature and pressure are constant, the configuration's weakness to changes in these two parameters can be overlooked.

Exercise $11.7.2A$

Which type of REACTION is generally much easier to control?

a. Endothermic
b. Distillation of alcohol
c. Cold Fusion
d. Exothermic

Answer available at the end of this section.

Exercise $11.7.2B$

By using the _____ as the manipulated variable, it linearizes the temperature control system which is not the case when the steam flowrate is the manipulated variable.

a. feed flow rate
b. steam pressure
c. product composition
d. rate of mixing

References

- Riggs, James B. and Karim, M. Nazmul. *Chemical and Bio-Process Control*. Chapter 18: Control Case Studies. Third Edition. Ferret Publishing.
- Svrcek, William Y., Mahoney, Donald P., and Young, Brent R. *A Real-Time Approach to Process Control*. Chapter 7: Common Control Loops. Second Edition. John Wiley and Sons, Ltd.

Answers

Exercise $11.7.2A$
endothermic

This page titled 11.7: Common control architectures and model for reactors is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Brandon Quigley, Brian McQuillan, Crystal Miranda, & John Zhang via source content that was edited to the style and standards of the LibreTexts platform.

12: MULTIPLE INPUT, MULTIPLE OUTPUT (MIMO) CONTROL

12.1: Determining if a system can be decoupled
12.2: MIMO control using RGA
12.3: MIMO using Model Predictive Control
12.4: Neural Networks for automatic model construction
12.5: Understanding MIMO Control Through Two Tanks Interaction

This page titled 12: Multiple Input, Multiple Output (MIMO) Control is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

12.1: Determining if a system can be decoupled

Authors: Wen Chung Lim, James Bennett, Jamila Grant, Ajay Bhasin

Introduction

A system of inputs and outputs can be described as one of four types: SISO (single input, single output), SIMO (single input, multiple output), MISO (multiple input, single output), or MIMO (multiple input, multiple output).

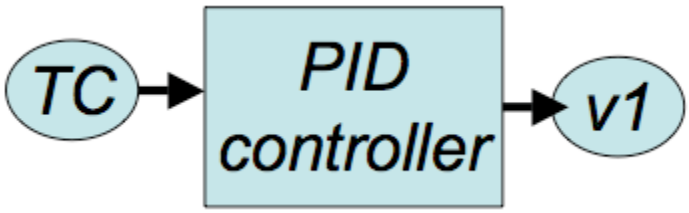
Multiple input, multiple output (MIMO) systems describe processes with more than one input and more than one output which require multiple control loops. Examples of MIMO systems include heat exchangers, chemical reactors, and distillation columns. These systems can be complicated through loop interactions that result in variables with unexpected effects. Decoupling the variables of that system will improve the control of that process.

An example of a MIMO system is a jacketed CSTR in which the formation of the product is dependent upon the reactor temperature and feed flow rate. The process is controlled by two loops, a composition control loop and a temperature control loop. Changes to the feed rate are used to control the product composition and changes to the reactor temperature are made by increasing or decreasing the temperature of the jacket. However, changes made to the feed would change the reaction mass, and hence the temperature, and changes made to temperature would change the reaction rate, and hence influence the composition. This is an example of loop interactions. Loop interactions need to be avoided because changes in one loop might cause destabilizing changes in another loop. To avoid loop interactions, MIMO systems can be decoupled into separate loops known as single input, single output (SISO) systems. Decoupling may be done using several different techniques, including restructuring the pairing of variables, minimizing interactions by detuning conflicting control loops, opening loops and putting them in manual control, and using linear combinations of manipulated and/or controlled variables. If the system can't be decoupled, then other methods such as neural networks or model predictive control should be used to characterize the system.

There are two ways to see if a system can be decoupled. One way is with mathematical models and the other way is a more intuitive educated guessing method. Mathematical methods for simplifying MIMO control schemes include the relative gain array (RGA) method, the Niederlinski index (NI) and singular value decomposition (SVD). This article will discuss the determination of whether a MIMO control scheme can be decoupled to SISO using the SVD method. It will also discuss a more intuitive way of decoupling a system using a variation of the RGA method.

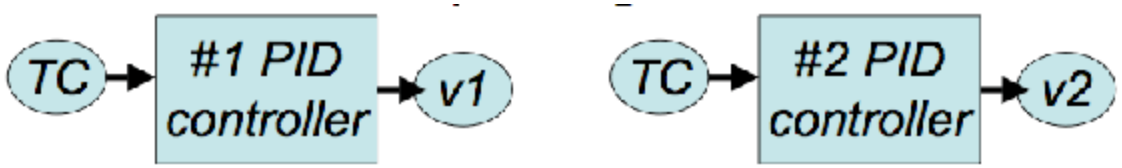
Definitions of Input and Output System Types

1. **SISO**- Single Input, Single Output



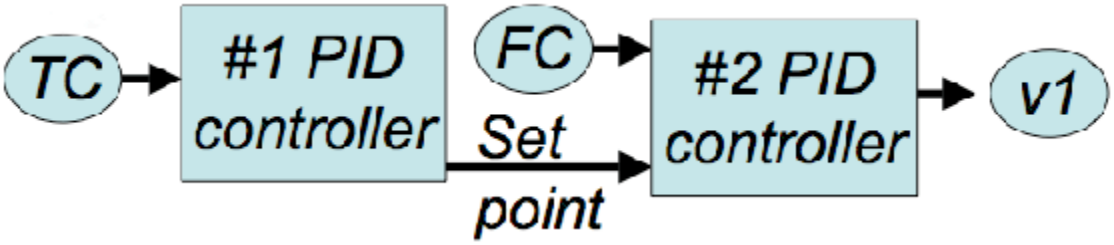
These systems use data/input from one sensor to control one output. These are the simplest to design since they correspond one sensor to one actuator. For example, temperature (TC) is used to control the valve state of v1 through a PID controller.

2. **SIMO**- Single Input, Multiple Output



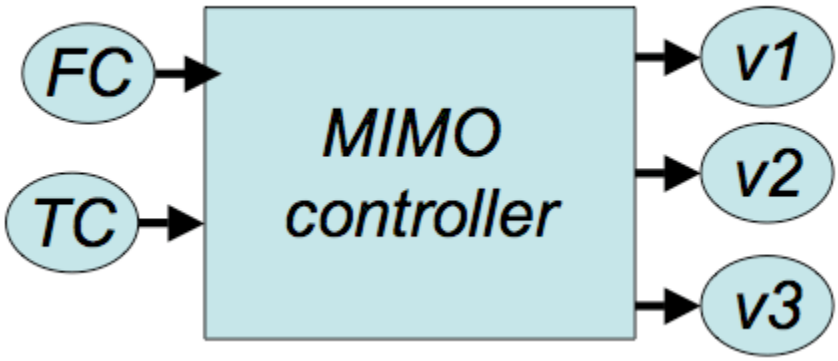
These systems use data/input from one sensor to control multiple outputs. For example, temperature (TC) is used to control the valve state of v1 and v2 through PID controllers.

3. **MISO**- Multiple Input, Single Output



These systems use data/input from multiple sensors to control one ouput. For example, a cascade controller can be considered MISO. Temperature (TC) is used in a PID controller (#1) to determine a flow rate set point i.e. FCset. With the FCset and FC controller, they are used to control the valve state of v1 through a PID controller (#2).

4. **MIMO**- Multiple Input, Multiple Output



These systems use data/input from multiple sensors to control multiple outputs. These are usually the hardest to design since multiple sensor data is integrated to coordinate multiple actuators. For example, flow rate (FC) and temperature (TC) are used to control multiple valves (v1, v2, and v3). Often, MIMO systems are not PID controllers but rather designed for a specific situation.

Singular Value Decomposition

Singular value decomposition is a useful mathematical way to determine if a system will be prone to interactions due to sensitivity problems that result from small errors in process gains. It is a matrix technique that determines if a system is able to be decoupled.

Two input two output system

The SVD method starts with the steady state gain matrix, which is obtained from the relative gain array (RGA) method for a given 2 input-2 output control loop system:

$$[G] = \begin{bmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{bmatrix},$$

Using [G], we obtain the eigenvalues for the system. This can be done through two methods; either the eigenvalues can be obtained through numerical software as described in the Eigenvalues and Eigenvectors section or the eigenvalues can be calculated by hand. The hand calculations are shown here. [G] gives the following values for the system, which have been derived from theoretical work and are used to find the eigenvalues:

$b = g_{11}^2 + g_{12}^2$ ----------(1)
$c = g_{11}g_{21} + g_{12}g_{22}$---------(2)
$d = g_{21}^2 + g_{22}^2$----------(3)

b, c, and d are parameters created for this method that do not have a direct physical meaning. Using the formulas below with the values of b, c, and d gives the eigenvalues for the system:

$\lambda_1 = s_1^2 = \dfrac{b + d + \sqrt{(b-d)^2 + 4c^2}}{2}$ ------(4)

$\lambda_2 = s_2^2 = \dfrac{bd - c^2}{s_1^2}$ ------(5)

s_1 and s_2 are the positive square roots of the respective eigenvalues. The condition number CN is defined as the ratio of the larger of the two values to the smaller value:

CN = $\dfrac{s_1}{s_2}$ if $s_1 \geq s_2$, or CN = $\dfrac{s_2}{s_1}$ if $s_2 \geq s_1$

The greater the CN value, the harder it is for the system in question to be decoupled. As a rule of thumb, a system with a CN number of more than 50 is impossible to decouple. Such a system has manipulated variables which all have significant and similar relative impact on the controlled variables they affect and hence the control equations for the system cannot be restructured to obtain a simpler control system. In the worst case scenario for a two input two output system, using the exact same control variable twice would give a CN number of infinity, because both control variables would have the same effect on the controlled variables. An ideal system would have a CN number of one, where each control variable controls a single distinct output variable.

MIMO systems with two or more inputs and outputs

To use this method, it is necessary to know how to create a transpose matrix. A transpose matrix is created by writing the columns of the original matrix as the rows of the transpose matrix. Mathematically, the transpose matrix is denoted by $[A]^T$, where [A] is the original matrix. An example is shown here: $\begin{bmatrix} i & j \\ k & l \end{bmatrix}^T = \begin{bmatrix} i & k \\ j & l \end{bmatrix}$

To use singular value decomposition on systems with two or more inputs and outputs, the m x m steady state gain matrix [G] of the form

$$\([G] = \begin{bmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,m} \\ g_{2,1} & g_{2,2} & \cdots & g_{2,m} \\ \vdots & & & \\ g_{n,1} & g_{n,2} & \cdots & g_{n,m} \end{bmatrix}\)$$

is broken down into the form:

$$\(G = U\Sigma V^\top\)$$

where:

- U is an m x m matrix , the column vectors of which are the unit eigenvectors of the m x m matrix GG^T,G^T being the transpose matrix of G.
- V is an m x m matrix, the column vectors of which are the unit eigenvectors of the m x m matrix G^TG.
- Σ is an m x m diagonal matrix containing singular values, where the values are arranged in descending order from the top left most corner to the bottom right corner.

$\Sigma = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix}$ for a 2 input 2 output system if $s_1 \geq s_2$, or $\Sigma = \begin{bmatrix} s_2 & 0 \\ 0 & s_1 \end{bmatrix}$ for a 2 input 2 output system if $s_2 \geq s_1$.

The CN number can then be determined by the ratio of the largest value in Σ to the smallest value. Notice in the previous two input two output example, there was no need to calculate the eigenvalue matrix Σ for a two by two matrix, since the eigenvalues can be calculated directly with the parameters b,c, and d.

For more information on how to derive the three matrices U, Σ and V by hand, refer to the Eigenvalues and Eigenvectors section.

Mathematica can also be used to find U, Σ and V for a given m x m Matrix G. This is accomplished by typing in the command

Map[MatrixForm, {u, w, v}=SingularValueDecomposition[G]],

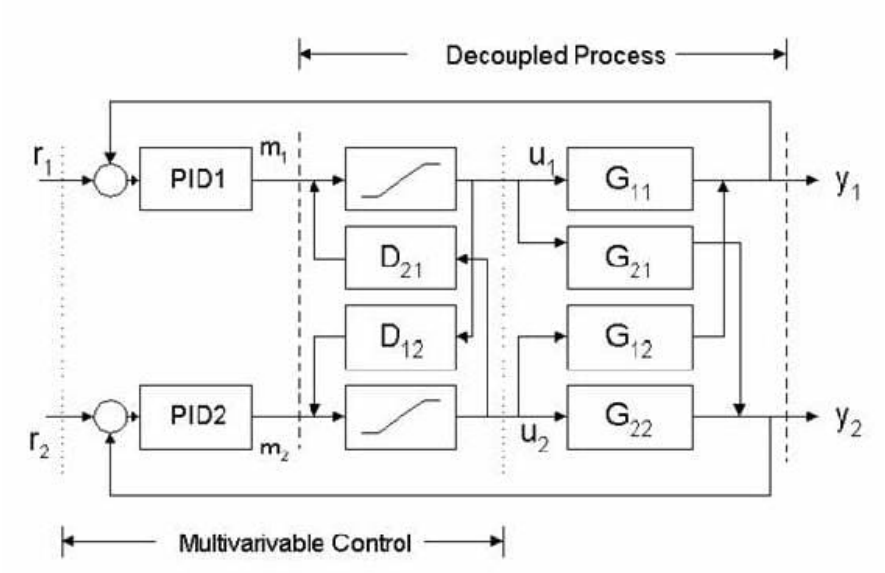
giving U, Σ and V in that order.

Intuitive decoupling using the RGA

With this technique, the first step is to construct the RGA by placing all of the controllers in manual, changing the value of the output manually, and then recording the final value of the controller after making the change. This will give the gain matrix, which is then multiplied by its inverse to give the relative gain matrix. The goal here is to pair measurements and control valves so that the relative gain elements are close to 1, and all other combinations are close to zero. A values between 0 and 1 tells that the net effect of the other loops on this pairing is to change the measurement signal in the same direction as the control valve signal (cooperation). A value greater than one means that the measurement signal will go in the opposite direction of the control valve signal (conflict). Finally a relative gain element less than zero means that the net effect is conflict, but here the gain will change sign, not the measurement signal. Therefore, if a loop is stable when the other loops are in manual, it will probably not be stable when the other loops are in automatic, and vice versa. To intuitively decouple this system then would require a guess and check of different pairings, with the goal being to get the pairings to a relative gain element close to one. If this is not possible and no pairings give a value close to one, then the system will need to be decoupled using decoupling control.

Decoupling a system using decoupling control

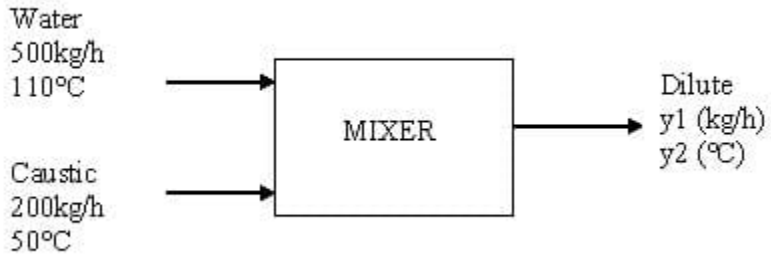
The goal of decoupling control is to eliminate complicated loop interactions so that a change in one process variable will not cause corresponding changes in other process variables. To do this a non-interacting or decoupling control scheme is used. In this scheme, a compensation network called a decoupler is used right before the process. This decoupler is the inverse of the gain array and allows for all measurements to be passed through it in order to give full decoupling of all of the loops. This is shown pictorially below.



For a simple jacketed reactor example, if changing the flow rate of reactant to obtain a desired composition also increases the reactor temperature by five degrees, the decoupler would decrease the temperature of the jacket to decrease the temperature inside the reactor by five degrees, offsetting the original change. This jacket temperature decrease also changes the composition inside the reactor, so a properly designed controller accounts for both effects in the control scheme to obtain the desired composition without changing the reactor temperature.

Worked out Example 1

Hot water at 110°C is pumped into a mixer at 500 kg/h. The water is used to dilute a concentrated caustic substance at 50°C flowing in at 200 kg/h. Below is a diagram of the process. Check, by the singular value decomposition method, to see if decoupling of the system is feasible. Assume that the heat capacities of both streams are the same with no heat of mixing.



$boldSolution:$

From the question you know:

T_1 = 383.15 K

T_2 = 323.15 K

m_1 = 500 kg/h

m_2 = 200 kg/h

y_1 = outgoing mass flow rate (kg/h)

y_2 = temperature of outflowing stream (K)

From mass balances you know:

$$y_1 = (m_1 + m_2)$$

$$y_2 = \frac{T_1 m_1 + T_2 m_2}{y_1}$$

So:

y_1 = 700 kg/h

y_2 = 366.01 K

The steady state gain matrix (G) needs to be calculated as explained in the RGA method:

$$G = \begin{bmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{bmatrix}$$

$$g_{1,1} = \left.\frac{\partial y1}{\partial m1}\right|_{m2=constant} = \frac{\partial}{\partial m1}(m1 + m2) = 1$$

$$g_{2,1} = \left.\frac{\partial y2}{\partial m1}\right|_{m2=constant} = \frac{\partial}{\partial m1}\left(\frac{T1m1 \mid T2m2}{m1 + m2}\right) = \left[\frac{(T1 \quad T2)m2}{(m1 + m2)^2}\right]$$

$$g_{1,2} = \left.\frac{\partial y1}{\partial m2}\right|_{m1=constant} = \frac{\partial}{\partial m2}(m1 + m2) = 1$$

$$g_{2,2} = \left.\frac{\partial y2}{\partial m2}\right|_{m1=constant} = \left[\frac{(T2 - T1)m1}{(m1 + m2)^2}\right]$$

By plugging in all known values, the steady state gain matrix becomes:

$$G = \begin{bmatrix} 1 & 1 \\ 0.0245 & -0.0612 \end{bmatrix}$$

From here you can solve for the values b, c, d λ_1, λ_2, s_1 and s_2 by using the equations (1) through (5) in the wiki:

b = $g_{11}^2 + g_{12}^2 = 1^2 + 1^2 = 2$

c = $g_{11}g_{21} + g_{12}g_{22}$ = (1)(.0245) + (1)(– .0612) = -0.0367

d = $g_{21}^2 + g_{22}^2 = .0245^2 + - .0612^2$ = 0.004346

$$\lambda_1 = \frac{b+d+\sqrt{(b-d)^2+4c^2}}{2} = \frac{2+.004346+\sqrt{(2-.004346)^2+4(-.0367)^2}}{2} = 2$$

$$\lambda_2 = \frac{bd-c^2}{\lambda_1} = \frac{(2)(.004346)-(-.0367)^2}{2^2} = 0.003673$$

$$s_1 = \sqrt{\lambda_1} = \sqrt{2} = 1.414$$

$$s_2 = \sqrt{\lambda_2} = \sqrt{.003673} = 0.0606$$

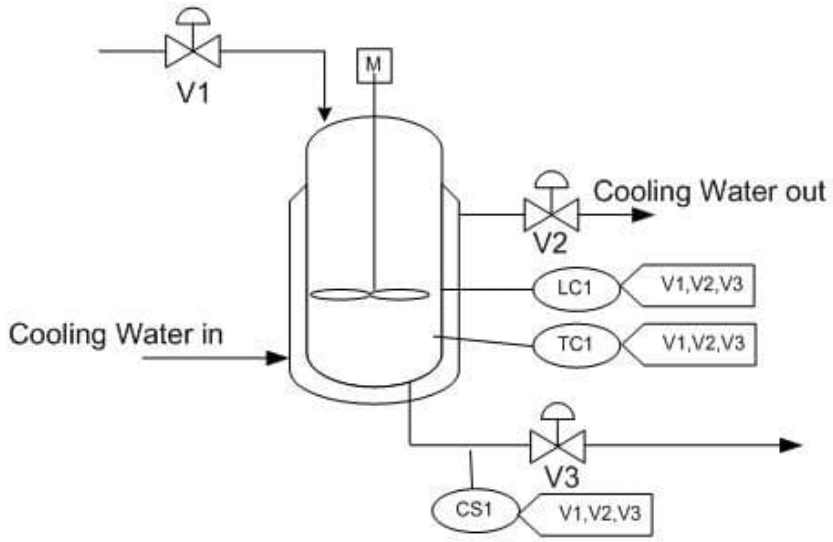
$$\Sigma = \begin{bmatrix} 1.414 & 0 \\ 0 & 0.0606 \end{bmatrix}$$

$$CN = \frac{s_1}{s_2} = 23.33$$

CN is less than 50, therefore decoupling for this system is feasible.

Worked out Example 2

A jacketed reactor with a single feed stream (Stream 1) is monitored by 3 proportional (P) controllers as indicated in the figure.



These controllers all act as inputs to each of 3 valves according to the relationships:

$$V_1 = \frac{1}{3}C_1 + \frac{7}{15}C_2 + \frac{1}{5}C_3$$

$$V_2 = \frac{4}{15}C_1 + \frac{4}{15}C_2 + \frac{1}{3}C_3$$

$$V_3 = \frac{1}{5}C_1 + \frac{4}{15}C_2 + \frac{1}{3}C_3$$

Where:

$C_1 = K_{c1}e_1(t) + b_1$, C_1 is the output from the level controller

$C_2 = K_{c2}e_2(t) + b_2$, C_2 is the output from the temperature controller

$C_3 = K_{c3}e_3(t) + b_3$, C_3 is the output from the composition sensor

Is it feasible to decouple this system?

$$boldSolution:$$

The gain array is calculated according to the RGA method and is:

Input the following commands in *Mathematica* for SVD analysis:

G =

ParseError: EOF expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

G = G*1.0 Map[MatrixForm, {u,w,v}=SingularValueDecomposition[G]]

The output of the SVD decomposition is the following:

$$U = \begin{bmatrix} -0.657043 & -0.75205 & 0.052116 \\ -0.552104 & 0.432979 & -0.712538 \\ -0.513299 & 0.496941 & 0.699696 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} 0.897525 & 0 & 0 \\ 0 & 0.193259 & 0 \\ 0 & 0 & 0.0392887 \end{bmatrix},$$

$$V = \begin{bmatrix} -0.522439 & -0.18542 & -0.832273 \\ -0.658174 & -0.532849 & 0.531864 \\ -0.542094 & 0.825646 & 0.156343 \end{bmatrix}$$

The condition number is calculated from the highest and lowest values of the ∑ matrix:

$$CN = \frac{0.897525}{0.0392887} = 22.844$$

$CN < 50$, therefore the system may be decoupled.

As a more complicated problem, consider how this problem would change if the controllers used a cascade control style. For a P controller, the output may be similar to:

$C_2 = K_{c2}e_2(t) + b_2 = K_{c2} * (LC_{actual} - f(C_1)) + b$

What if we used a PID controller instead?

Multiple Choice Question 1

The CN, or condition number

a. Is the ratio of the smaller number (s_2) to the larger number (s_1)
b. Determines the feasibility of decoupling a system.
c. Is the unit eigenvectors of the m x m matrix G^TG
d. Is always less than 50.

Answer: b

Multiple Choice Question 2

For MIMO systems

a. Control loops are isolated
b. Each controlled variable is only manipulated by one variable
c. Decoupling the system makes it more complicated
d. Manipulated variables may affect several controlled variables

Answer: d

Sage's Corner

www.youtube.com/v/j-heeeVmEAM
Unnarated Powerpoint Presentation

References

- Mahoney, D.P., Svrcek, W.Y., and Young, B.R. (2006) A Real-Time Approach to Process Control, 2nd Edition. John Wiley & Sons, Ltd. ISBN 0470025336
- Gwak, K.W., Masada, G.Y. (2005) "Structural Analysis and Optimization of Nonlinear Control Systems Using Singular Value Decomposition". American Society of Mechanical Engineers Journal of Dynamic Systems, Measurement and Control, Vol. 127, No. 1, pp. 105-113, March 2005.
- McMillan, Gregory K. (1983) Tuning and Control Loop Performance. Instrument Society of America. ISBN 0-87664-694-1.
- Tham, M.T. (1999). "Multivariable Control: An Introduction to Decoupling Control". Department of Chemical and Process Engineering, University of Newcastle upon Tyne. Available online at: lorien.ncl.ac.uk/ming/mloop/mloop.pdf.
- Lee, Jay H., Choi, Jin Hoon, and Lee, Kwang Soon. (1997). "3.2 Interaction and I/O Pairing". Chemical Engineering Research Information Center. Available online at: www.cheric.org/education/practical/mpc16473.pdf.

This page titled 12.1: Determining if a system can be decoupled is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

12.2: MIMO control using RGA

Introduction

Single variable **I**nput or **S**ingle variable **O**utput (SISO) control schemes are just one type of control scheme that engineers in industry use to control their process. They may also use MIMO, which is a **M**ulti-**I**nput-**M**ulti-**O**utput control scheme. In MIMO, one or more manipulated variables can affect the interactions of controlled variables in a specific loop or all other control loops. A MIMO control scheme is important in systems that have multiple dependencies and multiple interactions between different variables- for example, in a distillation column, where a manipulated variable such as the reflux ratio could directly or indirectly affect the feed flow rate, the product composition, and the reboiler energy. Thus, understanding the dependence of different manipulated and controlled variables in a MIMO control scheme could be extremely helpful in designing and implementing a control scheme for a process.

One method for designing and analyzing a MIMO control scheme for a process in steady state is with a **R**elative **G**ain **A**rray (RGA). RGA is useful for MIMO systems that can be decoupled (see the article about determining if a system can be decoupled). For systems that cannot be decoupled, model predictive control or neural networks are better choices of analysis tool than RGA. A good MIMO control scheme for a system that can be decoupled is one that can control a process variable without greatly affecting the other process variables. It must also be stable with respect to dynamic situations, load changes, and random disturbances. The RGA provides a quantitative approach to the analysis of the interactions between the controls and the output, and thus provides a method of pairing manipulated and controlled variables to generate a control scheme.

What is RGA?

Relative **G**ain **A**rray is an analytical tool used to determine the optimal input-output variable pairings for a multi-input-multi-output (MIMO) system. In other words, the RGA is a normalized form of the gain matrix that describes the impact of each control variable on the output, relative to each control variable's impact on other variables. The process interaction of open-loop and closed-loop control systems are measured for all possible input-output variable pairings. A ratio of this open-loop 'gain' to this closed-loop 'gain' is determined and the results are displayed in a matrix.

$$RGA = \Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2n} \\ \vdots & & & \\ \lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn} \end{bmatrix}$$

The array will be a matrix with one column for each input variable and one row for each output variable in the MIMO system. This format allows a process engineer to easily compare the relative gains associated with each input-output variable pair, and ultimately to match the input and output variables that have the biggest effect on each other while also minimizing undesired side effects.

Understanding the Results of the RGA

- The closer the values in the RGA are to 1 the more decoupled the system is
- The maximum value in each row of the RGA determines which variables should be coupled or linked
- Also each row and each column should sum to 1

- Example

The table below includes the RGA results. The values highlighted in red are the maximum values in the row. These values indicated that the valve for that row should be used to control the variable that is listed in the corresponding column.

RGA

P	T	F	Ca	Cb	Valves
0.907911	0.015328	0.021940	0.015297	0.040420	V1
0.084549	0.228626	-0.0313343	0.230685	0.487474	V2
0.001492	0.139518	-0.002314	0.763468	0.097835	V3
-0.027496	0.584721	0.029974	-0.025906	0.438707	V4
0.037542	0.031807	0.981733	0.016856	0.067439	V5

V1->P
V2 >Cb
V3->Ca
V4->T
V5->F

Calculating RGA

There are two main ways to calculate RGA:

1. Experimentally determine the effect of input variables on the output variables, then compile the results into an RGA matrix.
2. Use a steady-state gain matrix to calculate the RGA matrix.

Method **(1)** should be used when it is possible to carry out the experiments as detailed in the **Calculating RGA with Experiments** section. This method will generally yield the most accurate RGA matrix of the system because it is based on actual data taken from the operating control system. If performing these extensive experiments is not possible, method **(2)** can be used. If a process model is available, method **(2)** can be used with no experimental data. If there is no process model available, some experimental data must be taken (though less extensively than in method **(1)**) and used in conjunction with method **(2)**.

Method 1: Calculating RGA with Experiments

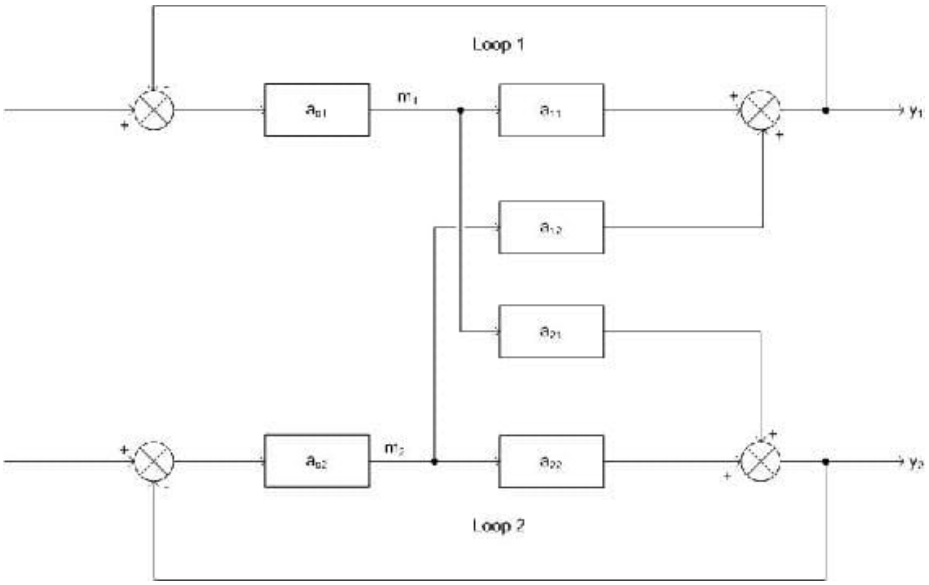
This method of calculating the RGA can be used when it is possible to run experiments on each of the input-output pairings. Below is a step-by-step explanation of how to experimentally compile the RGA for a simple MIMO system.

The simplest MIMO system is one that has two inputs and two outputs. Remember that by definition, a change in one of the inputs in a MIMO system will change *both* of the outputs. This system can be expressed mathematically as written below.

$$y_1 = a_{11}m_1 + a_{12}m_2$$

$$y_2 = a_{21}m_1 + a_{22}m_2$$

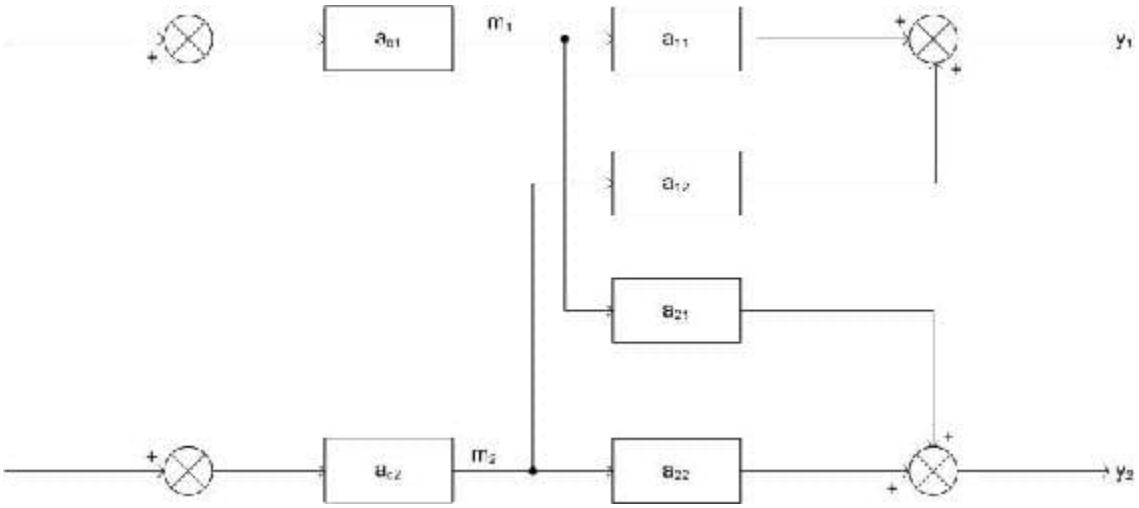
Where y_i is the output for loop *i*, the *m* variables are the inputs for each loop, and the *a* variables are the transfer functions. It also helps to see this system as a control diagram, as shown here.



Clearly, both y_1 and y_2 are affected by both inputs (m_1 and m_2). Because of this, there are two choices for control. y_1 could be controlled by m_1, or y_1 could be controlled by m_2. y_2 would then be controlled by the input that is not controlling y_1. For a MIMO system with n input variables, there are $n!$ possible control configurations. The question is: Which configuration will give you the best control? The answer can be determined by finding the relative gain array (RGA). There are a few ways to determine the RGA, one of which is by doing two experiments, repeated for every possible configuration. The RGA can be assembled from the data collected. These experiments are explained below using the above drawing and the configuration where m_1 controls y_1.

Experiment 1

For the first experiment, the objective is to observe the effect of m_1 on y_1. To do this, *all* of the loops must be open. In other words, the feedback loop is removed and the system is run manually with no control. This configuration is shown below.



Now, since the system is under manual control, it is possible to introduce a step change Δm_1 while keeping m_2 constant. Because m_2 is held constant, the change in y_1, Δy_1, can be attributed entirely to m_1. Although a change in m_1 will also affect y_2, it is the relationship between m_1 and y_1 that must be observed at this point. Then, the gain of y_1 with respect to m_1 with all loops open is defined as g_{11}. The calculation of g_{11} is shown below.

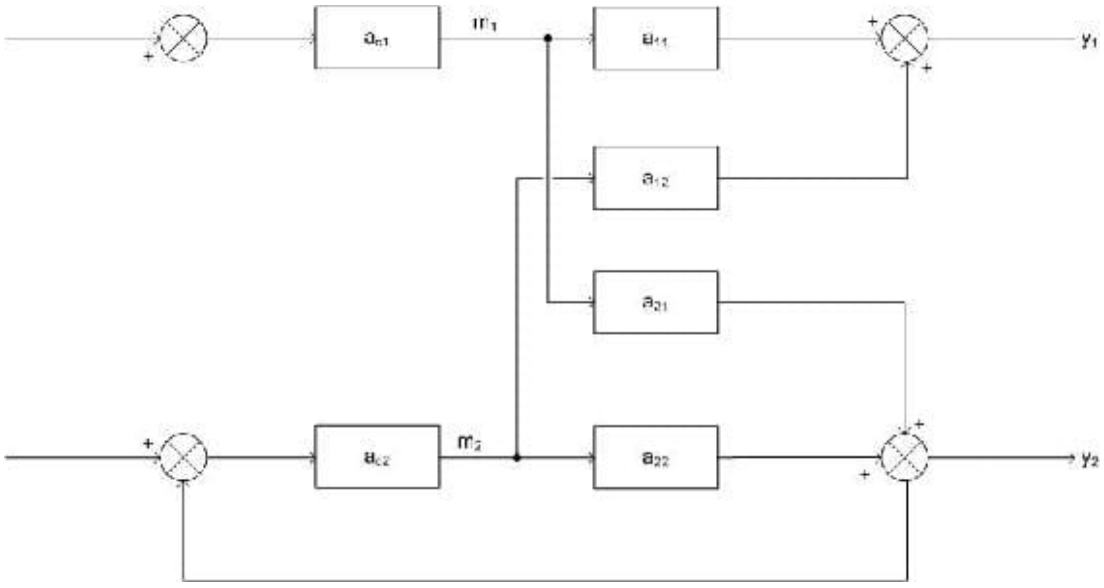
$$g_{11} = \frac{\Delta y_1(\text{ all - loops -open })}{\Delta m_1(\text{ all } - \text{ loops -open })}$$

The final objective of experiment one is obtaining the value g_{ij} where the controlled variable i is controlled by manipulated variable

j.

Experiment 2

The goal of experiment two is to determine the effect of m_2 on y_1. To do this, loop 1 must remain open, but all other loops remain closed. This configuration is shown below.



Now, the *same* step change that was introduced to m_1 in experiment one must be introduced again. The difference is, now loop 2 is in *perfect control* which means that when m_1 changes, it affects y_2, so the control loop on loop two will adjust m_2 in order to compensate and keep y_2 constant. When m_2 changes it, in turn, affects y_1. It is this change that is being observed. The amount y_1 changes is Δy_1 and the amount m_1 was stepped up is Δm_1. Now:

$$g_{11}^{*}=\frac{\Delta y_1(\text{ only }-\text{ loop }-\text{ one }-\text{ open })}{\Delta m_1(\text{ only }-\text{ loop }-\text{ one }-\text{ open })}$$

The objective of experiment two is to obtain this value of g_{ij}^{*} where controlled variable *i* is being controlled by manipulated *j*.

Compiling the Array

Once the experiments are run for every possible configuration, the results can be used to compile the relative gain array (RGA). To do this, we first have to find the *relative gain*. The relative gain is the ratio of g_{ij} to g_{ij}^{*} and is represented by λ_{ij}.

$$\lambda_{ij}=\frac{g_{ij}}{g_{ij}^{*}}=\frac{\left(\frac{\delta y_i}{\delta m_j}\right)_{\text{all-loops-open}}}{\left(\frac{\delta y_i}{\delta m_j}\right)_{\text{only-loop-i-open}}} \tag{12.2.1}$$

The value for λ must be computed for every possible combination of *i* and *j*. After every value is found, the RGA matrix can be formed:

$$RGA=\Lambda=\begin{bmatrix}\lambda_{11} & \lambda_{12} & \cdots & \lambda_{1n}\\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2n}\\ \vdots & & & \\ \lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn}\end{bmatrix} \tag{12.2.2}$$

Method 2: Calculating RGA with Steady-State Gain Matrix

Sometimes it is not convenient or possible to perform the experiments described above on every control pairing possibility in

the system. If that is the case, a *steady-state gain matrix* can be used to determine the RGA. If a process model is available, the steady-state gain matrix can be calculated from the model equations. If a process model is not available, the steady-state gain matrix can be calculated from experimental data (fewer experiments are required than when the RGA is calculated directly from experimental data as above). Once the steady-state gain matrix is calculated, it can be used to compute the RGA. Beware that a RGA has to have the same number of controlled variables and manipulated variables (same number of inputs and outputs) to be in an *m* x *m* matrix form.

Process Model Available

If a process model is available, the steady-state gain matrix relates the manipulated variables to the controlled variables according to the following equation:

$$\mathbf{y} = G\mathbf{m}$$

where $\mathbf{y}$ is a vector of the controlled variables:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$\mathbf{m}$is a vector of the manipulated variables:

$$\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{bmatrix}$$

and *G* is the steady-state gain matrix:

$$G = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & & & \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix}$$

The values of g_{ij} are calculated by taking partial derivatives of the equations governing the controlled variables. Specifically:

$$g_{ij} = \frac{\partial y_i}{\partial m_j}$$

During differentiation, hold constant all manipulated variables other than m_j. To fill in matrix *G* with numerical values, simply substitute the appropriate steady state values for the process into the expressions derived from differentiation.

Now that the steady-state gain matrix has been found, the RGA can be calculated with just a little bit more math. However, now that you have the steady state gain matrix, this might be a good time to determine if your system can even be decoupled! See the article on singular value decomposition for methods to check this. Assuming SVD gives a condition number of less than about 50, we can proceed to calculation of the RGA. First, define a matrix *R* to be the transpose of the inverse of matrix *G*:

$$R = \left(G^{-1}\right)^{\mathrm{T}}$$

The calculation of the **inverse** of a matrix is quite complicated for anything larger than a 2x2 matrix. Fortunately Mathematica will do this with the following command for a square matrix *M*:

Inverse[*M*]

The **transpose** of a matrix is when the rows become columns and the columns become rows. For a square matrix, this just means reflecting across the diagonal running from the top left to the bottom right corner of the matrix. The **transpose** can be found in Mathematica with the following command:

Transpose[*M*]

The RGA can now be obtained one element at a time according to this equation:

$$\lambda_{ij} = g_{ij} r_{ij}$$

Note that this is not your usual matrix multiplication! Here you multiply corresponding elements of the *G* and *R* matrices to get the corresponding element of the RGA. This is the type of multiplication Mathematica does with the standard multiplication operator.

See Example 1 below for a worked-out RGA calculation using this approach.

Process Model Not Available

In case there is no process model available and it is not feasible to determine the RGA by carrying out both experiments necessary for full experimental determination, it still may be possible to develop a steady-state gain matrix experimentally which can then be used to derive the RGA. In this case, just carry out experiment 1 as described above to determine the elements of the steady-state gain matrix:

$$g_{ij} = \frac{\Delta y_i}{\Delta m_j}$$

with all loops open.

Each element of the steady-state gain matrix can be determined this way, and then the RGA can be calculated from the steady-state gain matrix as shown in the previous section.

Interpreting the RGA

There are some important properties and guidelines in understanding and analyzing the RGA, and what the different values of the RGA mean:

1) **All elements of the RGA across any row, or down any column will sum up to one:**

$$\sum_{i=1}^{n} \lambda_{ij} = \sum_{j=1}^{n} \lambda_{ij} = 1$$

This makes calculating the RGA easier because:

- in 2X2 case, only 1 element must be calculated to determine all elements
- in 3X3 case, only 4 elements must be calculated to determine all elements

and so on.

2) The λ_{ij} calculated from steady-state matrix is dimensionless and unaffected by scaling.

3) Each of the rows in the RGA represent one of the outputs. Each of the columns represent a manipulated variable.

- **If $\lambda_{ij} = 0$:** The manipulated variable (m_j) will have no effect on the output or the controlled variable (y_i).
- **If $\lambda_{ij} = 1$:** The manipulated variable m_j affects the output y_i without any interaction from the other control loops in the system. From the definition of λ_{ij}, this implies that the gain loop with all loops open is equal to the gain loop with all other loops closed, ie: $g_{11} = g^*_{11}$.
- **If $\lambda_{ij} < 0$:** The system will be unstable whenever m_j is paired with y_i, and the opposite response in the actual system may occur if other loops are opened in the system.
- **If $0 < \lambda_{ij} < 1$:** This implies that other control loops (m_j- y_i) are interacting with the manipulated and controlled variable control loop.

Three different relationships based on λ=0.5 imply different interpretations of pairing and the RGA:

- **If λ_{ij} = 0.5:** The control pairing effect is equal to the retaliatory effect of other loops.
- **If λ_{ij} < 0.5:** The other control loops are influencing the control pair, and the influence from the other control loops are greater than the influence from the control pair.
- **If λ_{ij} > 0.5:** This means that the control pair has a greater influence on the system than the other control loops.
- **If λ_{ij} >1:** The open-loop gain of the control pair is greater than the gain with all other loops closed, ie: $g_{11} > g^{*}_{11}$. The positive value of RGA indicates that the control pair is dominant in the system, but the other loops are still affecting the control pair in the opposite direction. The higher the value of λ_{ij}, the more correctional effects the other control loops have on the pair.

λ_{ij}	Possible Pairing
$\lambda_{ij} = 0$	Avoid pairing m_j with y_i
$\lambda_{ij} = 1$	Pair m_j with y_i
$\lambda_{ij} < 0$	Avoid pairing m_j with y_i
λ_{ij} = or < 0.5	Avoid pairing m_j with y_i
λ_{ij} >1	Pair m_j with y_i

Analysis with RGA

The **Niederlinski Index** (NI) is a calculation used to analyze the stability of the control loop pairings using the result of the RGA, evaluated at Steady State:

$$NI = \frac{|G|}{\prod_{i=1}^{n} g_{ii}}$$

A negative NI value indicates instability in the control loop. For a 2 $\times$ 2 matrix, a positive NI value indicates stability in the pairings, but this is not necessarily true for larger matrices! For matrices larger than 2 $\times$ 2, a conclusion can only be drawn from a negative NI, which indicates instability. NI should not be used for systems that have time delays (dead time), as the NI stability predictions assume immediate feedback to the controller. However, the index can be adapted for systems with dead time, although that application is beyond the scope of this article.

Here's an example NI calculation: given the steady state gain matrix

$$G = \begin{bmatrix} -0.002 & 0.001 \\ 0.002 & 0.003 \end{bmatrix}$$

the NI can be calculated from the following expression:

$$NI = \frac{(-0.002 \times 0.003) - (0.001 \times 0.002)}{-0.002 \times 0.003} = 1.333$$

Since this is a 2 $\times$ 2 matrix, the positive value of the NI indicates stability in the control loop pairings.

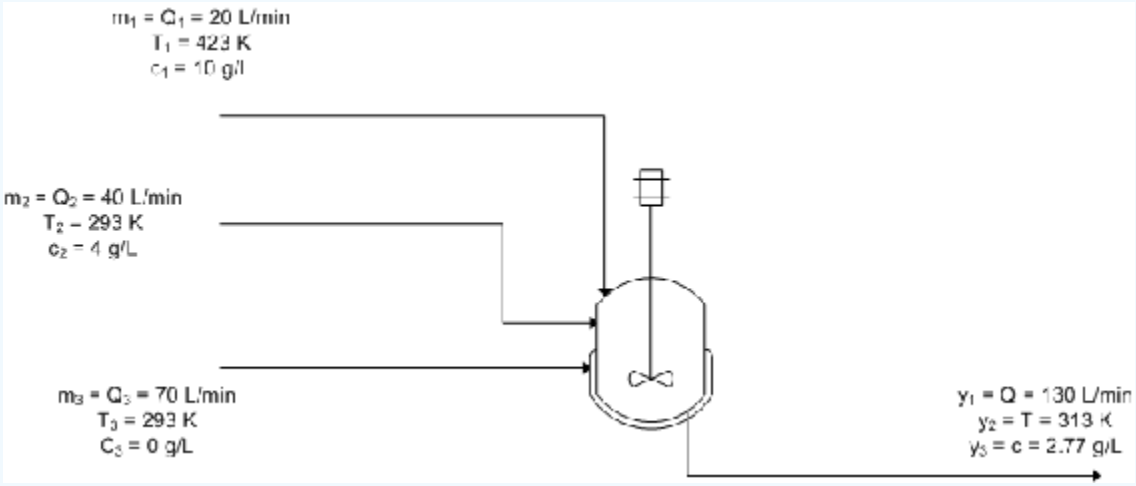
Optimizing a MIMO Control Scheme: Pairing Rules

The goal of the RGA and NI analysis is to quantitatively determine the optimal variable pairing for a given process. Some basic rules to remember when attempting to obtain an optimal pairing of control loops in a system are:

- **Rule 1:** The **positive** RGA elements that are closest to 1.0 should have the corresponding manipulated and controlled variables paired. When the CN number is large, implying a less decoupled system, one should look for the max RGA elements.
- **Rule 2:** If the NI value is negative, the loop pairing for that control system configuration is unacceptable.

Example 12.2.1

A blending unit is used to dilute and cool down the product stream of a reactor. Three streams are combined in the mixer: the hot, concentrated stream from the reactor, a room temperature stream containing none of the product A, and a second room temperature stream containing some A produced elsewhere in the process. It is desired to control the flow rate, temperature, and concentration of A in the product stream by manipulating the flow rates of the three input streams. A schematic of the process follows, filled in with the steady state values of flow rate, temperature, and concentration. The manipulated variables are indicated with the letter *m* and the controlled variables are indicated with the letter *y*:



The equations used to model the system are:

$$y_1 = m_1 + m_2 + m_3$$

$$y_2 = \frac{T_1 m_1 + T_2 m_2 + T_3 m_3}{y_1} = \frac{T_1 m_1 + T_2 m_2 + T_3 m_3}{m_1 + m_2 + m_3}$$

$$y_3 = \frac{c_1 m_1 + c_2 m_2 + c_3 m_3}{y_1} = \frac{c_1 m_1 + c_2 m_2 + c_3 m_3}{m_1 + m_2 + m_3}$$

Calculate the RGA using the steady-state gain matrix method and use it to determine the best control scheme for the system.

Solution

First you must take the partial derivatives of the control variable equations in order to determine the elements of the steady-state gain matrix:

$$g_{11} = \frac{\partial y_1}{\partial m_1} = 1$$

$$g_{21} = \frac{\partial y_2}{\partial m_1} = \frac{(T_1 - T_2)\, m_2 + (T_1 - T_3)\, m_3}{(m_1 + m_2 + m_3)^2}$$

$$g_{31} = \frac{\partial y_3}{\partial m_1} = \frac{(c_1 - c_2)\, m_2 + (c_1 - c_3)\, m_3}{(m_1 + m_2 + m_3)^2}$$

$$g_{12} = \frac{\partial y_1}{\partial m_2} = 1$$

$$g_{22} = \frac{\partial y_2}{\partial m_2} = \frac{(T_2 - T_1)\, m_1 + (T_2 - T_3)\, m_3}{(m_1 + m_2 + m_3)^2}$$

$$g_{32} = \frac{\partial y_3}{\partial m_2} = \frac{(c_2 - c_1)\, m_1 + (c_2 - c_3)\, m_3}{(m_1 + m_2 + m_3)^2}$$

$$g_{13} = \frac{\partial y_1}{\partial m_3} = 1$$

$$g_{23} = \frac{\partial y_2}{\partial m_3} = \frac{(T_3 - T_1)\, m_1 + (T_3 - T_2)\, m_2}{(m_1 + m_2 + m_3)^2}$$

$$g_{33} = \frac{\partial y_3}{\partial m_3} = \frac{(c_3 - c_1)\, m_1 + (c_3 - c_2)\, m_2}{(m_1 + m_2 + m_3)^2}$$

Now you can substitute in the given steady state values of the variables to find the numerical values of g_{ij}, obtaining:

$$g_{11} = 1$$

$$g_{21} = \frac{11}{13}$$

$$g_{31} = \frac{47}{845}$$

$$g_{12} = 1$$

$$g_{22} = -\frac{2}{13}$$

$$g_{32} = \frac{8}{845}$$

$$g_{13} = 1$$

$$g_{23} = -\frac{2}{13}$$

$$g_{33} = -\frac{18}{845}$$

In matrix form:

$$G = \begin{bmatrix} 1 & 1 & 1 \\ \frac{11}{13} & -\frac{2}{13} & -\frac{2}{13} \\ \frac{47}{845} & \frac{8}{845} & -\frac{18}{845} \end{bmatrix}$$

Now calculate matrix *R*:

$$R = \left(G^{-1}\right)^{\mathrm{T}} = \begin{bmatrix} \frac{2}{13} & \frac{4}{13} & \frac{7}{13} \\ 1 & -\frac{5}{2} & \frac{3}{2} \\ 0 & \frac{65}{2} & -\frac{65}{2} \end{bmatrix}$$

You can now calculate the RGA by multiplying matrices *G* and *R* piecewise, according to the formula:

$$\lambda_{ij} = g_{ij} r_{ij}$$

So the RGA is:

$$RGA = \begin{bmatrix} \frac{2}{13} & \frac{4}{13} & \frac{7}{13} \\ \frac{11}{13} & \frac{5}{13} & -\frac{3}{13} \\ 0 & \frac{4}{13} & \frac{9}{13} \end{bmatrix}$$

Note that all rows and columns sum to 1, as they should.

Now it's time to determine the optimal control scheme for the system. The closer an element in the matrix is to 1, the better the pairing is between the manipulated and controlled variables represented by that element. With this in mind, we can immediately see that m_1, the flow rate of stream 1, should be used to control y_2, the temperature of the output stream, because the value of λ_{21} is the closest to 1 in the matrix ($\frac{11}{13}$). Next in line is λ_{33}, with a value of $\frac{9}{13}$. This indicates that m_3, the flow rate of stream 3, should control y_3, the concentration of A in the product stream. So far so good, but what should control y_1, the flow rate of the product stream? The best option would be m_3, but this is already controlling the concentration of A in the product, so the next best option is m_2, the flow rate of stream 2. This is not optimal because the relative gain is less than 0.5, but it is the best option available given the constraints of the system.

Example 12.2.2

In a distillation column, reflux rate (R) and steam rate (S)are used to control the compositions in the distillate (x_D)and the bottom product (x_B). The control loops are opened in order to conduct experiment one as described above, and the following set of data is collected:

Trial	R (kg/min)	S (kg/min)	x_D	x_B
1	75	20	0.96	0.05
2	85	20	0.94	0.06
3	95	20	0.92	0.08
4	85	15	0.95	0.06
5	85	25	0.97	0.03

Calculate the RGA for this system, and determine which is the best pairing between the manipulated (R and S) and controlled variables (x_D and x_B).

Solution

From the experiment data, it is observed that the reflux rates (R) are constant for trials 2, 4 and 5, while the steam rates (S) are constant for trials 1, 2 and 3.
First, calculate the steady-state gain matrix for this system:

$$g_{11} = \frac{\Delta x_D}{\Delta R}_{\text{with S constant}} = \frac{0.96 - 0.94}{75 - 85} = -0.002$$

$$g_{12} = \frac{\Delta x_B}{\Delta R}_{\text{with S constant}} = \frac{0.05 - 0.06}{75 - 85} = 0.001$$

$$g_{21} = \frac{\Delta x_D}{\Delta S}_{\text{with R constant}} = \frac{0.95 - 0.97}{15 - 25} = 0.002$$

$$g_{22} = \frac{\Delta x_B}{\Delta S}_{\text{with R constant}} = \frac{0.06 - 0.03}{15 - 25} = 0.003$$

Organize the gain value into the matrix form and compute R:

$$G = \begin{bmatrix} -0.002 & 0.001 \\ 0.002 & 0.003 \end{bmatrix}$$

$$R = \left(G^{-1}\right)^T = \begin{bmatrix} -375 & 250 \\ 125 & 250 \end{bmatrix}$$

Multiply each value in the R matrix by its corresponding gain value to obtain λ_{ij}:

$$\lambda_{11} = g_{11} R_{11}$$

λ_{ij} is then organized into the RGA matrix form:

$$RGA = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$$

The RGA matrix indicates that the composition of the distillate (x_D) would be best controlled by the reflux flow (R); while x_B would be best controlled by the steam rate (S)from the reboiler.

Example 12.2.3

An example problem based from Fall 2008 second exam:

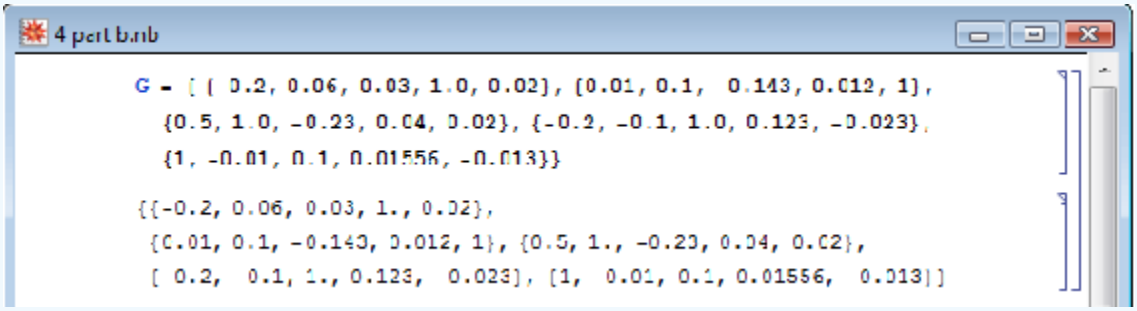
"Imagine that you are developing a control scheme for a new low temperature reactor system. The physics governing the system is complicated, so you decide to start out with an experiment on the system to see if you can develop a simple set if SISO controllers. To evaluate this possibility, you begin with your five valves at their steady state operating conditions (v1=0.5, v2=0.55, v3=0.5, v4=0.6, v5=0.4) and measure the five key variables of this process T1, T2, T3, P1 and P2. Next you slightly adjust each valve to see the steady state effects on each output variable"

As an approximation, you fit this model to the following system:

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ P_1 \\ P_2 \end{bmatrix} = \begin{bmatrix} -0.2 & 0.06 & 0.03 & 1.0 & 0.02 \\ 0.01 & 0.1 & -0.143 & 0.012 & 1.0 \\ 0.5 & 1.0 & -0.23 & 0.04 & 0.02 \\ -0.2 & -0.1 & 1.0 & 0.123 & -0.023 \\ 1.0 & -0.01 & 0.1 & 0.01556 & -0.013 \end{bmatrix} \begin{bmatrix} \mathrm{V}_1 \\ \mathrm{v}_2 \\ \mathrm{v}_3 \\ \mathrm{v}_4 \\ \mathrm{v}_5 \end{bmatrix}$$

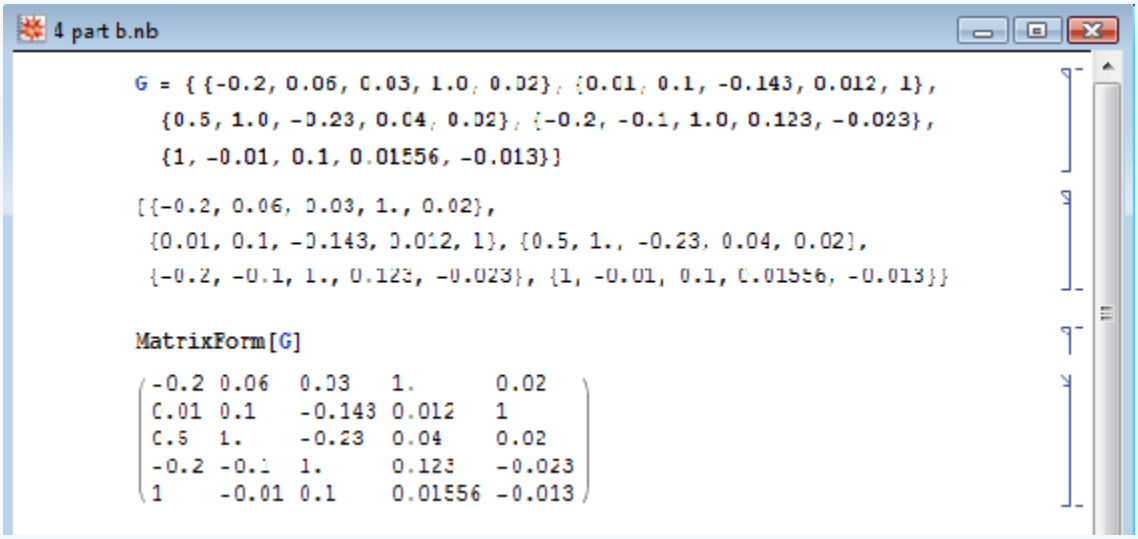
"Perform an RGA analysis to find the best pairing between valves (v1, v2, v3, v4, and v5) and outputs (T1,T2, T3, P1, and P2)."

In Mathematica, you can input the matrix of coefficients by the following:

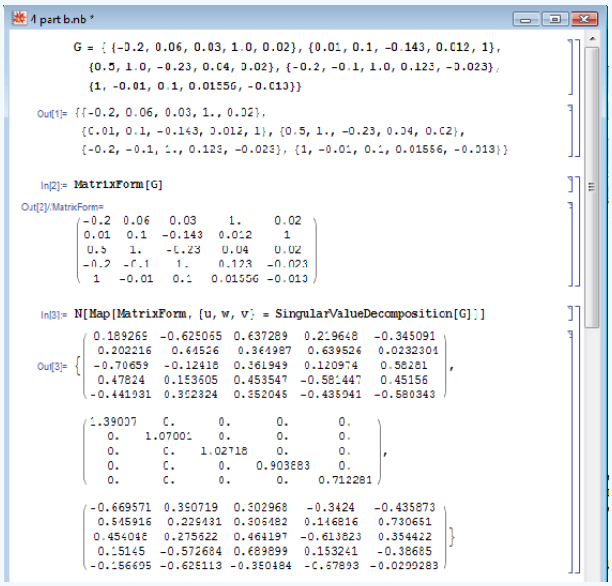


Note: I assign the matrix of coefficients to the variable, G.

It can be viewed in a matrix form in Mathematica by using function, "MatrixForm[]":

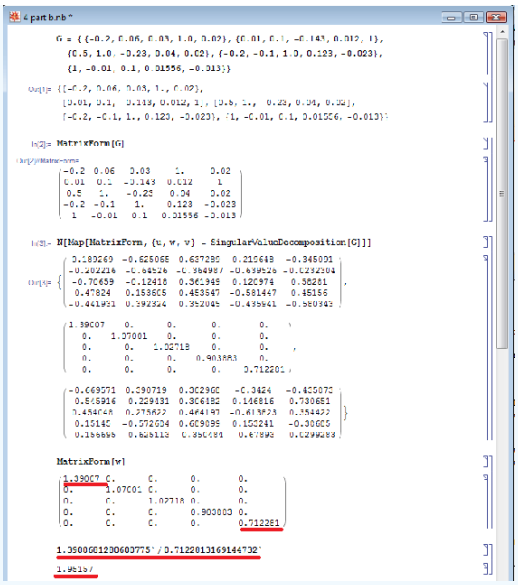


With the gain matrix, G, you can use the a singular value decomposition (SVD) of the gain matrix to determine how linearly dependent the model is. The linearity is based on the condition number (CN). Mathematica has a built-in function to do singular value decomposition. The following shows how to do singular value decomposition for this example.

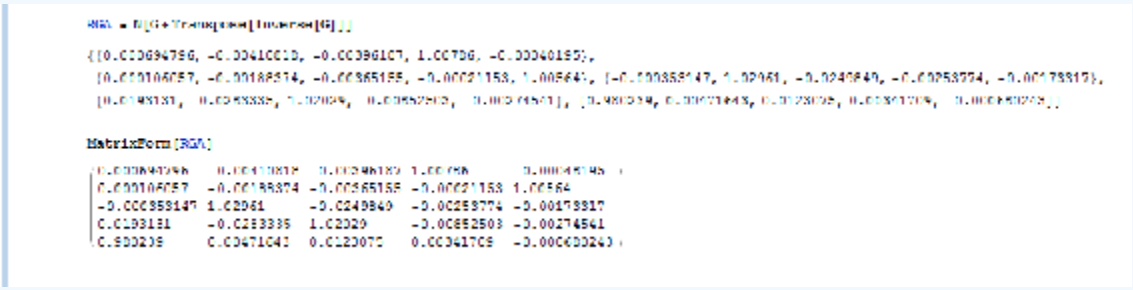


The Mathematica syntax is : N[Map[MatrixForm, {u, w, v} = SingularValueDecomposition[G]]]. The "SingularValueDecomposition[]" pops out three arrays assigned to variable *u, w, and v* . The "Map[MatrixForm, ..." portion pops out the three arrays into a matrix view and "N[.." portion forces numerical values in the matrices.

To the condition number(CN), you divide the biggest value by the smallest value in matrix, *w*. The general rule thumb is that if the condition number(CN) is greater than 50, then the system/model is too difficult to decouple.



This illustrates that this system can be decoupled because the CN number is below 50. To determine the best pairing, RGA analysis is used in which it is a normalized form of the gain matrix, "G" that determines the impact of each control variable on its respective output e.g. how much impact does v1 have on T1, etc. The Mathematica syntax is N[G*Transpose[Inverse[G]]]. The RGA normalization is essential since at times, not in this case however, the values of the gain matrix may represent that a valve has great influence on multiple variables (ie. a column may contain the highest value from multiple rows). In order to distinguish unique separations between these influences, the RGA is used to normalize the gain matrix as seen below. This allows us to determine the different magnitudes of influence the valve has on an output in order to find the predominant effect on an output. The RGA illustrates these marginal differences in impact by the valves on the variable outputs, and lets us make the best possible pairings.



The best pairing is discovered by taking the maximum value of RGA Matrix for each row.

$$\text{RGA} = \begin{bmatrix} 0.000694 & -0.00411 & -0.00396 & \mathbf{1.0079} & -0.000482 \\ 0.000106 & -0.00188 & -0.00365 & -0.000212 & \mathbf{1.00564} \\ -0.000353 & \mathbf{1.0296} & -0.0249 & -0.00254 & -0.00173 \\ 0.0193 & -0.0283 & \mathbf{1.0202} & -0.00853 & -0.00275 \\ \mathbf{0.9802} & 0.00471 & 0.0123 & 0.00342 & -0.000680 \end{bmatrix} \begin{bmatrix} V1 \\ V2 \\ V3 \\ V4 \\ V5 \end{bmatrix}$$

The bold values correspond to the valve that is the best pairing for the individual output. By looking at the bold values, V1->P2 V2->T3, V3-> P1, V4-> T1, V5-> T2.

Exercise 12.2.1

For a MIMO control system containing 4 different manipulated and 4 different controlled variables, how many control configurations are possible for the system?

a. 24
b. 4
c. 64
d. 256

Answer available at the end of this section.

Exercise 12.2.2

In a case of a 2x2 matrix system, if one RGA element in first row calculated to be $\frac{3}{4}$. According to the RGA properties, what is the other element in that same row?

a. $\frac{3}{4}$
b. $\frac{1}{4}$
c. $\frac{2}{4}$
d. Cannot be determined with given information.

Answer available at the end of this section.

Exercise 12.2.3

RGA is useful for process engineers because it allows us to determine:

a. Which matrix size to use.
b. What eigenvalues and eigenvectors are paired.
c. What input and output pairs to use.
d. When feed forward control is preferred over feedback control.

Answer available at the end of this section.

References

- Berber, Ridvan. *Methods of Model Based Process Control*, Kluwer Academic Publishers, 1994.
- Lee, P.L., Newell, R.B., Cameron, I.T. *Process Control and Management*, Blackie Academic and Professional.
- Romagnoli, Jose A., Palazoglu, Ahmet. *Introduction to Process Control*, Taylor and Francis Group.
- Shinskey, F.G. *Process Control Systems: Application, Design, and Tuning, Fourth Edition*, McGraw-Hill, Inc., 1996.
- Svrcek, W.Y., Mahoney, D.P., and Young, B.R. *A Real-Time Approach to Process Control*, John Wiley and Sons, Ltd.
- Edinburgh University, Exercise: Relative Gain Arrays, 2006
- Edinburgh University, Multivariable Systems, 2007

Answers

Exercise 12.2.1

a

Explanation: Mathematically, for an *nxn* manipulated by controlled variable system, the number of control configurations possible will be *n*! or *n* factorial. For a 4x4 variable system, 4x3x2x1 = 24. Physically, the number of control configurations translates to the number of possible input-output pairings.

Exercise 12.2.2

b

Explanation: As explained above, all elements of the RGA across any row, or down any column should sum up to unity (one). For a 2x2 matrix, if one element in a row is $\frac{3}{4}$, the other element must be 1- $\frac{3}{4}$ = $\frac{1}{4}$

Exercise 12.2.3

c

Explanation: Using RGA, a process engineer will assess the relative gain of pairing each input with each output to minimize undesired interactions and maximize desired interactions between controlled and manipulated variables in a system. See "Interpreting the RGA" section if you are confused on how this analysis is achieved.

This page titled 12.2: MIMO control using RGA is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Angela Yeung, Ben van Kuiken, Curt Longcore, & Jeffrey Carey via source content that was edited to the style and standards of the LibreTexts platform.

12.3: MIMO using Model Predictive Control

Introduction

This section describes how to control a system with multiple inputs and outputs using **Model Predictive Control** (MPC). MPC is a linear algebra method for predicting the result of a sequence of control variable manipulations. Once the results of specific manipulations are predicted, the controller can then proceed with the sequence that produces the desired result. One can compare this controller method to "look ahead" in chess or other board games. In look ahead, you foresee what an action might yield some time in the future using specific knowledge of the process (or game in the case of chess), and are thereby able to optimize your actions to select for the best long term outcome. MPC methods can prevent an occurrence with conventional PID controllers in which actions taken achieve short term goals, but end up very costly in the end. This phenomenon can be described as "winning the battle but losing the war."

The open ended nature of MPC allows the process control engineer use MPC to control any system for which models can be generated.

Model Predictive Control

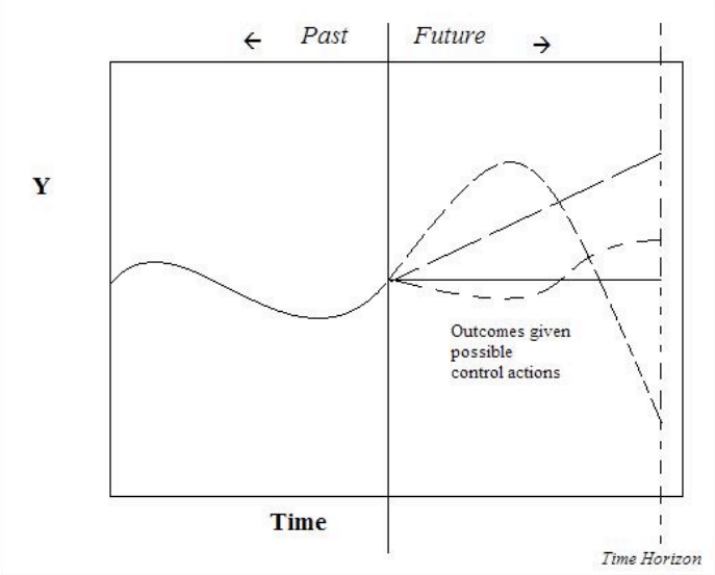
MPC is a widely used means to deal with large multivariable constrained control issues in industry. The main aim of MPC is to minimize a performance criterion in the future that would possibly be subject to constraints on the manipulated inputs and outputs, where the future behavior is computed according to a model of the plant. The model predictive controller uses the models and current plant measurements to calculate future moves in the independent variables that will result in operation that honors all independent and dependent variable constraints. The MPC then sends this set of independent variable moves to the corresponding regulatory controller set-points to be implemented in the process.

MPC uses the mathematical expressions of a process model to predict system behavior. These predictions are used to optimize the process over a defined time period. An MPC controller can operate according to the following algorithm.

1. Development of a process model by the control engineers.
2. At time t, previous process inputs and outputs are used, along with the process model, to predict future process outputs $u(f)$ over a "prediction horizon."
3. The control signals that produce the most desired behavior are selected.
4. The control signal is implemented over a pre-defined time interval.
5. Time advances to the next interval, and the procedure is repeated from step 2.

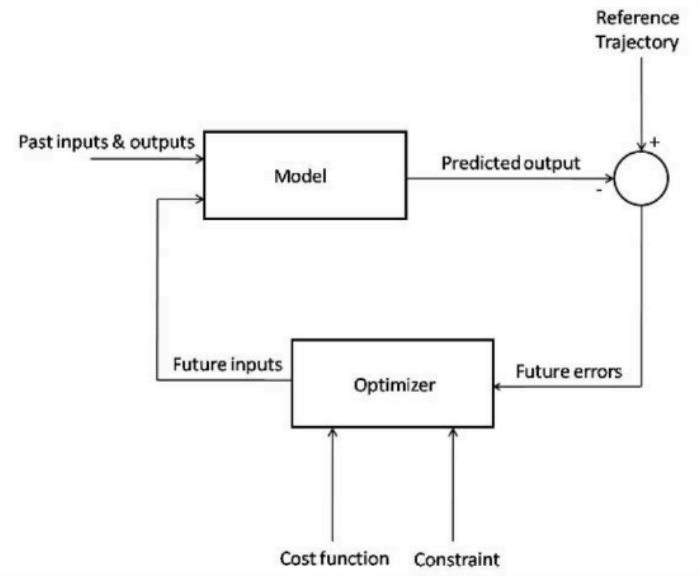
This is one of the many algorithm possibilities, which can be applied to systems with any number of inputs or outputs. The process model can be variable as well. Examples include physical models, input-output models, and state models which are all derived from the specific system being controlled.

When comparing predicted behavior to desired behavior, there are multiple techniques. A common procedure is to generate a second mathematical model that describes your desired behavior. When process behavior is predicted in step 2 of the MPC algorithm, the control signals which produce the predicted behavior that minimize deviations from your desired behavior over the $[t, t+h]$ interval, are selected. The generation of models and optimization process is repeated continuously as the algorithm is repeated.



As seen in the figure above, depending on the algorithm, MPC may generate possible outcomes given possible controller action. These generations are either based on past process outputs, or the process model. After many possible outcomes are generated, the controller can pick one based on the optimization goals. This generation and optimization process is repeated at every time step.

The flow diagram below depicts the flow of information used by the controller.



The figure above shows the basic structure of a Model Predictive Controller. The model takes data from past inputs and outputs, and combines it with the predicted future inputs, and gives a predicted output for the time step. This predicted output is combined with the reference trajectory, giving the predicted future errors of the system. These errors are fed into an optimizer, which enforces the constraints of the system (for instance, ensuring that a flow rate calculation for the model is not greater than the maximum flow) on the predicted outputs and minimizes the operating cost function. This gives the predicted future inputs, which are fed back into the main model, restarting the cycle.

Motivation

The motivation for the development and implementation of MPC controllers is compelling. MPC is very simple for sampled systems in which the control signal is constant over the interval [t,t+h]. The value of h can then be taken as the sampling interval,

and the prediction horizon can become a small number of sampling intervals. This can reduce the computational power needed to implement a model predictive controller.

A highly desired feature of MPC is that constraints can be implemented in the controller. These constraints include heaters and valves that have a finite operating range, actuators with finite states (on/off or low/high), and cost or energy limits for the process. MPC can incorporate these constraints and eliminate the possibility of variables exceeding their limits. This helps the process operate efficiently, prevents damage to equipment, and prevents the system from running away by continuously increasing a variable's setting when the limit for the equipment has already been reached. Another benefit of MPC controllers incorporating system constraints is the ability of the model to dynamically react to system changes. For example, if a valve is stuck open, it can be added as a constraint, and the model will compensate accordingly. This allows the controller to continue effectively controlling the system after an equipment malfunction. In a simpler control, such as PID control, this would not be possible.

MPC is a highly specific method for controlling a process. Each controller is specific to the system it was designed for and the model equations, constraints, and set points will change for different systems. This means that a controller developed for a tank reactor will not be able to control an evaporation unit as the process model will be very different. This weakness in MPC brings out its biggest strength. The specificity and customizable nature of the controller will empower you with the freedom to design for exactly what is desired.

MPC is a flexible control technique that uses discrete time segments and is the most commonly applied advanced control technique in the chemical process industry. MPC helps to simplify or completely eliminate controller design and instead works as a system modeling controller. With MPC the designer does not have to worry about optimizing control parameters such as with PID control. The ability of MPC controllers to handle constraints in an optimal fashion is also a contributor to its success.

Model Predictive Control Example

To demonstrate the concepts of MPC, a general example for the development of a sampled process and a first-order system will be shown.

General Model

Take Equation 12.3.1 as our process model:

$$y(t) + a_1 y(t-h) + ... + a_n y(t-nh) = b_1 u(t-h) + b_2 u(t-2h) + ... + b_n (t-nh) \tag{12.3.1}$$

where u is controller input, y is process output, and h is the time interval. This is a general equation that relates previous process output $y(t-h)$ and previous controller input $u(t-h)$. In some situations this equation may be created using fitted experimental data, but is most often a derivation using knowledge of your specific system and fundamentals of chemical engineering. At time t, total previous behavior y_p is shown as Equation 12.3.2.

$$y_p = f(y(t), y(t-h), \dots, u(t-h), u(t-2h), ...) \tag{12.3.2}$$

Future process output y_f can be predicted using current and future control signals, $u(t)$ and $u(t+h)$ respectively:

$$y_f = f(u(t), u(t-h), \dots, u(t+Nh)) \tag{12.3.3}$$

Both y_p and y_f could possibly be created by fits of experimental data, but are more likely to be derived from specific equations related to your system. Deviations from the desired behavior y_d, either specified by another mathematical model or reference trajectory, produce an error function $e(t) = y(t) - y_d(t)$ for increments of control actions $\Delta u(t) = u(t) - u(t-h)$. The loss function J to be minimized is shown as Equation 12.3.4.

$$J(u(t), u(t-h), \dots, u(t+Nh)) = \sum_{l=1}^{t|N} e(t+kh)^2 + \rho(\Delta u(t+(k-1)h))^2 \tag{12.3.4}$$

The control inputs that minimize Equation 12.3.4 are then applied to the system by the controller over the time interval, and the process is repeated. The control input function F in Equation 12.3.5 is determined implicitly by the optimization.

$$u(t) = F(y(t), y(t-h), ... y(t-nh), u(t-h), y(t-2h), \dots, u(t-nh)) \tag{12.3.5}$$

This general model is meant to be a guideline, and the equations listed representative, for the thought process required to create

a model predictive controller.

First-Order System Example

Take the process model to be Equation 12.3.6:

$$\Delta y(t+h) = -a\Delta y(t) + b\Delta u(t) \tag{12.3.6}$$

Let us define $\Delta y(t) = y(t) - y(t-h)$ and $\Delta u(t) = u(t) - u(t-h)$. Let us also define our desired system behavior y_d as a function which starts at $y(t)$ and exponentially approaches a set point y_{sp} with time constant T. Our desired behavior y_d then becomes Equation 12.3.7:

$$y_d(t+h) = y(t) + \left(1 - e^{\frac{-h}{T}}\right)(y_{sp} - y(t)) \tag{12.3.7}$$

Assuming that our controller can take as much action as needed to produce the desired behavior, the desired behavior can be realized in the next sampling period. This is done by setting $y(t+h)$ equal to $y_d(t+h)$, and can be seen in Equation 12.3.8:

$$y(t+h) = y(t) + \Delta y(t+h) = y(t) - \alpha\Delta y(t) + b\Delta u(t) = y(t) + \left(1 - e^{\frac{-h}{T}}\right)(y_{sp} - y(t)) \tag{12.3.8}$$

Solving Equation 12.3.8 for $u(t)$ gives Equation 12.3.9:

$$\Delta u(t) = \frac{a}{b}\Delta y(t) + \frac{1 - e^{\frac{-h}{T}}}{b}(y_{sp} - y(t)) \tag{12.3.9}$$

Upon examination of this result, you can see that we have produced a PI controller with gains $k = \frac{a}{b}$ and $k_i = \frac{1-e^{\frac{-h}{T}}}{b}$. It should be noted that the proportional gain *k* will only depend upon the developed process model, and the integral gain k_i depends on both the process model and the desired response rate *T*.

This process can be modified to include multiple inputs and outputs via the process model and desired behavior. We may also add system constraints to our MPC example by adding conditional statements limiting the maximum controller output discussed in the logical programs wiki. As discussed before, there are as many variations on the MPC process as you can think of. See the worked examples 1 and 2 for more MPC instances.

Question

Is the above expression a PI control system or more of a PD control system? The first term is k*(y(present0)-y(past)) which seems more like a derivative term. The second term is more like k*(y(set)-y(present)) which seems more like a P term.

Differences from Other Controllers Types

Model predictive control uses a mathematical model to simulate a process. This model then fits the inputs to predict the system behavior. In this way, MPC is a type of feed forward control. It uses system inputs as a basis of control. MPC is more complex than most other feed forward control types because of the way these predictions are used to optimize a process over a defined amount of time. Most feed forward control types do not take into account the process outputs much past a residence time. The MPC algorithm will compare predicted outputs to desired outputs and select signals that will minimize this difference over the time selected. This control type can see ahead into multiple time steps in the future in order to optimize the process. Normal PID type controllers use mathematical expressions based on error from a set point. The governing equation for MPC controllers are based on set points, system properties, and desired outcomes and optimization.

MPC is very specific to the process it is modeling. Unlike ratio or cascade control set ups, where it is simple to implement and change set points in various situations, MPC will model one specific process and optimize it. As previously mentioned this can either be an advantage or disadvantage. MPC is great for selecting one type of operation on one system and perfecting it to the desired conditions. This also has a downside in that the model equation will work for one and only one situation.

Limitations of MPC

Advantages of MPC

1. MPC can be used to handle multivariable control programs.
2. MPC can consider actuator limitations.
3. MPC can increase profits by allowing for operation close to the system constraints.
4. MPC can perform online computations quickly.
5. MPC can be used for non-minimal phase and unstable processes.
6. MPC is easy to tune.
7. MPC is able to handle structural changes.

Disadvantages of MPC

1. Several MPC models are limited to only stable, open-loop processes.
2. MPC often requires a large number of model coefficients to describe a response.
3. Some MPC models are formulated for output disturbances, and they may not handle input disturbances well.
4. Some forms of MPC use a constant output disturbance assumption. This corrects for the fact that the output predicted by the model is not exactly equal to the actual measured output. This method assumes the correction term is constant in the future, which may not yield a good performance if there is a real disturbance at the plant input.
5. If the prediction horizon is not formulated correctly, control performance will be poor even if the model is correct.
6. Some systems have a wide range of operating conditions that change frequently. Some examples of this include exothermic reactors, batch processes, and any systems where different consumers have different product specifications. An MPC linear model will not be able to handle the dynamic behavior of these processes. A nonlinear model must be used for better control performance.

Industrial MPC Applications

There are many industrial applications that incorporate model predictive control in order to effectively control a multivariable system. In order to effectively do this, one needs to set up a working model by testing many different parameters in a plant. This is usually done by starting up a plant, varying many different parameters, and having the MPC program analyze the data.

In order to test the plant, one may vary parameters such as:

1. feed flow rate/composition
2. steam pressure
3. heat duty
4. recycle ratio
5. reactor temperatures

This is only a small sample of parameters that can be changed. Once these parameters have been changed, the data is analyzed, and downstream effects of these parameters are characterized as a function of these variables. This relationship can be a combination of many different relationships (linear, nonlinear, logarithmic, exponential, power, etc.).

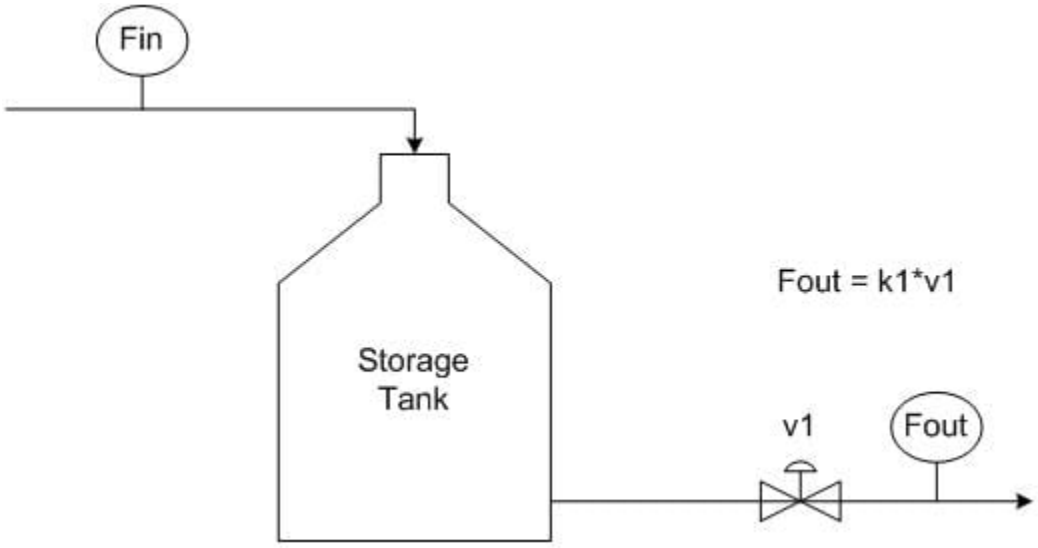
By testing the plant thoroughly and coming up with a robust model, the engineers ensure that an MPC controller will be able to much more effectively run the plant. Barring any major process changes, this model should be accurate for normal use.

Some industrial MPC applications are:

1. Model Predictive Heuristic Control by Richard et al. 1976 (Adersa)
2. Dynamic Matrix Control (DMC) by Cutler and Ramaker 1979 (Shell Oil)
3. Quadratic-Program Dynamic Matrix (QDMC) Control by Cutler et al. 1983 (Shell Oil)
4. IDCOM-M by Setpoint, Inc (part of ASPEN Technology)
5. Generalized Predictive Control (GPC)

Implementing MPC using Excel

In MPC, values of the control variables will be optimized for a given time interval in order to best tell the system how it should act. The control variables will be optimized by optimizing some characteristic. Usually, this characteristic is simply the least squared error between an actual state and a "set" or desired state. This can easily be done using the Solver tool in Excel. See MPC_in_Excel for an example of implementing MPC to control the valve, v1, in the following system (keep reading to learn about everything in the Excel file).



To describe this system in words, there is a variable feed, Fin, with time that pours into a storage tank. The flow out of the tank, Fout, is controlled by the valve, v1, multiplied by a constant of proportionality, k1 = 5Ø. The tank volume is to be maintained at 15Ø.

There are three scenarios contained in the example Excel file. In the first scenario, the valve is continuous and in the second, it is binary (can be open or closed only). In the third scenario, the valve is still binary; however, Solver cannot recognize this setup even though this setup may seem intuitive (this is discussed further below). As it turns out, it doesn't really matter if a binary or continuous valve is used. Both types of valves yield very similar results; however, the continuous valve yields slightly tighter control over the tank volume (see screenshot below).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
19				MPC assuming continuous valve					MPC assuming binary valve								
20									Using MPC *correctly* (letting Solver think initially it can choose any values for v1)					Using MPC *incorrectly* (look how the Fout cells check to see if v1 is 1 or 0)			
21																	
22						sum of least squares:	408858				sum of least squares:	409588.2458				sum of least squares:	616273.152
23						sum of least squares from the time the controller first acts until the end:	24.28831				sum of least squares from the time the controller first acts until the end:	754.5180095			sum of least squares from the time the controller first acts until the end:	207439.424
24	time	Fin		V	Fout	v1 (MPC)	(V-Vset)^2		V	Fout	v1 (MPC)	(V-Vset)^2		V	Fout	v1 (MPC)	(V-Vset)^2
99	7.4	31.609		150.056	26.6655	0.53331	0.003157		142.989	0	0	49.15103561		197.989	0	0	2302.96524
100	7.5	25.8309		150.551	25.8154	0.51631	0.303087		146.15	0	0	14.82155646		201.15	0	0	2616.33504
101	7.6	15.1411		150.552	25.0974	0.50195	0.304789		148.733	0	0	1.604756631		203.733	0	0	2887.25788
102	7.7	29.7642		149.556	27.7989	0.55598	0.196739		150.247	50	1	0.061165848		205.247	0	0	3052.26607
103	7.8	29.8362		149.753	29.7791	0.59558	0.06102		148.224	50	1	3.155098484		208.224	0	0	3390.00394
104	7.9	34.5313		149.759	31.3583	0.62717	0.058232		146.207	50	1	14.38411759		211.207	0	0	3746.34093
105	8	32.3695		150.076	31.2803	0.62561	0.005773		144.66	50	1	28.51040645		214.66	0	0	4180.97849
106	8.1	23.6817		150.185	30.231	0.60462	0.034187		142.897	0	0	50.44642486		217.897	0	0	4610.06172
107	8.2	36.8266		149.53	30.7764	0.61553	0.220927		145.266	0	0	22.41449205		220.266	0	0	4937.2553
108	8.3	25.6309		150.135	28.8694	0.57739	0.018225		148.948	50	1	1.10614036		223.948	0	0	5468.34633
109	8.4	30.0267		149.811	27.6976	0.55395	0.035666		146.511	50	1	12.17064756		226.511	0	0	5853.98736
110	8.5	24.0554		150.044	25.6788	0.51358	0.001941		144.514	0	0	30.09588238		229.514	0	0	6322.48074
111	8.6	24.8008		149.882	24.3336	0.48667	0.013992		146.92	0	0	9.489065109		231.92	0	0	6710.81552
112	8.7	16.7253		149.928	23.1774	0.46355	0.005122		149.4	50	1	0.360425963		234.4	0	0	7123.3001
113	8.8	32.4343		149.283	24.4854	0.48971	0.513764		146.072	0	0	15.42777862		236.072	0	0	7408.41998
114	8.9	26.3788		150.078	23.4468	0.46894	0.006102		149.316	0	0	0.468396354		239.316	0	0	7977.27732
115	9	18.7908		150.371	21.7067	0.43413	0.137879		151.953	50	1	3.816116295		241.953	0	0	8455.444
116	9.1	20.1458		150.08	21.2798	0.4256	0.006357		148.833	0	0	1.362898635		243.833	0	0	8804.55072
117	9.2	18.5492		149.966	21.535	0.4307	0.001133		150.847	0	0	0.717666524		245.847	0	0	9186.67655
118	9.3	25.7376		149.668	22.8006	0.45601	0.110388		152.702	0	0	7.301175146		247.702	0	0	9545.69422
119	9.4	22.3079		149.961	22.7419	0.45484	0.001486		155.276	50	1	27.83436294		250.276	0	0	10055.2417
120	9.5	16.9677		149.918	22.0778	0.44156	0.006715		152.507	0	0	6.28312927		252.507	0	0	10507.6065
121	9.6	23.1784		149.407	21.5435	0.43087	0.351595		154.203	50	1	17.66846993		254.203	0	0	10858.3461
122	9.7	16.7697		149.571	19.9593	0.39919	0.184438		151.521	0	0	2.314139423		256.521	0	0	11346.7724
123	9.8	24.4947		149.252	17.5396	0.35079	0.560147		153.198	50	1	10.22844946		258.198	0	0	11706.8493
124	9.9	36.2061		149.947	10.9221	0.21844	0.0028		150.648	50	1	0.419471391		260.648	0	0	12242.906
125	10	37.2061		152.475	50	1	6.128025		149.268	0	0	0.535425704		264.268	0	0	13057.238

Sheet1 / Sheet2 / Sheet3

As you can see from this screen shot, the sum of the least squares between the actual volume and the set volume (Row 22) is lowest for the 1st scenario, when the valve is continuous. This measure of deviation from the set value is only slightly lower than

in the 2nd scenario, when the valve is binary. In the 3rd scenario, Solver didn't work properly, resulting in all of the v1 variables (Column P) to be set to 0.

At the start of the simulation, the volume of the tank is 0, so v1 stays closed until time = 5.4 when the volume is very close to the set volume. If the deviations that occurred during the time v1 was closed (which is the same time length for all 3 scenarios) is removed from the sum of the least squares (Row 23) the 1st scenario suddenly looks much better at controlling the volume than the 2nd scenario.

There are a couple things to note while using Solver in Excel to perform MPC:

1) This note only applies if your system has a binary control variable. If you intend to use a conditional statement (such as an IF() function) that depends on the optimization of the control variable, you must leave open the possibility that this control variable can be any real number, whether it's binary or not. This may seem unnecessary because Solver will output only optimized discreet numbers. To put this in the context of the example Excel file, you might think you could have simply done the following to control Fout:

- = IF(v1 = 1, k1*1, 0)

(The variables "v1" and "k1" are not recognized by Excel. Instead of typing the variables into Excel, cells containing the variable values must be referenced.)

This statement says that if the valve is open, allow a flow of k1 (or 50), otherwise, the flow is 0. This would not allow Solver to work properly. If Solver chose a value of v1 = 0.1, measured the target cell, and then it chose v1 = 0.2, it wouldn't see any difference because in both cases the flow would be 0. Solver needs to be able to see a change in the system, whether it's practical or not, to determine the actual gradient of your system. This gradient cannot be determined if an IF() statement turns the continuous output to discreet output.

The proper way to handle this is to program a constraint into solver that allows only binary outputs of v1 values (select bin from the pull down menu when adding the constraint). F_{out} can then be programmed as:

- = v1 * k1

Using this method, Solver can calculate the gradient of the system to define how it would react if v1 changed to a non-binary value. By using this method, Solver will output only binary values for the valve state as is desired.

2) Solver can only handle manipulating 200 variables at a time, so if your model requires looking at more than 200 time steps, you will have to optimize your system in sections.

Example 12.3.1

You have a tank outside your home used to supply water for various household functions (i.e. flushing the toilet). The optimal water pressure from the tank is achieved when the height of water in the tank is at 15, and the tank can hold water up to a height of 30, because when it rains, it pours. When it does rain, the level of water in this tank exceeds 15 and a model predictive controller and valve is in place to release water so the height will drop back down to this value (don't worry, this water is not wasted, but used as drinking water on your llama farm). At time t=25, there is a downpour that increases the flow into the tank by 150 and lasts for 100 time steps. Will this tank overflow, and how long will it take for the tank to return to the desired set point? Your controller output is 25 maximum.

Other key pieces of info:

- H(0)= 15
- A= 100
- B= 1
- T= 15
- Hset= 15
- h= 1

Solution

To obtain the required information, we will produce a model predictive controller for the water tank system, and simulate how the system will behave during the downpour. The first step is to generate a mathematical model that describes the relevant parameters. Equation 12.3.10 below relates the fluid height in the tank H to the flow in, and controller signal u.

$$A\frac{dH}{dt} = F_{in} - Bu\sqrt{H(t)} \tag{12.3.10}$$

where A and B are physical constants of the system. Next we will define the way we wish the height of fluid in the tank to act in Equation 12.3.11.

$$H(t+h) = H(t) + \left(1 - e^{-h/T}\right)(H_{set} - H(t)) \tag{12.3.11}$$

Where h is the time interval, T is an exponential time constant, and H_{set} is the desired liquid set point of 15. It can be seen by examining equation 11 that as time proceeds, the exponential term will reduce to zero, resulting in $H(t + h) = H_{set}$. The next step is to figure out how to produce a controller input that will give the desired behavior. This may be done by calculating the error between process model and desired behavior, and optimizing for u. Fortunately for us, this is a relatively simple set of equations that can be solved for u explicitly. By expanding Equation 12.3.10 using the fundamental theorem of calculus to produce Equation 12.3.12, and plugging Equation 12.3.11 in for $H(t + h)$, equation 13 results solved for u.

$$A\frac{H(t+h) - H(t)}{h} = F_{in} - Bu\sqrt{H(t)} \tag{12.3.12}$$

$$u = \frac{\left(-AH_{set} + AH(t) + Ae^{\frac{-h}{T}} * H_{set} - Ae^{\frac{-h}{T}}H(t) + F_{\text{in}}\, h\right)}{\left(B\sqrt{H_{set}}h\right)} \tag{12.3.13}$$

These equations can be implemented in a spreadsheet to model the behavior of the height of liquid in the tank, as well as the actions taken by the controller to restore the set point. Keep in mind the 0 to 25 output restriction on your controller.

This spreadsheet has been made with the relevant equations and can be used to solve this problem and model tank behavior. This spreadsheet can be found here:MPC example 1

So yes, the tank will overflow, and will return to its set point after 200 time points.

Example 12.3.2

What happens if the actuator on the valve breaks and must be replaced with an older one you happened to have in your basement? This new (although old) actuator cannot operate the valve fast enough, and B is decreased from 1 to 0.8. How does the graph from example 1 change? What does this tell you about the physical meaning of parameter B?

Solution

This B can easily be changed in the spreadsheet used in example 1. With this different actuator, and hence different B, we can see that the tank acquires more volume, and it takes longer for the tank to return to its optimal height of 15. This is because the valve on the tank cannot be opened as fast. Parameter B represents the physical properties/limitations of the valve, and pipe properties.

Sage's Corner

- www.youtube.com/v/XNccO4neVfc
- www.youtube.com/v/2HkpKgXlyfE

Exercise 12.3.1

Which of the following is a highly desired feature for model predictive control?

a. It is ideal in determining short-term behavior.
b. It is the simplest of all feed forward controller types.
c. It accounts for system constraints.
d. Controllers of this type are universal and can be used without modification for any process.

Answer available at the end of this section.

Exercise 12.3.2

Will a model predictive controller developed for a tank reactor be able to control an evaporation unit?

a. Yes, as they are both standard pieces of chemical engineering equipment.
b. Maybe, it depends if the reaction taking place in the reactor is exothermic or endothermic.
c. No, the governing equations will be very different.
d. Yes, because evaporation units and tank reactors can both use steam to supply heat.

Answer available at the end of this section.

Exercise 12.3.3

Which of the following statements about MPC is false?

a. MPC uses a defined time interval.
b. MPC utilizes feed forward control.
c. MPC requires a certain number of inputs and outputs.
d. MPC can generate more than one possible outcome.

Answer available at the end of this section.

References

- Astrom, Karl J. and Hagglund, Tore. *Advanced PID Control*, Department of Automatic Control, Lund Institute of Technology, Lund University.
- Bequette, B. Wayne. *Process Control: Modeling, Design and Simulation*. Prentice Hall PTR. pp. 487-511, December 26, 2002.
- Garcia C.E, Prett D.M. and Morari M. *Model Predictive Control: Theory and Practice - a Survey*. Automatica, Vol.25, No.3, pp. 335-348, 1989.
- Orukpe, P.E. *Basics of Model Predictive Control*. Imperial College, London. April 14, 2005.

Answers

Exercise 12.3.1
C

Exercise 12.3.2
C

Exercise 12.3.3
C

This page titled 12.3: MIMO using Model Predictive Control is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Aaron Bennick, Bradley Anderson, & Michael Salciccioli via source content that was edited to the style and standards of the LibreTexts platform.

12.4: Neural Networks for automatic model construction

Authors: Group E - Ardemis Boghossian, James Brown, Sara Zak

Introduction

Multiple Input Multiple Output (MIMOs) are systems that require multiple inputs and generate multiple outputs. MIMOs are controlled by controllers that combine multiple input readings in an algorithm to generate multiple output signals. MIMOs can be used with a variety of algorithms. The most versatile algorithm used to date is the neural network. Neural networks, which were initially designed to imitate human neurons, work to store, analyze, and identify patterns in input readings to generate output signals. In chemical engineering, neural networks are used to predict the ouputs of systems such as distillation columns and CSTRs. This article will discuss how neural networks work, the advantages and disadvantages of neural networks, and some common applications of the networks.

MIMOs

As mentioned, Multiple Inputs Multiple Outputs (MIMOs) are systems that require multiple inputs and generate multiple outputs, as shown schematically in Figure 1.

Figure 1. Schematic diagram of MIMOs

As shown in the figure, MIMOs are controlled by controllers that combine multiple input readings in an algorithm to generate multiple output signals. Typically, MIMOs do not require that the the number of inputs and outputs be the same. Instead, any number of input readings could be used to generate any number of output signals.

Neural Networks

Various types of controllers can be used to control MIMOs. One of the most accurate and versatile controllers used for MIMOs is the neural network. Neural networks are controllers that crudely imitate the human neuron. Initially, these networks were designed to model neural brain activity. However, as people began to recognize the advantages of neural networks, these networks were applied to controller algorithm design. Like a human neuron, neural networks work to store, analyze, and identify patterns in data by performing learning tasks. The ability of these networks to "learn" parallels the human neuron's ability to learn, making these automatic controllers the closest analog to a human controller.

Neurons

Like neurons in the body, network neurons receive inputs, store this data, and transmit outputs to either another neuron or directly to the MIMO. In order to transmit this data, the neuron must relate the multiple inputs to the multiple outputs. A simple mathematical representation of this relationship is shown below.

$$y = f\left(w_1a_1 + w_2a_2 + w_3a_3 + ... + w_na_n\right) = f\left(\sum_{i=1}^{n} w_ia_i\right)$$

where

- w_i = weight
- a_i = input

- y = output
- f = sigmoid function (any nonlinear function)

According to this relationship, the multiple input parameters are each multiplied by the corresponding weight factor, w_i. These weight factors "weigh" the significance of each input, scaling each input proportionally to the effect it will have on the output. These weighted inputs are then added and the sum is input into the sigmoid function to generate an output. This output can then be sent to multiple neurons that, in turn, each generate its own output.

The sigmoid function in this relationship is a nonlinear, empirical function that relates the input readings to the output signals. This empirical function can take on many forms depending on the data set. The equation that is best able to predict the outputs for the given system will be used (polynomial, sine, logarithmic, etc.). For example, one form this function may take is the hyperbolic sine function, where

$$f(x) = \sinh(\alpha x)$$

where

x = sum of weighted inputs = $\sum_{i=1}^{n} w_i a_i$

α = empirical parameter

In this sigmoid function, α is an empirical parameter that adjusts the function outputs. The effect of α on this sigmoid function is shown in Figure 2.

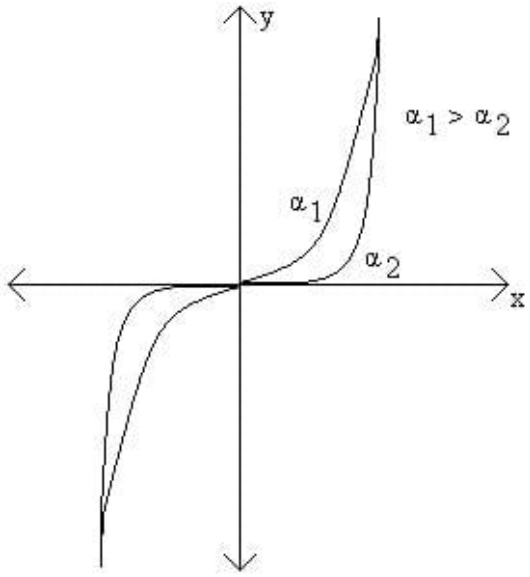


Figure 2. Plot of hyperbolic sine sigmoid function with varying α*x*

As shown in Figure 2, increasing α increases the output for this particular sigmoid function. Like most empirical functions, the hyperbolic sine sigmoid function can only be used within a specified range of values (x values between the vertical asymptotes). In this case, the range would depend on the value of α.

Combining Neurons into Neural Networks

Once neurons have been programmed to correlate input and output data, they can be connected in a feedforward series to produce a neural network, or neural net (NN). A schematic diagram of a neural network is shown in Figure 3.

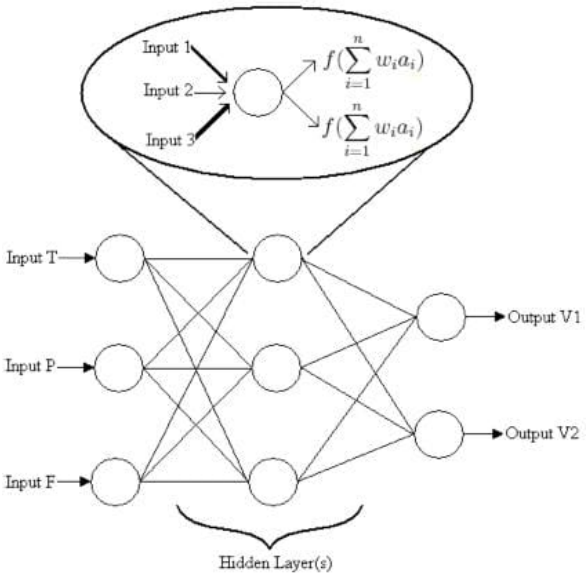


Figure 3. Schematic diagram of a neural network.

Figure 3 shows MIMO parameters, such as temperature, pressure, and flow readings, are first processed in the first layer of neurons. The outputs of the first layer of neurons then serve as the inputs to the second layer. The outputs of the second layer then become the inputs to the third layer, and so on, until the final output of the network is used to directly affect MIMO controls such as valves. The layers of neurons between the initial and final layers are known as hidden layers.

The way neurons within hidden layers correlate the inputs and outputs is analogous to the way individual neurons correlate these variables. As shown in the flowchart inset of the diagram, a neuron within the network receives multiple inputs. Each of these input parameters are then weighted, having the most effective parameters weighted more heavily. These weighted input values are added and input into the particular sigmoid function the neuron is programmed to follow. The output of this function is then sent to other neurons as input. These input values are then reweighed for that particular neuron. This process is continued until the final output of the neuron network is used to adjust the desired controls. Although the diagram shows only one hidden layer, these neuron networks can consist of multiple layers. Although almost all continuous functions can be approximated by a single hidden layer, incorporating multiple hidden layers decreases the numbers of weights used. Since more layers will result in more parameters, each of these individual parameters will be weighted less. With more parameters and less weights, the system becomes more sensitive to parameter change (a greater "rippling" effect within the network).

The "rippling" effect of a neuron network makes the system difficult to model analytically. A small change in a single input variable would result in multiple changes throughout the the entire network. Although modelling these complex networks is beyond the scope of the class, only a basic, qualitative understanding of how neural networks function is necessary to analyze neural network controllers and their effects on input and output parameters.

Learning Process

The ability of neural networks to learn distinguishes them from most automatic controllers. Like humans, neural networks learn by example, and thus need to be trained. Neural networks are usually configured to specific applications and have the ability to process large amounts of data. Complex trends and patterns can be detected by neural networks that would otherwise be imperceptible to humans or other computing programs. Within neural networks, there are learning procedures that allow the device to recognize a certain pattern and carry out a specific task. These learning procedures consist of an algorithm that enables the network to determine the weighting parameters in order to match the given data (inputs and outputs) with a function. In this iterative procedure, the initial input values are used to generate initial output values. Based on these input and output values, the weights within the network are adjusted to match the data. These adjusted weights are then used to correlate the next pair of input and output values. Again, these values are used to adjust weights. This process continues until the network obtains a good fit for the data. A flow chart summarizing this iterative process is shown in Figure 4.

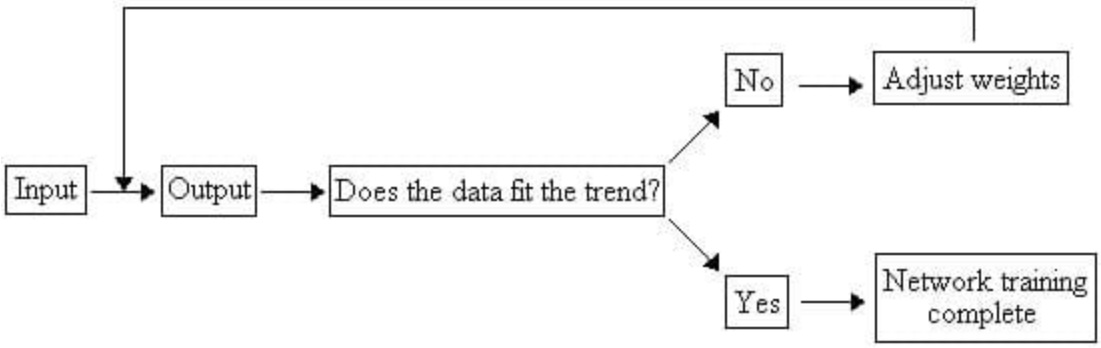


Figure 4. Flowchart for training a network.

The neural network learns by tracing the path of an object over a specific area and sequence. For this application, inputs for the neural network included the position, velocity, and direction. Once the the neural network has completed its training, it will use what it learned to predict trajectories. The neural network can predict where the object will move given the position, velocity, and direction of the object.

The trajectories shown in these videos demonstrate an important characteristic of neural networks. In these videos, the neural network can only predict outputs for objects that behave similar to the output it encountered during training. For instance, if a person was to start running forward, then stop, spin in a circle, and start running backwards, the neural network output would be unpredictable.

Once the neural network becomes capable of predicting system outputs, it is run like any other controller. In real-time, this often entails the chemical process and a computer system of some sort. Often, such as in LabVIEW, the chemical engineer is presented with a user-friendly data acquisition program that allows the user to set desired temperatures, flowrates, etc. and displays the system's input and outputs. Although the neural network continually sends signals to the system controllers (such as valves), the network algorithm is embedded in the acquisition program. Therefore, once the system is up and running, the user does not directly see the algorithm working.

Advantages and Disadvantages

Given the interesting, human-like behavior of neural networks, one would expect all process applications to be controlled by neural networks. However, the advantages and disadvantages of neural networks limit their use in applications. The following lists summarize these advantages and disadvantages.

Advantages

- Neural networks are very general and can capture a variety of patterns very accurately
- The static, non-linear function used by neural networks provide a method to fit the parameters of a particular function to a given set of data.
- A wide variety of functions can be used to fit a given set of data
- Neural networks do not require excessive statistical training
- There is no need to assume an underlying input data distribution when programming a neural network
- Neural networks can detect all possible, complex nonlinear relationships between input and outputs
- Neural networks are the closest thing to having an actual human operate a system (i.e. they can "learn")

Disadvantages

- Neural networks are difficult to design. One must determine the optimal number of nodes, hidden layers, sigmoid function, etc.
- Neural networks are diffcult to model analytically because a small change in a single input will affect the entire network
- The operation of neural networks is limited to the training process. If the network is trained poorly, then it will operate poorly and the outputs cannot be guaranteed.
- There is a great computational burden associated with neural networks
- Neural networks require a large sample size in order to empirically fit data
- Neural networks have a “black box” nature. Therefore, errors within the complex network are difficult to target.

- Outside of their data training range, neural networks are unpredictable. This occurs because neural networks may "overfit" data. For instance, during training, a neural network may fit a 10th order polynomial to only 5 data points. When using this model to make predictions of values outside this 5-point range, the neural network behaves unpredictably.

Applications of Neural Networks

Each neural network will generate a different algorithm based on the inputs and outputs to the system. Because neural networks fit a specific function to the given data, they can be used in a variety of applications. Within chemical engineering, neural networks are frequently used to predict how changing one input (such as pressure, temperature, etc.) on a distillation column will influence the compositions and flow rates of the streams exiting the column. The network training is performed on various inputs to the column, and thus can predict how changing one input will affect the product streams. Within a CSTR, neural networks can be used to determine the effect of one input parameter such as temperature or pressure on the products.

In addition to its applications to chemical equipment, neural networks can also be applied to model material responsess as a function of various loads under various conditions. These models can then be used in product development to create a device for a particular application, or to improve an existing device. For example, by modeling the corrosion of steel under different temperature and pH conditions, implanted biomedical devices can be manufactured or improved.

Neural networks are also often used in biology and biological applications to predict the outcome of a certain event. For example, neural networks can be used to predict the growth of cells and bacteria in cell culture labs, given a set of varying conditions, such as temperature and pH. In addition, neural network models have been used to predict the mortality rate in intensive care units in hospitals. Data was collected from different patients, and a neural network model was created to predict the mortality of future patients given a set of specified conditions. Neural networks have also been used to diagnose breast cancer in patients by predicting the effects of a tumor provided specified input conditions of the patient.

Neural networks are also used in applications beyond the chemical aspect of controllers. For instance, neural networks are used to predict travel time based on different travel conditions. Signs on the highways that give estimated travel times are examples of neural networks. They predict the amount of time required to reach a certain destination given the varying traffic volume, road conditions, and weather conditions.

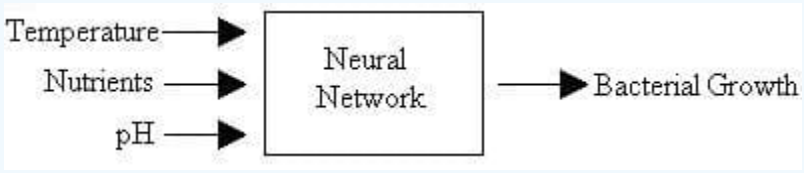
Neural networks are not limited to the applications listed above. They can be used to model most predictable events, and the complexity of the network will increase depending on the situation. For more information on the uses listed above, refer to the journal articles listed in the references section of this wiki.

Example 12.4.1

Hypothetical Industries has expanded to now include a biology lab! As an employee, you are working on developing a new antibiotic, and thus have been assigned to predict the growth of bacteria. Your boss wants to know how the growth of the bacteria is affected by different conditions, such as temperatures, pH, and provided nutrients. You don't really feel like manipulating all these conditions, and then sitting and watching bacteria grow on a petri dish, so you decide to come up with a way to predict the how the bacteria will grow. Using the information presented in the wiki, determine what the inputs and outputs to this neural network are.

Solution

The inputs to the neural network are each of the parameters that may affect the growth of the bacteria. In this situation, the inputs are the temperature, pH, and nutrients provided (such as sugars, amino acids, and antibiotics). The outputs from this system include the growth of the bacteria.



Example 12.4.2

As seen in the example above, a neural network can be used to predict bacterial growth. Given the information presented in the wiki, explain the advantages and disadvantages of using a neural network to model bacterial growth.

Solution

Advantages

- Because there are so many inputs (temperature, pH, etc.), a neural network fits a function to this data that is able to predict how future conditions would affect the bacterial growth.
- Neural networks provide a mechanical method of modeling the system of bacterial growth that is extremely similar to have an acutal human predict the growth of bacteria.
- Someone monitoring the system does not need much statistical training to use the neural network.

Disadvantages

- You must first run a large number of samples with varying conditions in order to obtain a good fit for the neural network. Based on how well you would like your network to predict bacterial growth, the number of samples that must be run in order to create the network changes. The network will only function as well as it is trained, so if a high degree of accuracy is desired, then more data inputs will be required. This becomes time comsuming and expensive.
- Because of the black box nature of the neural networks, it is difficult to determine how the individual parameters such as temperature or pH will affect the bacterial growth.
- Neural networks cannot be used to predict the growth patterns of the bacteria outside of the given data ranges.

Exercise 12.4.1

Why would someone want to increase the number of hidden layers when combining neurons into a neural network?

a. To decrease the amount of programming needed
b. To decrease the number of weights required
c. To increase the cost associated with the system
d. To increase the aesthetic structure of the system

Exercise 12.4.2

What does the neural network output if its inputs are outside its training range?

a. Zero
b. Input^2
c. sqrt(input)
d. The outputs outside of range are unpredictable

References

- Accurate freeway travel time prediction with state-space neural networks under missing data. Van-zuylen. Transportation research part c. 2005/10/12. 13(5-6) pp347-369.
- Astrom, K.J., Hagglund, T. Advanced PID Control. ISA- The Istrumentation Systems and Automation Society.
- Campos, Lucio P.A., Silva, Aristorfanes C. Barros, Allan Kardec. Diagnosis of breast cancer in digital mammograms using independent component analysis and neural networks. Lecture Notes in Computer Science, vol 3773, pub 2005, p460-469.
- Chan, CH, Chow, PY. Application of artificial neural networks to establish a predictive mortality risk model in children admitted to pediatric intensive care unit. Singapore Med J. Volume 47, Issue 11, Date 2006/10/31, pages 928-934.
- Yu, C. Davidson, VJ, Yang, SX. A neural network approach to predict durvival/death and growth/no-growth interfaces for Escherichia coli O157:H7. Food Microbiology [Food Microbiol.]. Vol. 23, no. 6, pp. 552-560. Sep 2006

This page titled 12.4: Neural Networks for automatic model construction is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

12.5: Understanding MIMO Control Through Two Tanks Interaction

We have been familiar with the models of single surge tank manipulated by first order process and two tanks in series manipulated by second order process, both of which are typical examples of **Single Input Single Output** (SISO) control. However, in the real chemical processes, there are always interations between the reactors. The following page will discuss the two tanks model by taking into consideration the interaction between the two tanks. To manipulate this model, we need to use Multiple Input Multiple Output (MIMO) control, which will add more complexity in understanding the overall process.

Two Tanks Interaction Model

The following figure, Figure 1, demonstrates how two tanks model works to interact each other. From Figure 1, we can see that the input of Tank1, x1, not only affects the level in Tank1, but also affects that in Tank2. The same goes for the input of Tank2, x2. We use the resistance, R3, to account for the interaction between the two tanks.

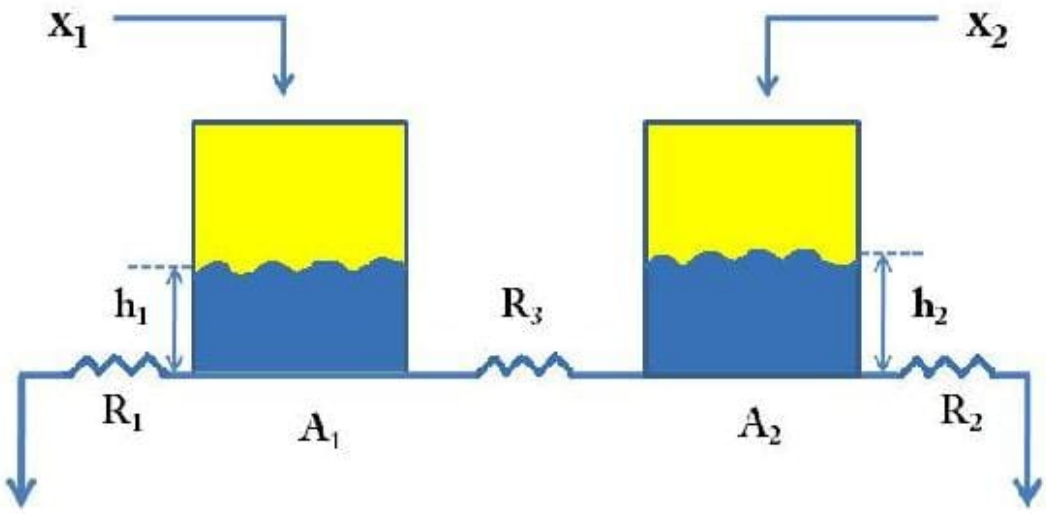


Figure 1: Two tanks interaction model

In the above figure,

- x_1 = Input of Tank1
- x_2 = Input of Tank2
- A_1 = Cross section area of Tank1
- A_2 = Cross section area of Tank2
- h_1 = Level of Tank1
- h_2 = Level of Tank2
- R_1 = Resistance of Tank1
- R_2 = Resistance of Tank2
- R_3 = Interaction Resistance

Mathematical Equations for the Process

Now, we will derive the mathematical equations to describe the process, now we assume the direction of the flow goes from Tank1 to Tank2:

To begin with, we write down the governing equation for each of the tanks, taking into consideration the interaction term:

$$A_1\frac{dh_1}{dt} = x_1 - \frac{h_1}{R_1}dt - \frac{(h_1-h_2)}{R_3} = x_1 + h_1\left(-\frac{1}{R_1}-\frac{1}{R_3}\right) + h_2\left(\frac{1}{R_3}\right)$$

$$A_2\frac{dh_2}{dt} = x_2 - \frac{h_2}{R_2}dt + \frac{(h_1-h_2)}{R_3} = x_2 + h_2\left(-\frac{1}{R_2}-\frac{1}{R_3}\right) + h_1\left(\frac{1}{R_3}\right)$$

Under steady state, the time derivatives, i.e. the left hand side of the above equations, go to zero:

$$0 = x_1(0) + h_1(0)\left(-\frac{1}{R_1}-\frac{1}{R_3}\right) + h_2(0)\left(\frac{1}{R_3}\right)$$

$$0 = x_2(0) + h_1(0)\left(\frac{1}{R_3}-\frac{1}{R_3}\right) + h_2(0)\left(-\frac{1}{R_2}-\frac{1}{R_3}\right)$$

Now, we define the deviation variables as follows:

- $X_1 = x_1 - x_1(0)$
- $Y_1 = h_1 - h_1(0)$
- $X_2 = x_2 - x_2(0)$
- $Y_2 = h_2 - h_2(0)$

Set $y_1 = h_1$ and $y_2 = h_2$, we can obtain the following equations:

$$A_1\frac{dy_1}{dt} = \frac{x_1}{A_1} + y_1\left(-\frac{1}{A_1R_1}-\frac{1}{A_1R_3}\right) + y_2\left(\frac{1}{A_1R_3}\right)$$

$$A_2\frac{dy_2}{dt} = \frac{x_2}{A_1} + y_1\left(\frac{1}{A_2R_3}\right) + y_2\left(-\frac{1}{A_2R_2}-\frac{1}{A_2R_3}\right)$$

It will be more general if we write down the above equations into a form of Matrix:

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{A_1} & 0 \\ 0 & \frac{1}{A_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} -\frac{1}{A_1R_1}-\frac{1}{A_1R_2} & \frac{1}{A_1R_3} \\ \frac{1}{A_2R_3} & -\frac{1}{A_2R_2}-\frac{1}{A_2R_3} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

We can represent the above equation using the following equation:

$$\vec{y} = \overrightarrow{\vec{B}}\,\vec{x} + \overrightarrow{\vec{A}}\,\vec{y}$$

$$\vec{y} - \vec{A}\vec{y} = \overrightarrow{\vec{B}}\,\vec{x}$$

$$\vec{x} = \left(\vec{B}^{-1}\right)\left(\frac{\partial}{\partial t} - \overrightarrow{\vec{A}}\right)\vec{y} = \overrightarrow{\vec{G}_p}^{-1}\,\vec{y}$$

Invert the above equation:

$$\overrightarrow{y} = \overrightarrow{\overrightarrow{G_p}}\,\overrightarrow{x}$$

Now, set

$$\overrightarrow{\overrightarrow{G_p}} = \begin{bmatrix} \hat{G_{p1,1}} & \hat{G_{p1,2}} \\ \hat{G_{p2,1}} & \hat{G_{p2,2}} \end{bmatrix}$$

We get the following equation:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \hat{G}_{p1,1} & \hat{G}_{p1,2} \\ \hat{G}_{p2,1} & \hat{G}_{p2,2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{12.5.1}$$

Control Diagram

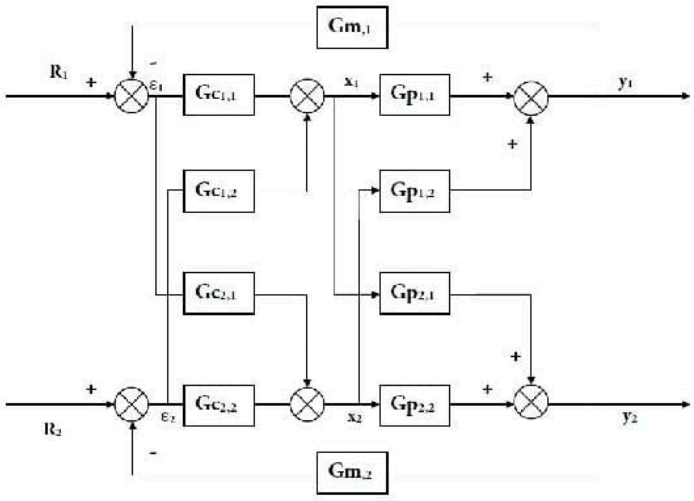


Figure 2: Contol Diagram for Two tanks interaction model.

Decouple the process

$$E_1 = R_1 + \widehat{G_{m1}}.Y_1$$

$$E_2 = R_2 + \widehat{G_{m2}}.Y_2$$

$$\begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} + \begin{bmatrix} \hat{G}_{m1} & 0 \\ 0 & \hat{G}_{m2} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

$$\vec{E} = \vec{R} - \widehat{\widehat{G_m}}.\vec{Y}$$

Summarizing All Equations

$$\left.\begin{aligned} \vec{Y} &= \widehat{\widehat{G_p}}.\vec{X} \\ \vec{X} &= \widehat{\widehat{G_c}}.\vec{E} \\ \vec{E} &= \vec{R} - \widehat{\widehat{G_m}}.\vec{Y} \end{aligned}\right\} \quad \vec{Y} = \widehat{\widehat{G_c}}\,\widehat{\widehat{G_p G_c}}\left(\vec{R} - \widehat{\widehat{G_m}}.\vec{Y}\right)$$

$$\vec{Y} + \widehat{\widehat{G_c}}\,\widehat{\widehat{G_p G_c}}\vec{Y} = \widehat{\widehat{G_p G_c}}\vec{R}$$

$$\left(\vec{I} + \widehat{\widehat{G_c}}\,\widehat{\widehat{G_p G_c}}\right)\vec{Y} = \widehat{\widehat{G_p G_c}}\vec{R}$$

$$\vec{Y} = \left(\vec{I} + \widehat{\widehat{G_c}}\,\widehat{\widehat{G_p G_c}}\right)^{-1}\widehat{\widehat{G_p G_c}}\vec{R}$$

If $\widehat{\widehat{G_p}}\,\widehat{\widehat{G_c}}$ is diagonal, i.e. $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$

Then $\vec{Y}$ = [Diagonal] $\vec{R}$

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$$

$$Y_1 = aR_1$$

$$Y_2 = aR_2$$

$$\widehat{\widehat{G_p}}\,\widehat{\widehat{G_c}} - \begin{bmatrix} G_{p1,1} & G_{p1,2} \\ G_{p2,1} & G_{p2,2} \end{bmatrix}.\begin{bmatrix} G_{c1,1} & G_{c1,2} \\ G_{c2,1} & G_{c2,2} \end{bmatrix} - \begin{bmatrix} G_{p1,1}.G_{c1,1} + G_{p1,2}.G_{c2,1} & G_{p1,1}.G_{c1,2} + G_{p1,2}.G_{c2,2} \\ G_{p2,1}.G_{c1,1} + G_{p2,2}.G_{c2,1} & G_{p2,1}.G_{c1,2} + G_{p2,2}.G_{c2,2} \end{bmatrix}$$

$$- \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

Using this identity, $G_{c1,2}$ & $G_{c2,1}$ can be calculated.

This page titled 12.5: Understanding MIMO Control Through Two Tanks Interaction is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Erin Baragar, Omar Ismail, & Yuan Chen via source content that was edited to the style and standards of the LibreTexts platform.

13: STATISTICS AND PROBABILITY BACKGROUND

13.1: Basic statistics- mean, median, average, standard deviation, z-scores, and p-value

13.2: SPC- Basic Control Charts- Theory and Construction, Sample Size, X-Bar, R charts, S charts

13.3: Six Sigma- What is it and what does it mean?

13.4: Bayes Rule, Conditional Probability and Independence

13.5: Bayesian Network Theory

13.6: Learning and analyzing Bayesian networks with Genie

13.7: Occasionally Dishonest Casino? Markov Chains and Hidden Markov Models

13.8: Continuous Distributions- normal and exponential

13.9: Discrete Distributions - Hypergeometric, Binomial, and Poisson

13.10: Multinomial Distributions

13.11: Comparisons of two means

13.12: Factor analysis and ANOVA

13.13: Correlation and Mutual Information

13.14: Random sampling from a stationary Gaussian process

This page titled 13: Statistics and Probability Background is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

13.1: Basic statistics- mean, median, average, standard deviation, z-scores, and p-value

Statistics is a field of mathematics that pertains to data analysis. Statistical methods and equations can be applied to a data set in order to analyze and interpret results, explain variations in the data, or predict future data. A few examples of statistical information we can calculate are:

- Average value (mean)
- Most frequently occurring value (mode)
- On average, how much each measurement deviates from the mean (standard deviation of the mean)
- Span of values over which your data set occurs (range), and
- Midpoint between the lowest and highest value of the set (median)

Statistics is important in the field of engineering by it provides tools to analyze collected data. For example, a chemical engineer may wish to analyze temperature measurements from a mixing tank. Statistical methods can be used to determine how reliable and reproducible the temperature measurements are, how much the temperature varies within the data set, what future temperatures of the tank may be, and how confident the engineer can be in the temperature measurements made. This article will cover the basic statistical functions of mean, median, mode, standard deviation of the mean, weighted averages and standard deviations, correlation coefficients, z-scores, and p-values.

What is a Statistic?

In the mind of a statistician, the world consists of populations and samples. An example of a population is all 7th graders in the United States. A related example of a sample would be a group of 7th graders in the United States. In this particular example, a federal health care administrator would like to know the average weight of 7th graders and how that compares to other countries. Unfortunately, it is too expensive to measure the weight of every 7th grader in the United States. Instead statistical methodologies can be used to estimate the average weight of 7th graders in the United States by measure the weights of a sample (or multiple samples) of 7th graders.

Parameters are to populations as statistics are to samples.

A parameter is a property of a population. As illustrated in the example above, most of the time it is infeasible to directly measure a population parameter. Instead a sample must be taken and statistic for the sample is calculated. This statistic can be used to estimate the population parameter. (A branch of statistics know as Inferential Statistics involves using samples to infer information about a populations.) In the example about the population parameter is the average weight of all 7th graders in the United States and the sample statistic is the average weight of a group of 7th graders.

A large number of statistical inference techniques require samples to be a single random sample and independently gathers. In short, this allows statistics to be treated as random variables. A in-depth discussion of these consequences is beyond the scope of this text. It is also important to note that statistics can be flawed due to large variance, bias, inconsistency and other errors that may arise during sampling. Whenever performing over reviewing statistical analysis, a skeptical eye is always valuable.

Statistics take on many forms. Examples of statistics can be seen below.

Basic Statistics

When performing statistical analysis on a set of data, the mean, median, mode, and standard deviation are all helpful values to calculate. The mean, median and mode are all estimates of where the "middle" of a set of data is. These values are useful when creating groups or bins to organize larger sets of data. The standard deviation is the average distance between the actual data and the mean.

Mean and Weighted Average

The mean (also know as average), is obtained by dividing the sum of observed values by the number of observations, *n*. Although data points fall above, below, or on the mean, it can be considered a good estimate for predicting subsequent data points. The formula for the mean is given below as Equation 13.1.1. The excel syntax for the mean is AVERAGE(starting cell: ending cell).

$$\bar{X} = \frac{\sum_{i=1}^{i=n} X_i}{n} \tag{13.1.1}$$

However, equation (1) can only be used when the error associated with each measurement is the same or unknown. Otherwise, the weighted average, which incorporates the standard deviation, should be calculated using equation (2) below.

$$X_{wav} = \frac{\sum w_i x_i}{\sum w_i} \tag{13.1.2}$$

where

$$w_i = \frac{1}{\sigma_i^2}$$

and x_i is the data value.

Median

The median is the middle value of a set of data containing an odd number of values, or the average of the two middle values of a set of data with an even number of values. The median is especially helpful when separating data into two equal sized bins. The excel syntax to find the median is MEDIAN(starting cell: ending cell).

Mode

The mode of a set of data is the value which occurs most frequently. The excel syntax for the mode is MODE(starting cell: ending cell).

Considerations

Now that we've discussed some different ways in which you can describe a data set, you might be wondering when to use each way. Well, if all the data points are relatively close together, the average gives you a good idea as to what the points are closest to. If on the other hand, almost all the points fall close to one, or a group of close values, but occasionally a value that differs greatly can be seen, then the mode might be more accurate for describing this system, whereas the mean would incorporate the occasional outlying data. The median is useful if you are interested in the range of values your system could be operating in. Half the values should be above and half the values should be below, so you have an idea of where the middle operating point is.

Standard Deviation and Weighted Standard Deviation

The standard deviation gives an idea of how close the entire set of data is to the average value. Data sets with a small standard deviation have tightly grouped, precise data. Data sets with large standard deviations have data spread out over a wide range of values. The formula for standard deviation is given below as Equation 13.1.3. The excel syntax for the standard deviation is STDEV(starting cell: ending cell).

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{i=n}\left(X_i - \bar{X}\right)^2} \tag{13.1.3}$$

Side Note: Bias Estimate of Population Variance

The standard deviation (the square root of variance) of a sample can be used to estimate a population's true variance. Equation 13.1.3 above is an unbiased estimate of population variance. Equation 13.1.4 is another common method for calculating sample standard deviation, although it is an bias estimate. Although the estimate is biased, it is advantageous in certain situations

because the estimate has a lower variance. (This relates to the bias-variance trade-off for estimators.)

$$\sigma_n = \sqrt{\frac{1}{n}\sum_{i=1}^{i=n}\left(X_i - \bar{X}\right)^2} \tag{13.1.4}$$

When calculated standard deviation values associated with weighted averages, Equation 13.1.5 below should be used.

$$\sigma_{wav} = \frac{1}{\sqrt{\sum w_i}} \tag{13.1.5}$$

The Sampling Distribution and Standard Deviation of the Mean

Population parameters follow all types of distributions, some are normal, others are skewed like the F-distribution and some don't even have defined moments (mean, variance, etc.) like the Chaucy distribution. However, many statistical methodologies, like a z-test (discussed later in this article), are based off of the normal distribution. How does this work? Most sample data are not normally distributed.

This highlights a common misunderstanding of those new to statistical inference. The distribution of the population parameter of interest and the *sampling distribution* are not the same. *Sampling distribution?!?* What is that?

Imagine an engineering is estimating the mean weight of widgets produced in a large batch. The engineer measures the weight of N widgets and calculates the mean. So far, one sample has been taken. The engineer then takes another sample, and another and another continues until a very larger number of samples and thus a larger number of mean sample weights (assume the batch of widgets being sampled from is near infinite for simplicity) have been gathered. The engineer has generated a sample distribution.

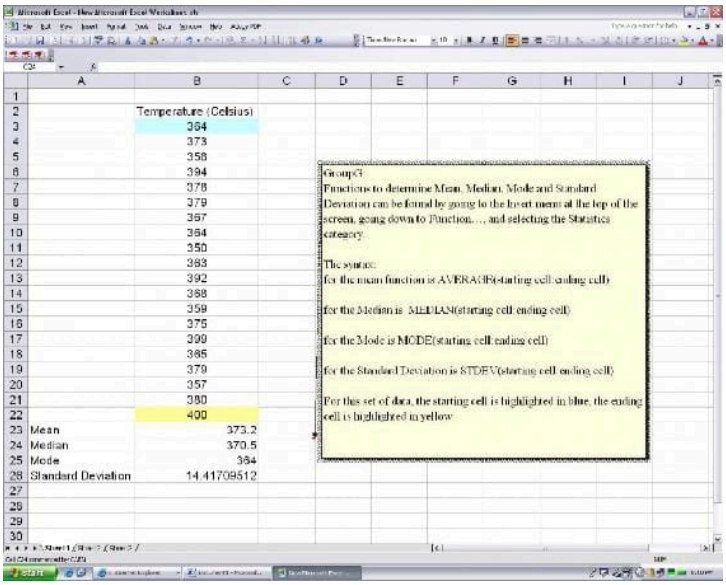
As the name suggested, a sample distribution is simply a distribution of a particular statistic (calculated for a sample with a set size) for a particular population. In this example, the statistic is mean widget weight and the sample size is N. If the engineer were to plot a histogram of the mean widget weights, he/she would see a bell-shaped distribution. This is because the Central Limit Theorem guarantees that as the sample size approaches infinity, the sampling distributions of statistics calculated from said samples approach the normal distribution.

Conveniently, there is a relationship between sample standard deviation (σ) and the standard deviation of the sampling distribution ($\sigma_{\bar{X}}$ - also know as the *standard deviation of the mean* or *standard error* deviation). This relationship is shown in Equation 13.1.6 below:

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{N}} \tag{13.1.6}$$

An important feature of the standard deviation of the mean, $\sigma_{\bar{X}}$ is the factor $\sqrt{N}$ in the denominator. As sample size increases, the standard deviation of the mean decrease while the standard deviation, σ does not change appreciably.

Microsoft Excel has built in functions to analyze a set of data for all of these values. Please see the screen shot below of how a set of data could be analyzed using Excel to retrieve these values.



Example by Hand

You obtain the following data points and want to analyze them using basic statistical methods. {1,2,2,3,5}

Calculate the average: Count the number of data points to obtain $n = 5$

$$mean = \frac{1+2+2+3+5}{5} = 2.6$$

Obtain the mode: Either using the excel syntax of the previous tutorial, or by looking at the data set, one can notice that there are two 2's, and no multiples of other data points, meaning the 2 is the mode.

Obtain the median: Knowing the n=5, the halfway point should be the third (middle) number in a list of the data points listed in ascending or descending order. Seeing as how the numbers are already listed in ascending order, the third number is 2, so the median is 2.

Calculate the standard deviation: Using Equation 13.1.3,

$$\sigma = \sqrt{\frac{1}{5-1}(1-2.6)^2 + (2-2.6)^2 + (2-2.6)^2 + (3-2.6)^2 + (5-2.6)^2} = 1.52$$

Example by Hand (Weighted)

Three University of Michigan students measured the attendance in the same Process Controls class several times. Their three answers were (all in units people):

- Student 1: A = 100 ± 3
- Student 2: A = 105 ± 4
- Student 3: A = 102 ± 2

What is the best estimate for the attendance A?

$$w_i = \frac{1}{{\sigma_i}^2}, \; w_1 = \frac{1}{9}, \; w_2 = \frac{1}{16}, \; w_3 = \frac{1}{4}$$

$$A_{wav} = \frac{\sum w_i A_i}{\sum w_i} = \frac{\frac{1}{9} * 100 + \frac{1}{16} * 105 + \frac{1}{4} * 102}{\frac{1}{9} + \frac{1}{16} + \frac{1}{4}} = 101.92 students$$

$$\sigma_{wav} = \frac{1}{\sqrt{\sum w_i}} = \frac{1}{\sqrt{\frac{1}{9} + \frac{1}{16} + \frac{1}{4}}} = .65$$

Therefore,

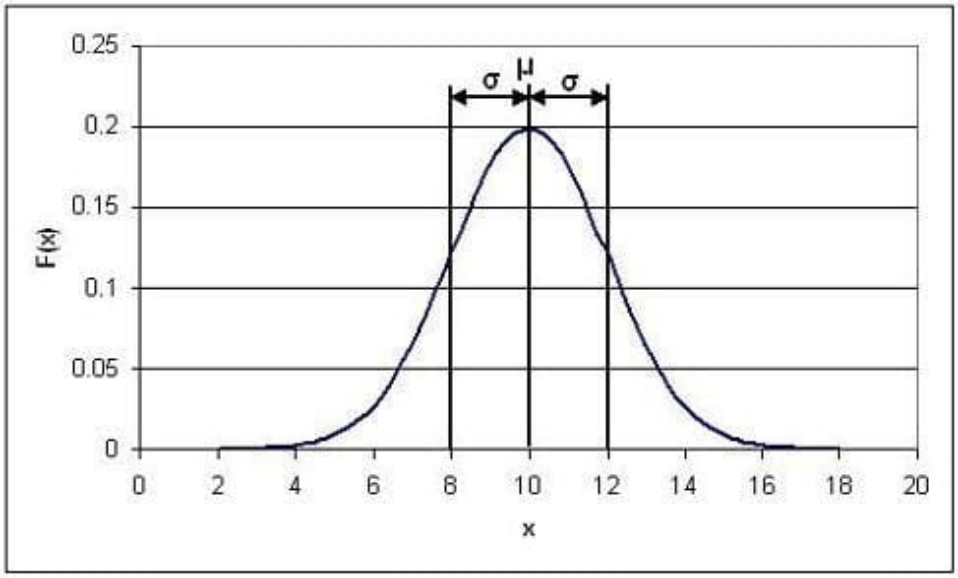
$$A = 101.92 \pm 0.65\, students$$

Gaussian Distribution

Gaussian distribution, also known as normal distribution, is represented by the following probability density function:

$$PDF_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where μ is the mean and σ is the standard deviation of a very large data set. The Gaussian distribution is a bell-shaped curve, symmetric about the mean value. An example of a Gaussian distribution is shown below.



In this specific example, μ = 10 and σ = 2.

Probability density functions represent the spread of data set. Integrating the function from some value *x* to *x* + *a* where *a* is some real value gives the probability that a value falls within that range. The total integral of the probability density function is 1, since every value will fall within the total range. The shaded area in the image below gives the probability that a value will fall between 8 and 10, and is represented by the expression:

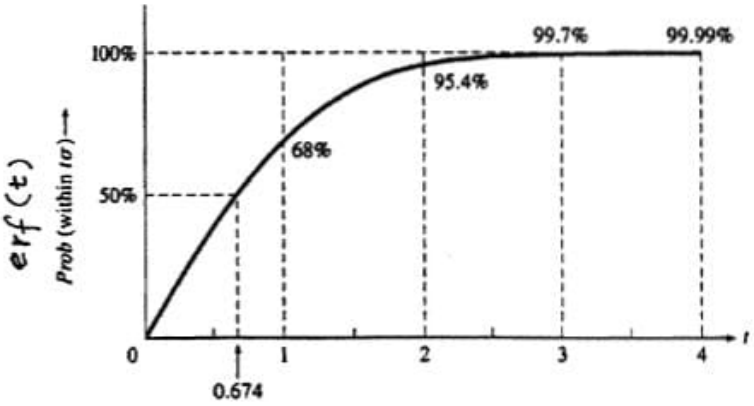
Gaussian distribution is important for statistical quality control, six sigma, and quality engineering in general. For more information see What is 6 sigma?.

Error Function

A **normal or Gaussian distribution** can also be estimated with a error function as shown in the equation below.

$$P(8 \leq x \leq 10) = \int_{8}^{10} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \text{erf}(t)$$

Here, erf(t) is called "error function" because of its role in the theory of normal random variable. The graph below shows the probability of a data point falling within t*σ of the mean.



t	0	0.25	0.5	0.75	1.0	1.25	1.5	1.75	2.0	2.5	3.0	3.5	4.0
Prob (%)	0	20	38	55	68	79	87	92	95.4	98.8	99.7	99.95	99.99

For example if you wanted to know the probability of a point falling within 2 standard deviations of the mean you can easily look at this table and find that it is 95.4%. This table is very useful to quickly look up what probability a value will fall into x standard deviations of the mean.

Correlation Coefficient (r value)

The linear correlation coefficient is a test that can be used to see if there is a linear relationship between two variables. For example, it is useful if a linear equation is compared to experimental points. The following equation is used:

$$r = \frac{\sum (X_i - X_{\text{mean}})(Y_i - Y_{\text{mean}})}{\sqrt{\sum (X_i - X_{\text{mean}})^2}\sqrt{\sum (Y_i - Y_{\text{mean}})^2}}$$

The range of r is from -1 to 1. If the r value is close to -1 then the relationship is considered anti-correlated, or has a negative slope. If the value is close to 1 then the relationship is considered correlated, or to have a positive slope. As the r value deviates from either of these values and approaches zero, the points are considered to become less correlated and eventually are uncorrelated.

There are also probability tables that can be used to show the significant of linearity based on the number of measurements. If the probability is less than 5% the correlation is considered significant.

Linear Regression

The correlation coefficient is used to determined whether or not there is a correlation within your data set. Once a correlation has been established, the actual relationship can be determined by carrying out a linear regression. The first step in performing a linear regression is calculating the slope and intercept:

\[\mathit{Slope} = \frac{n\sum_i X_iY_i -\sum_i X_i \sum_j Y_j }

ParseError: EOF expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

\nonumber \]

\[\mathrm{Intercept} = \frac{(\sum_i X_i^2)\sum_i(Y_i)-\sum_i X_i\sum_i X_iY_i }

ParseError: invalid ArgList (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

\nonumber \]

Once the slope and intercept are calculated, the uncertainty within the linear regression needs to be applied. To calculate the uncertainty, the standard error for the regression line needs to be calculated.

$$S = \sqrt{\frac{1}{n-2}\left(\left(\sum_i Y_i^2\right) - \text{ intercept } \sum Y_i - \text{slope}\left(\sum_i Y_i X_i\right)\right)}$$

The standard error can then be used to find the specific error associated with the slope and intercept:

$$S_{\text{slope}} = S\sqrt{\frac{n}{n\sum_i X_i^2 - \left(\sum_i X_i\right)^2}}$$

$$S_{\text{intercept}} = S\sqrt{\frac{\sum\left(X_i^2\right)}{n\left(\sum X_i^2\right) - \left(\sum_i X_i Y_i\right)^2}}$$

Once the error associated with the slope and intercept are determined a confidence interval needs to be applied to the error. A confidence interval indicates the likelihood of any given data point, in the set of data points, falling inside the boundaries of the uncertainty.

$$\beta = slope \pm \Delta slope \simeq slope \pm t^* S_{slope}$$

$$\alpha = intercept \pm \Delta intercept \simeq intercept \pm t^* S_{intercept}$$

Now that the slope, intercept, and their respective uncertainties have been calculated, the equation for the linear regression can be determined.

$$Y = \beta X + \alpha$$

Z-Scores

A z-score (also known as z-value, standard score, or normal score) is a measure of the divergence of an individual experimental result from the most probable result, the mean. Z is expressed in terms of the number of standard deviations from the mean value.

$$z = \frac{X - \mu}{\sigma} \tag{13.1.7}$$

- *X* = *ExperimentalValue*
- μ = *Mean*
- σ = *StandardDeviation*

Z-scores assuming the sampling distribution of the test statistic (mean in most cases) is normal and transform the sampling distribution into a standard normal distribution. As explained above in the section on sampling distributions, the standard deviation of a sampling distribution depends on the number of samples. Equation (6) is to be used to compare results to one another, whereas equation (7) is to be used when performing inference about the population.

Whenever using z-scores it is important to remember a few things:

- Z-scores normalize the sampling distribution for meaningful comparison.

- Z-scores require a large amount of data.
- Z-scores require independent, random data.

$$z_{obs} = \frac{X - \mu}{\frac{\sigma}{\sqrt{n}}} \tag{13.1.8}$$

- *n* = *SampleNumber*

P-Value

A p-value is a statistical value that details how much evidence there is to reject the most common explanation for the data set. It can be considered to be the probability of obtaining a result at least as extreme as the one observed, given that the null hypothesis is true. In chemical engineering, the p-value is often used to analyze marginal conditions of a system, in which case the p-value is the probability that the null hypothesis is true.

The null hypothesis is considered to be the most plausible scenario that can explain a set of data. The most common null hypothesis is that the data is completely random, that there is no relationship between two system results. The null hypothesis is always assumed to be true unless proven otherwise. An alternative hypothesis predicts the opposite of the null hypothesis and is said to be true if the null hypothesis is proven to be false.

The following is an example of these two hypotheses:

4 students who sat at the same table during in an exam all got perfect scores.

- *Null Hypothesis:* The lack of a score deviation happened by chance.
- *Alternative Hypothesis:* There is some other reason that they all received the same score.

If it is found that the null hypothesis is true then the Honor Council will not need to be involved. However, if the alternative hypothesis is found to be true then more studies will need to be done in order to prove this hypothesis and learn more about the situation.

As mentioned previously, the p-value can be used to analyze marginal conditions. In this case, the null hypothesis is that there is no relationship between the variables controlling the data set. For example:

- Runny feed has no impact on product quality
- Points on a control chart are all drawn from the same distribution
- Two shipments of feed are statistically the same

The p-value proves or disproves the null hypothesis based on its significance. A p-value is said to be significant if it is less than the level of significance, which is commonly 5%, 1% or .1%, depending on how accurate the data must be or stringent the standards are. For example, a health care company may have a lower level of significance because they have strict standards. If the p-value is considered significant (is less than the specified level of significance), the null hypothesis is false and more tests must be done to prove the alternative hypothesis.

Upon finding the p-value and subsequently coming to a conclusion to reject the *Null Hypothesis* or fail to reject the *Null Hypothesis*, there is also a possibility that the wrong decision can be made. If the decision is to reject the *Null Hypothesis* and in fact the *Null Hypothesis* is true, a type 1 error has occurred. The probability of a type one error is the same as the level of significance, so if the level of significance is 5%, "the probability of a type 1 error" is .05 or 5%. If the decision is to fail to reject the *Null Hypothesis* and in fact the *Alternative Hypothesis* is true, a type 2 error has just occurred. With respect to the type 2 error, if the *Alternative Hypothesis* is really true, another probability that is important to researchers is that of actually being able to detect this and reject the *Null Hypothesis*. This probability is known as the power (of the test) and it is defined as 1 - "probability of making a type 2 error."

If an error occurs in the previously mentioned example testing whether there is a relationship between the variables controlling the data set, either a type 1 or type 2 error could lead to a great deal of wasted product, or even a wildly out-of-control process. Therefore, when designing the parameters for hypothesis testing, researchers must heavily weigh their options for level of significance and power of the test. The sensitivity of the process, product, and standards for the product can all be sensitive to the smallest error.

Important Note About Significant P-values

If a P-value is greater than the applied level of significance, and the null hypothesis should not just be blindly accepted. Other tests should be performed in order to determine the true relationship between the variables which are being tested. More information on this and other misunderstandings related to P-values can be found at P-values: Frequent misunderstandings.

Calculation

There are two ways to calculate a p-value. The first method is used when the z-score has been calculated. The second method is used with the Fisher's exact method and is used when analyzing marginal conditions.

First Method: Z-Score

The method for finding the P-Value is actually rather simple. First calculate the z-score and then look up its corresponding p-value using the standard normal table.

This table can be found here: Media:Group_G_Z-Table.xls

This value represents the likelihood that the results are not occurring because of random errors but rather an actual difference in data sets.

To read the standard normal table, first find the row corresponding to the leading significant digit of the z-value in the column on the lefthand side of the table. After locating the appropriate row move to the column which matches the next significant digit.

Example:
If your z-score = 1.13

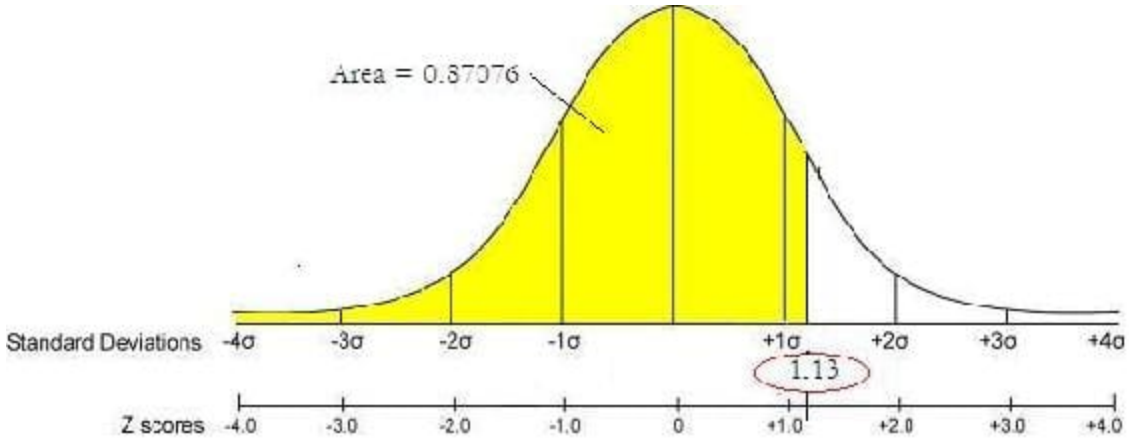
Follow the rows down to 1.1 and then across the columns to 0.03. The P-value is the highlighted box with a value of 0.87076.

z	0	0.01	0.02	0.03	0.04	0.05
0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63683
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894
1	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435
1.3	0.9032	0.9049	0.90658	0.90824	0.90988	0.91149
1.4	0.91924	0.92073	0.9222	0.92364	0.92507	0.92647
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943

Values in the table represent area under the standard normal distribution curve to the left of the z-score.

Using the previous example:

Z-score = 1.13, P-value = 0.87076 is graphically represented below.



Second Method: Fisher's Exact

In the case of analyzing marginal conditions, the P-value can be found by summing the Fisher's exact values for the current marginal configuration and each more extreme case using the same marginals. For information about how to calculate Fisher's exact click the following link:Discrete_Distributions:_hypergeometric,_binomial,_and_poisson#Fisher.27s_exact

Chi-Squared Test

A Chi-Squared test gives an estimate on the agreement between a set of observed data and a random set of data that you expected the measurements to fit. Since the observed values are continuous, the data must be broken down into bins that each contain some observed data. Bins can be chosen to have some sort of natural separation in the data. If none of these divisions exist, then the intervals can be chosen to be equally sized or some other criteria.

The calculated chi squared value can then be correlated to a probability using excel or published charts. Similar to the Fisher's exact, if this probability is greater than 0.05, the null hypothesis is true and the observed data is not significantly different than the random.

Calculating Chi Squared

The Chi squared calculation involves summing the distances between the observed and random data. Since this distance depends on the magnitude of the values, it is normalized by dividing by the random value

$$\chi^2 = \sum_{k=1}^{N} \frac{(observed - random)^2}{random}$$

or if the error on the observed value (sigma) is known or can be calculated:

$$\chi^2 = \sum_{k=1}^{N} \left(\frac{\text{observed} - \text{theoretical}}{\text{sigma}} \right)^2$$

Detailed Steps to **Calculate** Chi Squared by Hand

Calculating Chi squared is very simple when defined in depth, and in step-by-step form can be readily utilized for the estimate on the agreement between a set of observed data and a random set of data that you expected the measurements to fit. Given the data:

	1	2	3	4	5	6	7
x_i	12	14	17	28	26	40	45
y_i	9	18	23	31	47	51	53
σ_i	3	4	2	4	4	3	5

Step 1: Find χ_o^2

$$\chi_o^2 = \sum_i \frac{(y_i - A - Bx_i)^2}{\sigma_{yi}^2}$$

When:

$$A = \frac{S_{xx}S_y - S_x S_{xy}}{\Delta}$$

$$B = \frac{SS_{xy} - S_x S_y}{\Delta}$$

$$\Delta = SS_{xx} - (S_x)^2$$

$$S = \sum \frac{1}{\sigma_{yi}^2}$$

$$S_x = \sum \frac{x_i}{\sigma_{yi}^2}$$

$$S_y = \sum \frac{y_i}{\sigma_{yi}^2}$$

$$S_{xy} = \sum \frac{x_i y_i}{\sigma_{yi}^2}$$

$$S_{xx} = \sum \frac{x_i^2}{\sigma_{yi}^2}$$

The Excel function CHITEST(actual_range, expected_range) also calculates the χ_o^2 value. The two inputs represent the range of data the actual and expected data, respectively.

Step 2: Find the Degrees of Freedom

$$df = n - k$$

When: df = Degrees of Freedom

- n = number of observations
- k = the number of constraints

Step 3: Find $\tilde{\chi}_o^2 = \frac{\chi_o^2}{df}$

$\tilde{\chi}_o^2$

= the established value of $\tilde{\chi}^2$ obtained in an experiment with df degrees of freedom

Step 4: Find $P(\tilde{\chi}^2 \geq \tilde{\chi}_o^2)$ using Excel or published charts.

The Excel function CHIDIST(x,df) provides the p-value, where x is the value of the chi-squared statistic and df is the degrees of freedom. Note: Excel gives only the p-value and not the value of the chi-square statistic.

$P(\tilde{\chi}^2 \geq \tilde{\chi}_o^2)$ = the probability of getting a value of $\tilde{\chi}^2$ that is as large as the established $\tilde{\chi}_o^2$

Step 5: Compare the probability to the significance level (i.e. 5% or 0.05), if this probability is greater than 0.05, the null hypothesis is true and the observed data is not significantly different than the random. A probability smaller than 0.05 is an indicator of independence and a significant difference from the random.

Chi Squared Test versus Fisher's Exact

- For small sample sizes, the Chi Squared Test will not always produce an accurate probability. However, for a random null, the Fisher's exact, like its name, will always give an exact result.
- Chi Squared will not be correct when:

1. fewer than 20 samples are being used
2. if an expected number is 5 or below and there are between 20 and 40 samples

- For large contingency tables and expected distributions that are not random, the p-value from Fisher's Exact can be a difficult to compute, and Chi Squared Test will be easier to carry out.

Binning in Chi Squared and Fisher's Exact Tests

When performing various statistical analyzes you will find that Chi-squared and Fisher's exact tests may require binning, whereas ANOVA does not. Although there is no optimal choice for the number of bins (k), there are several formulas which can be used to calculate this number based on the sample size (N). One such example is listed below:

$$k = 1 + log_2 N$$

Another method involves grouping the data into intervals of equal probability or equal width. The first approach in which the data is grouped into intervals of equal probability is generally more acceptable since it handles peaked data much better. As a stipulation, each bin should contain at least 5 or more data points, so certain adjacent bins sometimes need to be joined together for this condition to be satisfied. Identifying the number the bins to use is important, but it is even more important to be able to note which situations call for binning. Some Chi-squared and Fisher's exact situations are listed below:

- **Analysis of a continuous variable:**

This situation will require binning. The idea is to divide the range of values of the variable into smaller intervals called bins.

- **Analysis of a discrete variable:**

Binning is unnecessary in this situation. For instance, a coin toss will result in two possible outcomes: heads or tails. In tossing ten coins, you can simply count the number of times you received each possible outcome. This approach is similar to choosing two bins, each containing one possible result.

- **Examples of when to bin, and when not to bin:**

- You have twenty measurements of the temperature inside a reactor: as temperature is a continuous variable, you should bin in this case. One approach might be to determine the mean (X) and the standard deviation (σ) and group the temperature data into four bins: T < X – σ, X – σ < T < X, X < T < X + σ, T > X + σ
- You have twenty data points of the heater setting of the reactor (high, medium, low): since the heater setting is discrete, you should not bin in this case.

Example 13.1.1

Say we have a reactor with a mean pressure reading of 100 and standard deviation of 7 psig. Calculate the probability of measuring a pressure between 90 and 105 psig.

Solution

To do this we will make use of the z-scores.

$$\Pr(a \leq z \leq b) = F(b) - F(a) = F\left(\frac{b-\mu}{\sigma}\right) - F\left(\frac{a-\mu}{\sigma}\right)$$

where a is the lower bound and b is the upper bound

Substitution of z-transformation equation (3)

$$z = \frac{X - \mu}{\sigma}$$

$$z_{upper} = \frac{b-\mu}{\sigma} = \frac{105 - 100}{7} = 0.714$$

$$z_{lower} = \frac{a-\mu}{\sigma} = \frac{90 - 100}{7} = -1.43$$

Look up z-score values in a standard normal table. Media:Group_G_Z-Table.xls

$$z_{upper} = 0.714 \rightarrow F(z_{upper}) = 0.76155$$
$$z_{lower} = -1.43 \rightarrow F(z_{lower}) = 0.07636$$

So:

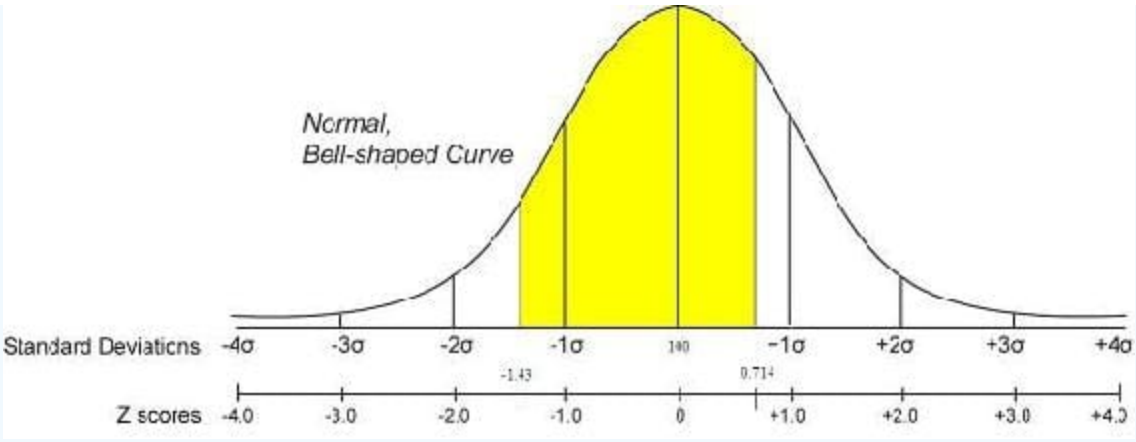
$$Pr(a \leq z \leq b) = F(.714) - F(-1.43)$$

= 0.76155 - 0.07636

= 0.68479.

The probability of measuring a pressure between 90 and 105 psig is **0.68479**.

A graphical representation of this is shown below. The shaded area is the probability



Alternate Solution

We can also solve this problem using the probability distribution function (PDF). This can be done easily in Mathematica as shown below. More information about the PDF is and how it is used can be found in the Continuous Distribution article

PDF is described by the following equation :

$$f(x) = \left(\frac{1}{\sigma * \sqrt{2\pi}}\right) * \text{Exp}\left(\frac{-1}{2} * \left(\frac{x-\mu}{\sigma}\right)^2\right)$$

Given $\mu = 100$ and $\sigma = 7$ we can plug these values in the equation to get :

$$f(x) = \left(\frac{1}{7 * \sqrt{2\pi}}\right) * e^{\left(\frac{-1}{2} * \left(\frac{x-100}{7}\right)^2\right)}$$

To find the probability between the two points 90 and 105, we preform the following integration:

$$f(90 \le x \le 105) = \int_{90}^{105} \left(\frac{1}{7 * \sqrt{2\pi}}\right) * e^{\left(\frac{-1}{2} * \left(\frac{x-100}{7}\right)^2\right)} dx$$

This results in the following expression :

$$f(90 \le x \le 105) = \frac{1}{2} * \left(\text{Erf}\left[\frac{(105 - 100)}{7 * \sqrt{2}}\right] - \text{Erf}\left[\frac{(90 - 100)}{7 * \sqrt{2}}\right]\right)$$

$$= \frac{1}{2}\left(\text{Erf}\left[\frac{5}{7\sqrt{2}}\right] + \text{Erf}\left[\frac{5\sqrt{2}}{7}\right]\right)$$

Or Simply entered in mathematica as

```
In[3]:= f = (1/(7 * Sqrt[2 * Pi])) * E^((1/2) * ((x - 100)/7)^2)
         NIntegrate[f, {x, 90, 105}]
Out[3]= E^((1/98) (100 - x)^2) / (7 Sqrt[2 Pi])
Out[4]= 0.685911
```

As you can see the the outcome is approximately the same value found using the z-scores.

Example 13.1.2

You are a quality engineer for the pharmaceutical company “Headache-b-gone.” You are in charge of the mass production of their children’s headache medication. The average weight of acetaminophen in this medication is supposed to be 80 mg, however when you run the required tests you find that the average weight of 50 random samples is 79.95 mg with a standard deviation of .18.

a. Identify the null and alternative hypothesis.
b. Under what conditions is the null hypothesis accepted?
c. Determine if these differences in average weight are significant.

Solution

a)

- Null hypothesis: This is the claimed average weight where H_0=80 mg
- Alternative hypothesis: This is anything other than the claimed average weight (in this case H_a<80)

b) The null hypothesis is accepted when the p-value is greater than .05.

c) We first need to find Z_{obs} using the equation below:

$$z_{obs} = \frac{X - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Where *n* is the number of samples taken.

$$z_{obs} = \frac{79.95 - 80}{\frac{.18}{\sqrt{50}}} = -1.96$$

Using the z-score table provided in earlier sections we get a p-value of .025. Since this value is less than the value of significance (.05) we reject the null hypothesis and determine that the product does not reach our standards.

Example 13.1.3

15 students in a controls class are surveyed to see if homework impacts exam grades. The following distribution is observed.

	High Homework Grades	Low Homework Grades	Totals
High Exam Grades	6	3	9
Low Exam Grades	2	4	6
Totals	8	7	15

Determine the p-value and if the null hypothesis (Homework does not impact Exams) is significant by a 5% significance level using the P-fisher method.

Solution

To find the p-value using the p-fisher method, we must first find the p-fisher for the original distribution. Then, we must find the p-fisher for each more extreme case. The p-fisher for the original distribution is as follows.

$$p_{\text{fisher}} = \frac{9!6!8!7!}{15!6!3!2!4!} = 0.195804$$

To find the more extreme case, we will gradually decrease the smallest number to zero. Thus, our next distribution would look like the following.

	High Homework Grades	Low Homework Grades	Totals
High Exam Grades	7	2	9
Low Exam Grades	1	5	6
Totals	8	7	15

The p-fisher for this distribution will be as follows.

$$p_{\text{fisher}} = \frac{9!6!8!7!}{15!7!2!1!5!} = 0.0335664$$

The final extreme case will look like this.

	High Homework Grades	Low Homework Grades	Totals
High Exam Grades	8	1	9
Low Exam Grades	0	6	6
Totals	8	7	15

The p-fisher for this distribution will be as follows.

$$p_{\text{fisher}} = \frac{9!6!8!7!}{15!8!1!0!6!} = 0.0013986$$

Since we have a 0 now in the distribution, there are no more extreme cases possible. To find the p-value we will sum the p-fisher values from the 3 different distributions.

$$p_{value} = 0.195804 + 0.0335664 + 0.0013986 = 0.230769$$

Because p-value=0.230769 we cannot reject the null hypothesis on a 5% significance level.

Application: What do p-values tell us?

Population Example

Out of a random sample of 400 students living in the dormitory (group A), 134 students caught a cold during the academic school year. Out of a random sample of 1000 students living off campus (group B), 178 students caught a cold during this same time period.

Population table

	Graup A	Group B	
Sick	$a = 134$	$b = 178$	$a + b = 312$
Not Sick	$c = 266$	$d = 822$	$c + d = 1088$
	$a + c = 400$	$b + d = 1000$	$a + b + c + d = 1400$

Fisher's Exact:

$$p_f = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{(a+b+c+d)!a!b!c!d!}$$

Solve:

$$p_f = \frac{(312)!(1088)!(400)!(1000)!}{(1400)!134!178!266!822!}$$

p_f = 2.28292 * 10 $^{-10}$

Comparison and interpretation of p-value at the 95% confidence level

This value is very close to zero which is much less than 0.05. Therefore, the number of students getting sick in the dormitory is significantly higher than the number of students getting sick off campus. There is more than a 95% chance that this significant difference is not random. Statistically, it is shown that this dormitory is more condusive for the spreading of viruses. With the knowledge gained from this analysis, making changes to the dormitory may be justified. Perhaps installing sanitary dispensers at common locations throughout the dormitory would lower this higher prevalence of illness among dormitory students. Further research may determine more specific areas of viral spreading by marking off several smaller populations of students living in different areas of the dormitory. This model of significance testing is very useful and is often applied to a multitude of data to determine if discrepancies are due to chance or actual differences between compared samples of data. As you can see, purely mathematical analyses such as these often lead to physical action being taken, which is necessary in the field of Medicine, Engineering, and other scientific and non-scientific venues.

Exercise 13.1.1

You are given the following set of data: {1,2,3,5,5,6,7,7,7,9,12} What is the mean, median and mode for this set of data? And then the z value of a data point of 7?

a. 5.82, 6, 7, 0.373
b. 6, 7, 5.82, 6.82
c. 7, 6, 5, 0.373
d. 7, 6, 5.82, 3.16

Answer available at the end of this section.

Exercise 13.1.2

What is n and the standard deviation for the above set of data {1,2,3,5,5,6,7,7,7,9,12}? And then consulting the table from above, what is the p-value for the data "12"?

a. 12, 3.16, 5.82
b. 7, 3.16, 0.83
c. 11, 3.16, 0.97
d. 11, 5.82, 0

Answer available at the end of this section.

References

- Woolf P., Keating A., Burge C., and Michael Y.. "Statistics and Probability Primer for Computational Biologists". Massachusetts Institute of Technology, BE 490/ Bio7.91, Spring 2004
- Smith W. and Gonic L. "Cartoon Guide to Statistics". Harper Perennial, 1993.
- Taylor, J. "An Introduction to Error Analysis". Sausalito, CA: University Science Books, 1982.
- http://www.fourmilab.ch/rpkp/experiments/analysis/zCalc.html

Answers

Exercise 13.1.1

a

Exercise 13.1.2

c

This page titled 13.1: Basic statistics- mean, median, average, standard deviation, z-scores, and p-value is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Andrew MacMillan, David Preston, Jessica Wolfe, & Sandy Yu via source content that was edited to the style and standards of the LibreTexts platform.

13.2: SPC- Basic Control Charts- Theory and Construction, Sample Size, X-Bar, R charts, S charts

Authors: Chris Bauman, Jennifer De Heck, Evan Leonard, Merrick Miranda
Stewards: Eric Black, Stacy Young, Dan Carter, Megan Boekeloo

Introduction

Control charts are one of the most commonly used methods of Statisical Process Control (SPC), which monitors the stability of a process. The main features of a control chart include the data points, a centerline (mean value), and upper and lower limits (bounds to indicate where a process output is considered "out of control").They visually display the fluctuations of a particular process variable, such as temperature, in a way that lets the engineer easily determine whether these variations fall within the specified process limits. Control charts are also known as Shewhart charts after Walter Shewhart, who developed them in the early 1900's.

Control Chart Background

A process may either be classified as in control or out of control. The boundaries for these classifications are set by calculating the mean, standard deviation, and range of a set of process data collected when the process is under stable operation. Then, subsequent data can be compared to this already calculated mean, standard deviation and range to determine whether the new data fall within acceptable bounds. For good and safe control, subsequent data collected should fall within three standard deviations of the mean. Control charts build on this basic idea of statistical analysis by plotting the mean or range of subsequent data against time. For example, if an engineer knows the mean (grand average) value, standard deviation, and range of a process, this information can be displayed as a bell curve, or population density function (PDF). The image below shows the control chart for a data set with the PDF overlay.

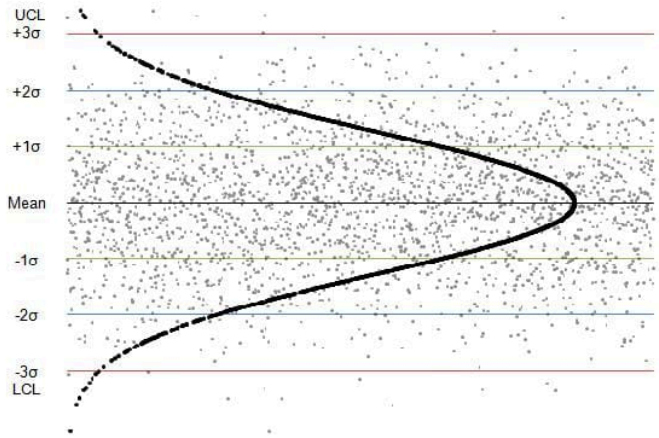


Figure I. Control chart showing PDF for a data set

The centerline is the mean value of the data set and the green, blue and red lines represent one, two, and three standard deviations from the mean value. In generalized terms, if data points fall within three standard deviations of the mean (within the red lines), the process is considered to be in control. These rules are discussed in greater detail later in this section.

Control Charts are commonly used in six sigma control today, as a means of overall process improvement. For more on six-sigma control, see six sigma.

Control Chart Functions

The main purpose of using a control chart is to monitor, control, and improve process performance over time by studying variation and its source. There are several functions of a control chart:

1. It centers attention on detecting and monitoring process variation over time.
2. It provides a tool for ongoing control of a process.
3. It differentiates special from common causes of variation in order to be a guide for local or management action.
4. It helps improve a process to perform consistently and predictably to achieve higher quality, lower cost, and higher effective capacity.
5. It serves as a common language for discussing process performance.

Sample Size and Subgrouping

There are a few key conditions that must be met when constructing control charts:

- The initial predictions for the process must be made while the process is assumed to be stable. Because future process quality will be compared to these predictions, they must be based off of a data set that is taken while the operation is running properly.
- Multiple subsets of data must be collected, where a subset is simply a set of n measurements taken over a specific time range. The number of subsets is represented as k. A subset average, subset standard deviation, and subset range will be computed for each subset.
- From these subsets, a grand average, an average standard deviation, and an average range are calculated. The grand average is the average of all subset averages. The average standard deviation is simply the average of subset standard deviations. The average range is simply the average of subset ranges.

The upper and lower control limits for the process can then be determined from this data.

- Future data taken to determine process stability can be of any size. This is because any point taken should fall within the statistical predictions. It is assumed that the first occurrence of a point not falling within the predicted limits shows that the system must be unstable since it has changed from the predictive model.
- The subsets are defined, based on the data and the process. For example, if you were using a pH sensor, the sensor would most likely output pages of data daily. If you know that your sensor has the tendency to drift every day, you might select a 30 minute subset of data. If it drifts monthly you might set your subset to be 24 hours or 12 hours.
- Finally, the population size, N is assumed to be infinite. Alternatively, if the population is finite but the sample size is less than 5% of the population size, we can still approximate the population to be near infinite. That is, n/N <= 0.05 where n is the sample size and N is the population size. [5]

X-Bar, R-Charts, and S-Charts

There are three types of control charts used determine if data is out of control, x-bar charts, r-charts and s-charts. An x-bar chart is often paired with either an r-chart or an s-chart to give a complete picture of the same set of data.

Pairing X-Bar with R-Charts

X-Bar (average) charts and R (range) -charts are often paired together. The X-Bar chart displays the centerline, which is calculated using the grand average, and the upper and lower control limits, which are calculated using the average range. Future experimental subsets are plotted compared to these values. This demonstrates the centering of the subset values. The R-chart plots the average range and the limits of the range. Again, the future experimental subsets are plotted relative to these values. The R-chart displays the dispersion of the subsets. X-Bar/R-Chart plot a subgroup average. Note that they should only be used when subgroups really make sense. For example, in a Gage R&R study, when operators are testing in duplicates or more, subgrouping really represents the same group.

Pairing X-Bar with S-Charts

Alternatively, X-Bar charts can be paired with S-charts (standard deviation). This is typically done when the size of the subsets are large. For larger subsets, the range is a poor statistic to estimate the distributions of the subsets, and instead, standard deviation

is used. In this case, the X-Bar chart will display control limits that are calculated using the average standard deviation. The S-Charts are similar to the R-charts; however, instead of the range, they track the standard deviation of multiple subsets.

Smoothing Data with a Moving Average

If it is desired to have smooth data, the moving average method is one option. This method involves taking the average of a number of points, and using that average for the middle data point. From this point on, the data is treated the same as any normal group of k subsets. Though this method will produce a smoother curve, it has a lag in detecting points, which may be problematic if the points are out of the acceptable range. This time lag would keep the control system from reacting to the problem until after the average is found. For this reason, moving average charts are appropriate mainly for slower processes that can handle the lag.

For example, let us calculate a value for a set of data which takes samples every second. We will use an average of 10 points to find this, however, in practice there is no set number of data points that should be used. For the point t = 50, we must wait until data has been collected through t = 54. The points are then averaged for t = 45-54 and used as the function value. For the next point, t = 51, the average of the points for t = 46-55 are used, and so on. If this is still confusing, please see moving average for a more detailed explanation.

Reading Control Charts

Control charts can determine whether a process is behaving in an "unusual" way.

Note: The upper and lower control limits are calculated using the grand average and either the average range and average sigma. Example calculations are shown in the Creating Control Charts Section.

The quality of the individual points of a subset is determined unstable if any of the following occurs:

- Rule 1: Any point falls beyond 3σ from the centerline(this is represented by the upper and lower control limts).
- Rule 2: Two out of three consecutive points fall beyond 2σ on the same side of the centerline.
- Rule 3: Four out of five consecutive points fall beyond 1σ on the same side of the centerline.
- Rule 4: Nine or more consecutive points fall on the same side of the centerline.

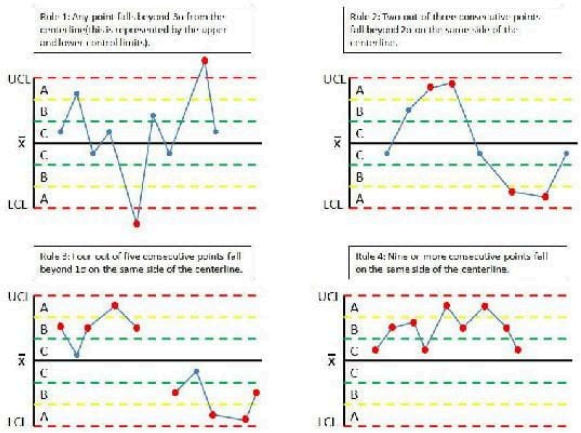


Figure III. Quality control rules.

The quality of a subset is determined unstable according to the following rules:

1. Any subset value is more than three standard deviations from the centerline.

2. Two consecutive subset values are more than two standard deviations from the centerline and are on the same side of the centerline.

3. Three consecutive subset values are more than one standard deviation from the centerline and are on the same side of the

centerline.

Creating Control Charts

To establish upper and lower control limits on control charts, there are a number of methods. We will discuss the method for the number of components in a subset, n, less than 15. For methods involving n > 15 and other techniques, see Process Control and Optimization, Liptak, 2.34. Here, the table of constants for computing limits, and the limit equations are presented below.

Please note that Table A below does NOT contain data for a sample problem. Any time you make a control chart, you refer to this table. The values in the table are used in the equations for the upper control limit (UCL), lower control limit (LCL), etc. This will be explained in the examples below. If you are interested in how these constants were derived, there is a more detailed explanation in Control Chart Constants.

Subgroup	X-bar chart		S-chart		R-chart	
	Using Ra	Using Sa				
n	A2	A3	B3	B4	D3	D4
2	1.886	2.659	0	3.267	0	3.268
3	1.023	1.954	0	2.568	0	2.574
4	0.729	1.628	0	2.266	0	2.282
5	0.577	1.427	0	2.089	0	2.114
6	0.483	1.287	0.03	1.97	0	2.004
7	0.419	1.182	0.118	1.882	0.076	1.924
8	0.373	1.099	0.185	1.815	0.136	1.864
9	0.337	1.032	0.239	1.761	0.184	1.816
10	0.308	0.975	0.284	1.716	0.223	1.777
11	0.285	0.927	0.322	1.678	0.256	1.744
12	0.266	0.886	0.354	1.646	0.283	1.717
13	0.249	0.85	0.382	1.619	0.307	1.693
14	0.235	0.817	0.407	1.593	0.328	1.672
15	0.223	0.789	0.428	1.572	0.347	1.653

Table A: Table of Constants

To determine the value for n, the number of subgroups

In order to determine the upper (UCL) and lower (LCL) limits for the x-bar charts, you need to know how many subgroups (n) there are in your data. Once you know the value of n, you can obtain the correct constants (A2, A3, etc.) to complete your control chart. This can be confusing when you first attend to create a x-bar control chart. The value of n is the number of subgroups within each data point. For example, if you are taking temperature measurements every min and there are three temperature readings per minute, then the value of **n** would be **3**. And if this same experiment was taking four temperature readings per minute, then the value of **n** would be **4**. Here are some examples with different tables of data to help you further in determining **n**:

Subset #	Values (kg)
1 (control)	1.02, 1.03, 0.98, 0.99
2 (control)	0.96, 1.01, 1.02, 1.01
3 (control)	0.99, 1.02, 1.03, 0.98
4 (control)	0.96, 0.97, 1.02, 0.98
5 (control)	1.03, 1.04, 0.95, 1.00
6 (control)	0.99, 0.99, 1.00, 0.97
7 (control)	1.02, 0.98, 1.01, 1.02
8 (experimental)	1.02, 0.99, 1.01, 0.99
9 (experimental)	1.01, 0.99, 0.97, 1.03
10 (experimental)	1.02, 0.98, 0.99, 1.00
11 (experimental)	0.98, 0.97, 1.02, 1.03

Example 1:

n= 4 since there are four readings of kg.

time (hours)	pH			
1	7.00	7.30	6.99	7.00
2	7.12	7.25	7.12	7.20
3	7.20	7.16	7.20	7.16
4	6.98	7.00	6.94	7.00
5	6.99	6.99	6.99	6.98
6	7.00	6.93	7.02	6.93
7	6.92	7.00	6.92	7.02
8	6.88	6.82	6.94	6.99
9	7.10	7.00	7.00	7.00
10	7.21	7.02	7.21	7.04
11	7.01	6.86	7.01	6.90
12	6.86	6.98	6.90	6.98
13	6.90	7.00	6.87	7.00
14	7.01	7.04	7.01	7.05
15	7.00	6.95	7.00	6.99
16	7.09	7.20	7.03	7.20
17	6.89	7.14	6.87	7.15
18	6.98	6.80	6.98	6.89
19	7.00	6.90	7.00	6.90
20	7.20	7.00	7.23	7.00
21	7.04	7.03	7.08	7.00
22	6.90	6.92	6.98	6.92
23	7.00	7.00	7.00	7.00
24	7.00	6.97	7.01	6.98

Example 2:

n= 4 since there are four readings of pH.

Example 3:

time (min)	T1	T2	T3
0	305.1578	311.1926	303.0032
1	308.6441	299.2898	307.9012
2	304.4789	308.7662	312.273
3	303.2384	303.7872	308.4915
4	316.6728	303.9563	303.3419
5	297.3459	308.0937	306.353
6	310.0358	304.9309	304.5568
7	302.2579	304.0973	317.315
8	305.5338	308.5081	308.1174
9	311.6743	302.4106	305.5727
10	303.535	312.9508	305.1281
11	307.5137	312.0491	307.6593
12	310.6001	305.5229	311.1861
13	307.6121	313.0331	313.4924
14	313.2346	312.1953	297.2964
15	306.0061	301.9239	298.6282
16	310.8455	308.7776	300.404
17	306.6952	299.0904	304.7548
18	305.2398	307.3239	297.1759
19	303.3781	305.8241	306.5276
20	309.3113	316.0451	309.9065

n= 3 since there are three readings of temperature.

After creating multiple control charts, determining the value of n will become quite easy.

Calculating UCL and LCL

For the X-Bar chart the following equations can be used to establish limits, where X_{GA} is the grand average, R_A is the average range, and S_A is the average standard deviation.

Calculating Grand Average, Average Range and Average Standard Deviation

To calculate the grand average, first find the average of the **n** readings at each time point. The grand average is the average of the averages at each time point.

To calculate the grand range, first determine the range of the **n** readings at each time point. The grand range is the average of the ranges at each time point.

To calculate the average standard deviation, first determine the standard deviation of the **n** readings at each time point. The average standard deviation is the average of the standard deviations at each time point.

Note: You will need to calculate either the grand range or the average standard deviation, not both.

For X-bar charts, the UCL and LCL may be determined as follows:

$$\text{Upper Control Limit (UCL)} = X_{GA} + A_2 R_A$$

$$\text{Lower Control Limit (LCL)} = X_{GA} - A_2 R_A$$

Alternatively, S_A *can be used as well to calculate UCL and LCL:*

$$\text{Upper Control Limit (UCL)} = X_{GA} + A_3 S_A$$

$$\text{Lower Control Limit (LCL)} = X_{GA} - A_3 S_A$$

The centerline is simply X_{GA} .

For R-charts, the UCL and LCL may be determined as follows:

$$\mathrm{UCL} = D_4 R_A$$

$$\mathrm{LCL} = D_3 R_A$$

The centerline is the value R_A .

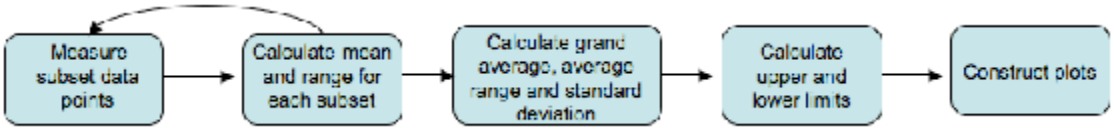
For S-charts, the UCL and LCL may be determined as follows:

$$\mathrm{UCL} = B_4 S_A$$

$$\mathrm{LCL} = B_3 S_A$$

The centerline is S_A .

The following flow chart demonstrates the general method for constructing an X-bar chart, R-chart, or S-chart:



Calculating Region Boundaries

To determine if your system is out of control, you will need to section your data into regions A, B, and C, below and above the grand average. These regions are shown in Figure III. To calculate the boundaries between these regions, you must first calculate the UCL and LCL. The boundaries are evenly spaced between the UCL and LCL. One way to calculate the boundaries is shown below.

Boundary between A and B above $X_{GA} = X_{GA} + (UCL - X_{GA}) * 2 / 3$

Boundary Between B and C above $X_{GA} = X_{GA} + (UCL - X_{GA}) * 1 / 3$

Boundary Between A and B below $X_{GA} = LCL + (X_{GA} - LCL) * 2 / 3$

Boundary Between B and C below $X_{GA} = LCL + (X_{GA} - LCL) * 2 / 3$

Example 1

Assume that in the manufacture of 1 kg Mischmetal ingots, the product weight varies with the batch. Below are a number of subsets taken at normal operating conditions (subsets 1-7), with the weight values given in kg. Construct the X-Bar, R-charts, and S-charts for the experimental data (subsets 8-11). Measurements are taken sequentially in increasing subset number.

Subset #	Values (kg)
1 (control)	1.02, 1.03, 0.98, 0.99
2 (control)	0.96, 1.01, 1.02, 1.01
3 (control)	0.99, 1.02, 1.03, 0.98
4 (control)	0.96, 0.97, 1.02, 0.98
5 (control)	1.03, 1.04, 0.95, 1.00
6 (control)	0.99, 0.99, 1.00, 0.97
7 (control)	1.02, 0.98, 1.01, 1.02
8 (experimental)	1.02, 0.99, 1.01, 0.99
9 (experimental)	1.01, 0.99, 0.97, 1.03
10 (experimental)	1.02, 0.98, 0.99, 1.00
11 (experimental)	0.98, 0.97, 1.02, 1.03

Solution:

First, the average, range, and standard deviation are calculated for each subset.

Subset #	Values (kg)	Average ($\bar{X}$)	Range (R)	Standard Deviation (s)
1 (control)	1.02, 1.03, 0.98, 0.99	1.0050	0.05	0.023805
2 (control)	0.96, 1.01, 1.02, 1.01	1.0000	0.06	0.027080
3 (control)	0.99, 1.02, 1.03, 0.98	1.0050	0.05	0.023805
4 (control)	0.96, 0.97, 1.02, 0.98	0.9825	0.06	0.026300
5 (control)	1.03, 1.04, 0.95, 1.00	1.0150	0.09	0.040509
6 (control)	0.99, 0.99, 1.00, 0.97	0.9875	0.03	0.012583
7 (control)	1.02, 0.98, 1.01, 1.02	1.0075	0.04	0.018930
8 (experimental)	1.02, 0.99, 1.01, 0.99	1.0025	0.03	0.015000
9 (experimental)	1.01, 0.99, 0.97, 1.03	1.0000	0.06	0.026820
10 (experimental)	1.02, 0.98, 0.99, 1.00	0.9975	0.04	0.017078
11 (experimental)	0.98, 0.97, 1.02, 1.03	1.0000	0.06	0.029439

Next, the grand average X_{GA}, average range R_A, and average standard deviation S_A are computed for the subsets taken under normal operating conditions, and thus the centerlines are known. Here n=4.

$$X_{GA} = 1.0004$$

$$R_A = 0.05428$$

$$S_A = 0.023948$$

X-Bar limits are computed (using R_A).

$$\text{UCL} = X_{GA} + A_2 R_A = 1.0004 + 0.729(0.05428) = 1.04$$

$$\text{LCL} = X_{GA} - A_2 R_A = 1.0004 - 0.729(0.05428) = 0.96$$

X-Bar limits are computed (using S_A).

$$\text{UCL} = X_{GA} + A_3 S_A = 1.0004 + 1.628(0.023948) = 1.04$$

$$\text{LCL} = X_{GA} - A_3 S_A = 1.0004 - 1.628(0.023948) = 0.96$$

Note: Since n=4 (a relatively small subset size), both R_A and S_A can be used to accurately calculate the UCL and LCL.

R-chart limits are computed.

$$\text{UCL} = D_4 R_A = 2.282(0.05428) = 0.12$$

$$\text{LCL} = D_3 R_A = 0(0.05428) = 0$$

S-chart limits are computed.

$$\text{UCL} = B_4 S_A = 2.266(0.023948) = 0.054266$$

$$\text{LCL} = B_3 S_A = 0(0.023948) = 0$$

The individual points in subsets 8-11 are plotted below to demonstrate how they vary with in comparison with the control limits.

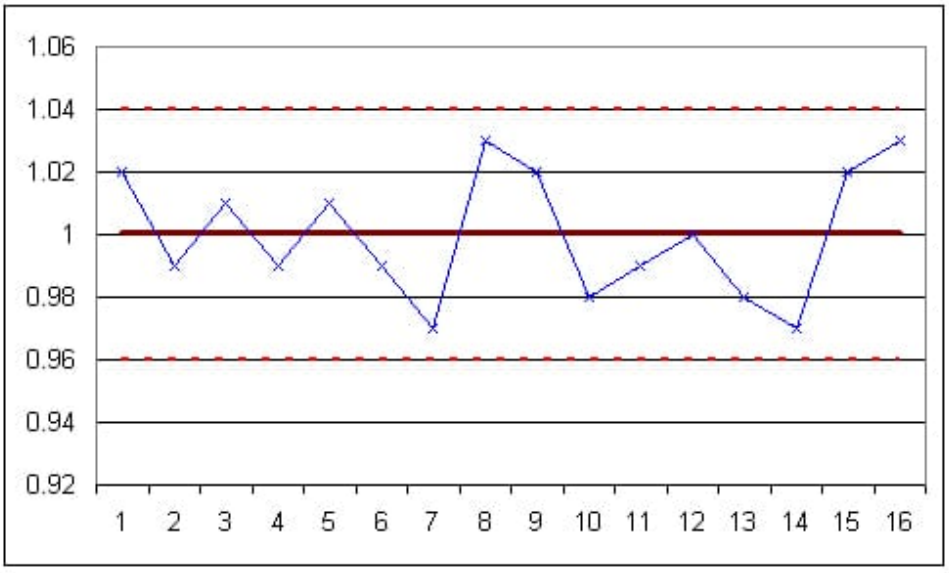


Figure E-1: Chart of individual points in subsets 8-11.

The subgroup averages are shown in the following X-Bar chart:

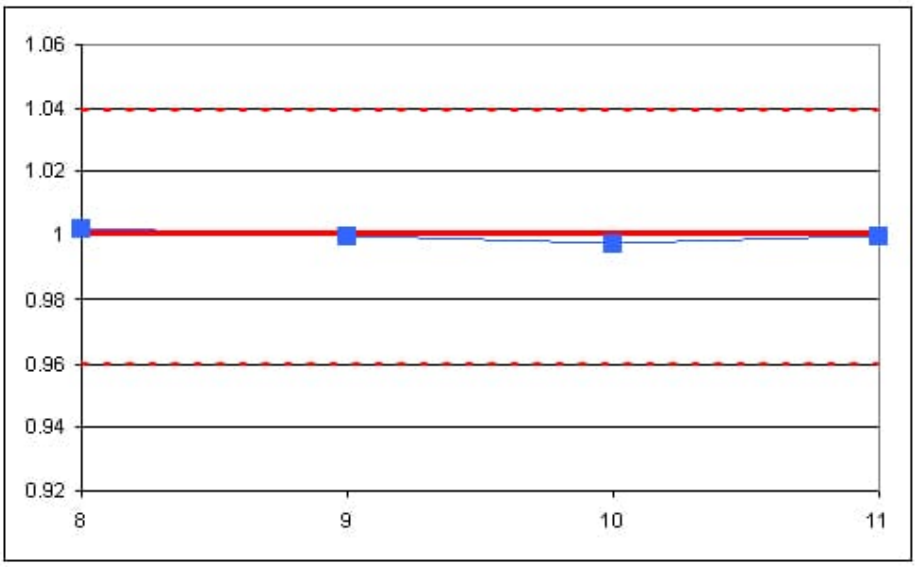


Figure E-2: X-Bar chart for subsets 8-11.

The R-chart is shown below:

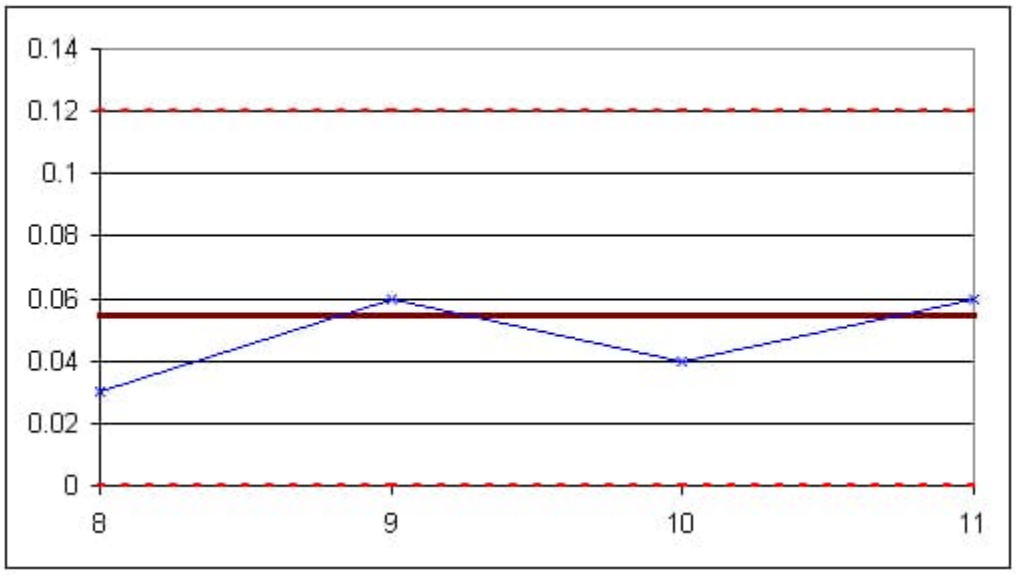


Figure E-3: R-chart for subsets 8-11.

The S-chart is shown below:

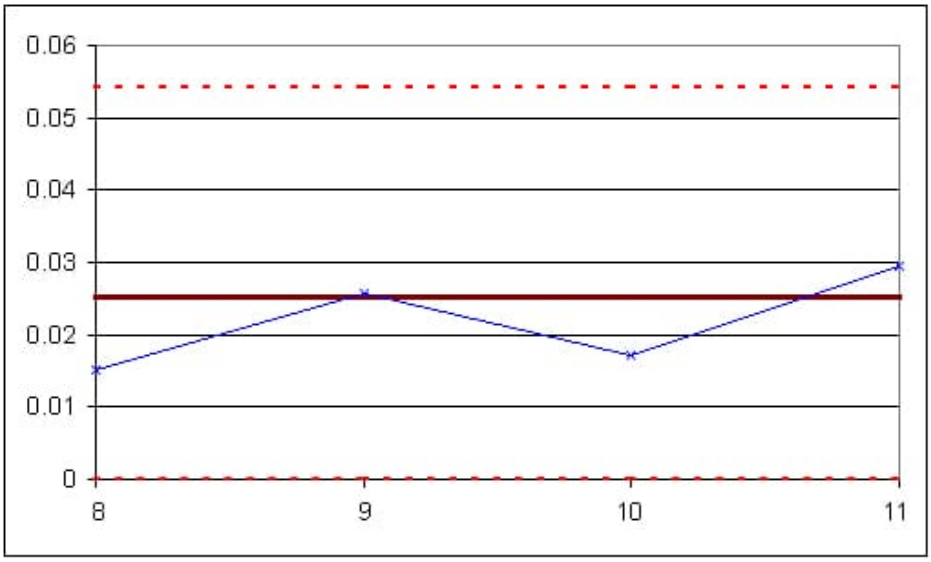


Figure E-4: S-chart for subsets 8-11.

The experimental data is shown to be in control, as it obeys all of the rules given above.

Example 2

It's your first day on the job as a chemical engineer in a plant, and one of your responsibilities is to monitor the pH of a particular process. You are asked by your boss to monitor the stability of the system. She gives you some baseline data for the process, and you collect data for the process during your first day. Construct X-bar and R-Charts to report your results.

Table 1: Baseline data

time (hours)	pH			
1	7.00	7.30	6.99	7.00
2	7.12	7.25	7.12	7.20
3	7.20	7.16	7.20	7.16
4	6.98	7.00	6.94	7.00
5	6.99	6.99	6.99	6.98
6	7.00	6.93	7.02	6.93
7	6.92	7.00	6.92	7.02
8	6.88	6.82	6.94	6.99
9	7.10	7.00	7.00	7.00
10	7.21	7.02	7.21	7.04
11	7.01	6.86	7.01	6.90
12	6.86	6.98	6.90	6.98
13	6.90	7.00	6.87	7.00
14	7.01	7.04	7.01	7.05
15	7.00	6.95	7.00	6.99
16	7.09	7.20	7.03	7.20
17	6.89	7.14	6.87	7.15
18	6.98	6.80	6.98	6.89
19	7.00	6.90	7.00	6.90
20	7.20	7.00	7.23	7.00
21	7.04	7.03	7.08	7.00
22	6.90	6.92	6.98	6.92
23	7.00	7.00	7.00	7.00
24	7.00	6.97	7.01	6.98

To be consistent with the baseline data, each hour you take four pH readings. The data you collect is displayed below.

Table 2: Experimental data

time (hours)	pH	pH	pH	pH
1	6.99	6.99	7.00	6.89
2	6.98	7.12	7.05	6.96
3	7.00	7.18	7.08	7.04
4	7.01	6.94	6.98	7.00
5	6.90	6.99	6.93	7.01
6	6.96	7.01	7.00	7.14
7	7.04	6.92	6.82	7.01
8	7.00	6.93	7.00	6.90
9	7.01	7.00	7.02	6.92
10	7.04	7.18	6.99	6.93
11	6.91	7.01	6.90	7.00
12	7.00	6.97	6.98	7.18
13	7.00	6.89	7.00	7.03
14	7.03	7.01	7.05	6.87
15	6.97	7.00	7.00	6.98
16	7.03	6.97	7.02	6.98
17	6.99	6.89	6.87	6.99
18	6.89	6.98	6.98	6.98
19	6.98	7.00	7.00	7.02
20	7.02	7.15	6.97	6.98
21	7.02	7.08	7.08	7.00
22	6.97	7.01	6.98	7.05
23	7.01	7.04	6.99	7.08
24	6.97	7.00	6.98	6.98

Solution

For this situation, there are k=24 subsets because there are 24 data sets. For each subset, n=4 because there are four pH measurements taken each hour. The first thing you do is calculate the mean and range of each subset. The means are calculated using the AVERAGE() Excel function and the ranges are calculated using MAX() – MIN(). Once these values are calculated, the Grand Average X_{GA} and average range R_A are calculated. These values are simply the means of each subset's mean and range. This data is displayed below.

Table 3: Data used to calculate and grand average and Range.

Control Data						
time (hours)	pH				x_ave	range
1	7.00	7.30	6.99	7.00	7.07	0.31
2	7.12	7.25	7.12	7.20	7.17	0.13
3	7.20	7.16	7.20	7.16	7.18	0.04
4	6.98	7.00	6.94	7.00	6.98	0.06
5	6.99	6.99	6.99	6.98	6.99	0.01
6	7.00	6.93	7.02	6.93	6.97	0.09
7	6.92	7.00	6.92	7.02	6.97	0.10
8	6.88	6.82	6.94	6.99	6.91	0.17
9	7.10	7.00	7.00	7.00	7.03	0.10
10	7.21	7.02	7.21	7.04	7.12	0.19
11	7.01	6.86	7.01	6.90	6.95	0.15
12	6.86	6.98	6.90	6.98	6.93	0.12
13	6.90	7.00	6.87	7.00	6.94	0.13
14	7.01	7.04	7.01	7.05	7.03	0.04
15	7.00	6.95	7.00	6.99	6.99	0.05
16	7.09	7.20	7.03	7.20	7.13	0.17
17	6.89	7.14	6.87	7.15	7.01	0.28
18	6.98	6.80	6.98	6.89	6.91	0.18
19	7.00	6.90	7.00	6.90	6.95	0.10
20	7.20	7.00	7.23	7.00	7.11	0.23
21	7.04	7.03	7.08	7.00	7.04	0.08
22	6.90	6.92	6.98	6.92	6.93	0.08
23	7.00	7.00	7.00	7.00	7.00	0.00
24	7.00	6.97	7.01	6.98	6.99	0.04
				Means:	7.01	0.12

Now that you know X_{GA} = 7.01 and R_A = 0.12, you can calculate the upper control limit, UCL, and lower control limit, LCL, for the X-bar control chart.

From Table A, A_2 = 0.729 when n=4. Using equations UCL and LCL for X-bar charts listed above:

$$\mathrm{UCL} = 7.01 + 0.729(0.12) = 7.0982$$

$$\text{LCL} = 7.01 - 0.729(0.12) = 6.9251$$

Then the UCL = 7.0982, LCL = 6.9251 and X_{GA} = 7.01 are plotted in Excel along with the average values of each subset from the experimental data to produce the X-bar control chart.

Experimental Data						
time (hours)	ph				x ave	range
1	6.99	6.99	7.00	6.89	6.97	0.11
2	6.98	7.12	7.05	6.96	7.03	0.16
3	7.00	7.18	7.08	7.04	7.08	0.18
4	7.01	6.94	6.98	7.00	6.98	0.07
5	6.90	6.99	6.93	7.01	6.96	0.11
6	6.96	7.01	7.00	7.14	7.03	0.18
7	7.04	6.92	6.82	7.01	6.95	0.22
8	7.00	6.93	7.00	6.90	6.96	0.10
9	7.01	7.00	7.02	6.92	6.99	0.10
10	7.04	7.18	6.99	6.93	7.04	0.25
11	6.91	7.01	6.90	7.00	6.96	0.11
12	7.00	6.97	6.98	7.18	7.03	0.21
13	7.00	6.89	7.00	7.03	6.98	0.14
14	7.03	7.01	7.05	6.87	6.99	0.18
15	6.97	7.00	7.00	6.98	6.99	0.03
16	7.03	6.97	7.02	6.98	7.00	0.06
17	6.99	6.89	6.87	6.99	6.94	0.12
18	6.89	6.98	6.98	6.98	6.96	0.09
19	6.98	7.00	7.00	7.02	7.00	0.04
20	7.02	7.15	6.97	6.98	7.03	0.18
21	7.02	7.08	7.08	7.00	7.05	0.08
22	6.97	7.01	6.98	7.05	7.00	0.08
23	7.01	7.04	6.99	7.08	7.03	0.09
24	6.97	7.00	6.98	6.98	6.98	0.03

Table 4: Average subset values and ranges plotted on the X-bar and R-chart

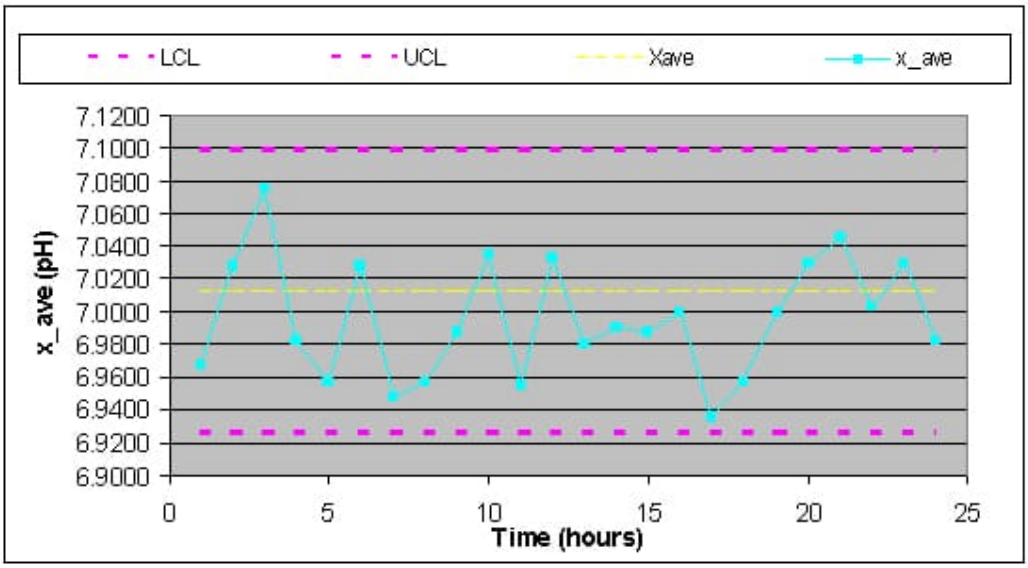


Figure E-5: X-bar control chart

Then, to construct the Range charts, the upper and lower control limits were found. For n=4, D_3 = 0 and D_4 = 2.282 so then:

$$\text{LCL} = D_3 R_A = 0(0.12) = 0$$

$$\text{UCL} = D_4 R_A = 2.282(0.12) = 0.2710$$

Figure E-6).

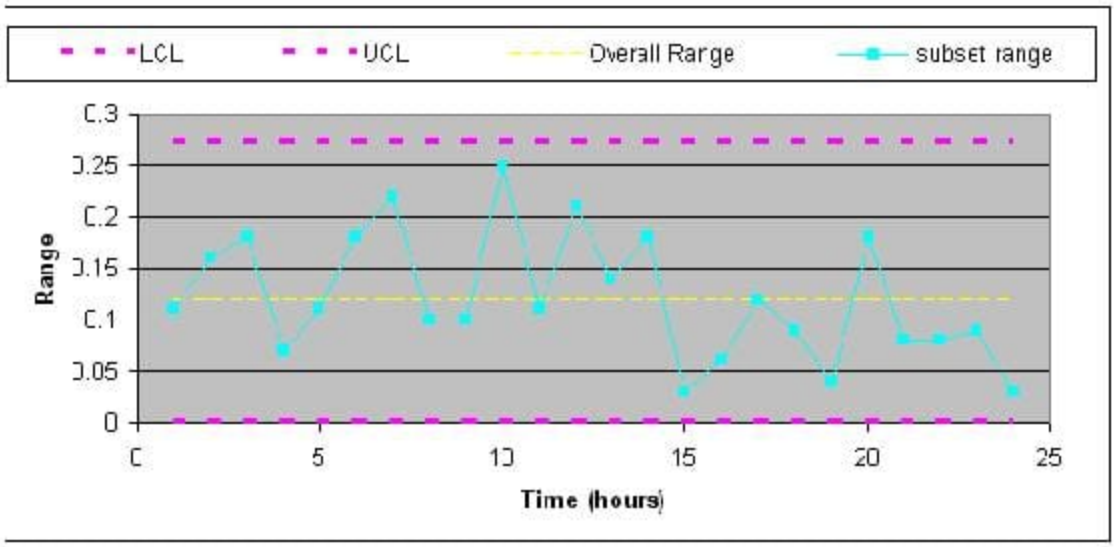


Figure E-6: Range control chart

From both of these charts, the process is in control because all rules for stabilty are met.

Rule 1: No point falls beyond the UCl and LCL.

Rule 2: Two out of three consecutive points do not fall beyond 2σ on the same side of the centerline.

Rule 3: Four out of five consecutive points do not fall beyond 1σ on the same side of the centerline.

Rule 4: Nine or more consecutive points do not fall on the same side of the centerline.

It's important that both of these charts be used for a given set of data because it is possible that a point could be beyond the control band in the Range chart while nothing is out of control on the X-bar chart.

Another issue worth noting is that if the control charts for this pH data did show some points beyond the LCL or UCL, this does not necessarily mean that the process itself is out of control. It probably just means that the pH sensor needs to be recalibrated.

Example 3

A simple out-of-control example with a sample constructed control chart.

You have been analyzing the odd operation of a temperature sensor in one of the plant's CSTR reactors. This particular CSTR's temperature sensor consists of three small thermocouples spaced around the reactor: T1, T2, and T3. The CSTR is jacketed and cooled with industrial water. The reaction taking place in the reactor is moderately exothermic. You know the thermocouples are working fine; you just tested them, but a technician suggests the CSTR has been operating out of control for the last 10 days. There have been daily samples taken and there is a control chart created from the CSTR's grand average and standard deviation from the year's operation.

You are assigned to see if the CSTR is operating out of control. The grand average is 307.47 units of temperature and the grand standard deviation is 4.67 units of temperature. The data is provided for construction of the control chart in Table 1 and the data from the last 10 troublesome days is shown in Table 2. You decide to plot the troublesome data onto the control chart to see if it violates any stability rules.

	Value	Calculate
Upper Control Limit	316.60	$\mu + R_3\sigma$
Upper Zone A Threshold	313.56	$\mu + (2/3)*R_3\sigma$
Upper Zone B Threshold	310.51	$\mu + (1/3)*R_3\sigma$
Grand Average Value	307.47	μ
Lower Zone B Threshold	304.42	$\mu - (1/3)*R_3\sigma$
Lower Zone A Threshold	301.38	$\mu - (2/3)*R_3\sigma$
Lower Control Limit	298.33	$\mu - R_3\sigma$

μ	307.47
σ	4.67
R_3	1.954

Table 3-1. Data for Construction of Control Chart

The way I found A_3 or in this case, R_3, I used the control charts constants table which is found on this wiki page. I decided to use the x-bar using the standard deviations) but you can also use use the range). I found that the value for n(number of subgroups) is three since the CSTR's temperature sensor consists of **three small thermocouples** (T1,T2,T3). Therefore by looking at the constant chart, I get A_3(or R_3 in this case) to be 1.954. Here's the table below:

Subgroup	X-bar chart	
	Using Ra	Using Sa
n	**A2**	**A3**
2	1.886	2.659
3	1.023	1.954
4	0.729	1.628
5	0.577	1.427
6	0.483	1.287

Also, you will notice if you used the range instead of the standard deviation to determine the UCL,LCL, etc. that the values will be roughly the same. Here's the table in comparing the values of UCL and LCL using either A_2 (range) or A_3(stdev):

	A3 (stdev)	**A2 (range)**
	1.954	1.023
	Using Sdev	Using Range
UCL	316.5997379	316.473784
LCL	298.3347317	298.460686

Note: These values were using the same grand average (307.47), the grand standard deviation (4.67) and the grand range (8.80)

Sample Day	T1	T2	T3
1	299.82	310.26	306.60
2	311.68	307.04	300.90
3	310.94	311.68	306.83
4	325.00	308.82	304.97
5	321.43	300.98	311.23
6	320.74	308.97	305.26
7	314.77	313.25	303.36
8	332.75	302.76	306.11
9	319.87	296.81	305.95
10	315.21	314.99	309.60

Table 3-2. Sample Data from Past 10 Troublesome Days

Solution

When the sample data was graphed onto the control chart, the image below was seen.

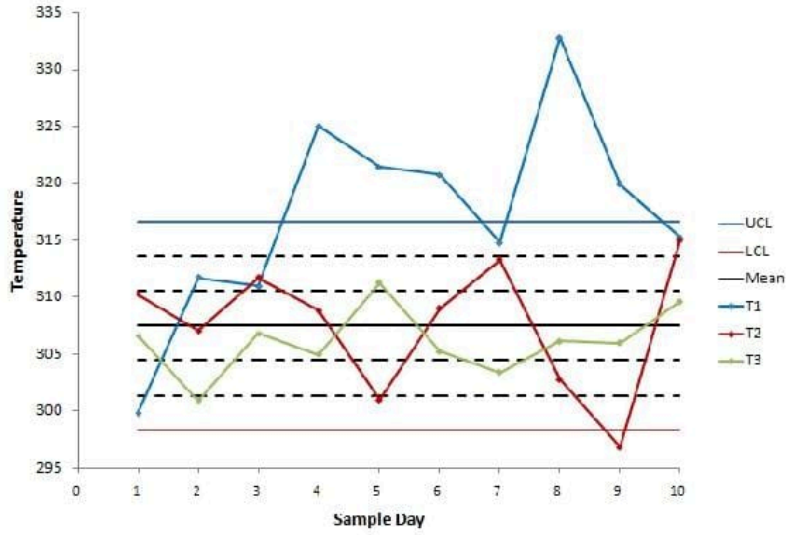


Figure 3-1. 10-Day Data Graphed Onto Control Chart

We can see from the control chart that the CSTR system is clearly out of control. Each thermocouple was tested to see which stability rules it violates.

The first thermocouple (**T1**) violates every stability rule.

- *Rule 1* - Several points from the T1 data fall above the upper control line.
- *Rule 2* - There are many instances where at least two out of three consecutive points fall above the zone AB threshold.
- *Rule 3* - There are eight consecutive points falling above the BC threshold.
- *Rule 4* - Nine consecutive points fall above the mean value.

Judging on this thermocouple's performance, we can say that the system is out of control, but we will analyze the other thermocouples' performance for good measure.

The second thermocouple (**T2**) violates stability rule 1, 2, and 3.

- *Rule 1* - One point falls below the lower control line.
- *Rule 2* - Two consecutive points (samples 9 and 10) fall beyond the AB threshold.
- *Rule 3* - Of the last five samples from T2, four are beyond the BC threshold.

The third thermocouple (**T3**) does not violate any stability rules and the results it displays are within control.

This system is **out of control** because the data from the thermocouples falls beyond the threshold rules for the unit's control chart. This could be explained with many potential situations. One is explained below.

If the CSTR's agitator is knocked loose, the agitation could become erratic. The erratic agitation could create eddy currents and hot spots in the CSTR.

The entire system is out of control because you know that the thermocouples are operating fine and more than one thermocouple violates the stability rules.

Multiple Choice Question 1

When is it useful to use a moving average data set?

A. Never

B. When instant response to process is desired

C. When you have a process that changes slowly

D. Always, as it is a far superior method

Multiple Choice Question 2

What does n signify?

A. The number of subsets in collected data

B. The number of data points in a subset

C. The upper control limit for an X-bar chart

D. The centerline for an R-chart

Multiple Choice Question 3

Of the four stability rules for reading control charts, how many must occur in order for a subset to be determined unstable?

A. 1

B. 2

C. 3

D. 4

Multiple Choice Answers

Question 1: C

Question 2: B

Question 3: A

Sage's Corner

SPC Control Presentation video.google.com/googleplayer...03583755378767
Slides without narration SPC control

References

- Wheeler, Donald J., and David S. Chambers. Understanding Statistical Process Control. 2nd ed. Knoxville: SPC P. 37-88. [1]
- Box, George E., William G. Hunter, and J S. Hunter. Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building. New York: John Wiley & Sons. 43-45. [2]
- Liptak, Bela G. "Process Control and Optimization." Instrument Engineers' Handbook 4: 405-413. [3]
- Woolf, Peter, Amy Keating, Christopher Burge, and Michael Yaffe. Statistics and Probability Primer for Computational Biologists. Massachusetts Institute of Technology. 2004. [4]
- Anderson, David R., Sweeney, Dennis J. and Williams, Thomas A. Statistics for Business and Economics, 10e. Thompson South-Western. 2008. [5]

This page titled 13.2: SPC- Basic Control Charts- Theory and Construction, Sample Size, X-Bar, R charts, S charts is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

13.3: Six Sigma- What is it and what does it mean?

Introduction

Every generation of business strives for a new level of quality. The quality program that is currently in vogue and being widely used and recognized by industry is the Six Sigma program. Six Sigma is a relatively new program, and was only started in 1986. It was first put into implementation at Motorola, but is now in use by most large corporations. Some of these other large companies include GE, Honeywell, and Bank of America. The Six Sigma program is in place to eliminate any abnormalities, failures, or defects that occur within a given process. These are problems that already exist. DFSS (Design for six sigma) starts earlier, to develop or redesign the process itself, so fewer wrinkles show up in the first place, thus systematically preventing downstream errors. Six Sigma is also used in developing new processes. The Six Sigma program strives to achieve six standard deviations between the mean and the closest specification limit on a short term study. Studies run to obtain this goal are short term capability studies, which include common cause or random variation, such as operator control, and long term studies, which include random and special types of variation. Both of these studies are evaluated on a Z-scale. The short term data variability which makes up long term variability tends to cause the mean to shift. This shift of the mean is by 1.5 standard deviations. Most companies are looking on a long term scale, because they would rather have a good/safe product in the long run/for a long time rather than for a short amount of time. Using this idea, the goal for the Six Sigma program is to have fewer than 3.4 failures per one million opportunities when the data is evaluated to include the shifted mean from process variability (6 standard deviations - 1.5 standard deviations = 4.5 standard deviations). The 4.5 vs 6 standard deviations is the same goal, but the 4.5 represents data variation in the long run, which is used in most processes. We will be using the shifted mean scenario for the rest of this article when referring to opportunity goals. This leaves very slight room for error on a process and leads to a very high level of quality in the products.

The often-used six sigma symbol.

To reiterate, the term "Six Sigma" comes from the standard deviation and the Gaussian distribution. 3.4 failures per one million opportunities represents 4.5 standard deviations (sigma) away from the median value, either up or down, under a bell curve. This will be discussed further below along with what the Six Sigma program represents, what mathematically this means, and finally what a Gaussian distribution is. After the basics have been covered we will move onto Statistical Process Control. Statistical Process Control is different ways to analyze the data that you have obtained. Finally, we will relate Six Sigma to process control and explain how you can tell if something is in Six Sigma control or not.

The Six Sigma Program

Six Sigma is defined as a measure of quality that organizations strive to achieve. It is a disciplined, data-driven approach used to eliminate defects in any process – from manufacturing to transactional and from product to service. Six Sigma does not refer to a process operating within 6 or 4.5 standard deviations of the desired state. To achieve Six Sigma, a process must not produce more than 3.4 defects per million opportunities, where a defect is defined as anything outside of the customer specification. One million opportunities, and only 3.4 defects? Sounds like an efficient process. Here is an example that will help demonstrate. Let's say a surgeon at the top of his field has a 5.51 sigma ranking. Sounds pretty good, right? That means that if he operates on 1,000,000 patients per year he loses 30 of them. When a process is operating at Six Sigma, it is seen that there are essentially zero defects within a process.

In chemical engineering, many different processes require strict, robust control. For example, a reactor may need to be maintained between a specified temperature range to keep the plant safe. However, there is also an optimum temperature at which the reaction will occur. Engineers will want to keep a tight control around this temperature to make the system more efficient, while also closely watching the variable and keeping track of how close it gets to the control limits. Crossing the control

limits is not the only problem. In order for a system to be in statistical control as stated by the Six Sigma program, the data points must not frequently approach the same control limit. The Six Sigma program helps engineers monitor processes to keep them within their desirable ranges. Many industries employ the Six Sigma system as well as its hierarchy of positions to champion data-driven problem solving.

Originally invented by Motorola, the Six Sigma system contains belt rankings that represent a person's achievements utilizing Six Sigma methodologies and associated roles and responsibilities. These rankings are modeled after the martial arts system, as detailed below.

Yellow Belts (YB)

Yellow belt is the first level in the Six Sigma program. To get yellow belt status one must go through a training course in Six Sigma and pass an online certification. Yellow belt team members do not lead projects on their own, but they are usually a core team member with special knowledge on a project. Yellow belts are responsible for:

- The development of process maps to support Six Sigma projects.
- Running smaller process improvement projects using the Plan, Do, Check, Act (PDCA) method
- Being a subject matter expert on the Six Sigma project

Yellow Belt Projects

- usually limited to identifying small projects the could benefit from improvement

Green Belts (GB)

Green belts are new team members that achieve certification after completing training in a Six Sigma course. To complete green belt training, new members must complete a relevant project that utilizes the **DMAIC model** outlined below. To maintain green belt certification, green belts generally complete DMAIC projects once every twelve months. Green belts are also responsible for:

- Recommending Six Sigma projects
- Participating on Six Sigma project teams
- Leading Six Sigma teams in local improvement projects

Green Belt Projects

Green belts complete six sigma projects while still holding their current job duties. Therefore, a six sigma project undertaken by a green belt could include:

- Statistical analysis to improve reliability on a reactor within the green belt's current responsibility
- Root cause analysis and elimination of quality control problems that chronically affect the green belt's department

Black Belts (BB)

Black belts are intermediate team members that employ in-depth technical expertise to achieve Six Sigma objectives. Furthermore, black belts can become temporary, full-time change agents for a moderate period of time (two-three years) to produce change using Six Sigma methods. Black belts are also responsible for:

- Demonstrating mastery of black belt knowledge
- Demonstrating proficiency at achieving results through the application of the Six Sigma approach
- Consultation for functional area Internal Process Improvement
- Coaching, Mentoring, and Recommending green belts

Black Belt Projects

Black belts complete six sigma projects outside of their previous job duties. Therefore, a six sigma project undertaken by a black belt could include:

- Plant-level process controls improvement at a facility that greatly affects the finished product at a facility

- Assessment of total warranty/quality control issues at a facility through large scale statistical data analysis and improvement

Master Black Belts (MBB)

The major duty of master black belts include acting as project leaders, as well as:

- Training and mentoring Black Belts and Green Belts
- Helping to decide large-scale Six Sigma projects
- Maintaining the quality of Six Sigma measurements
- Developing Six Sigma training

Master Black Belt Projects

Most master black belts would only be involved with directing black and/or green belts in their projects. Their effect on process would generally be limited to pointing out and suggesting large projects for appointed black belts to undertake.

To complete Six Sigma projects, belt candidates utilize the DMAIC model. DMAIC (Define, Measure, Analyze, Improve, Control) is an improvement system for existing processes falling below specification and looking for incremental improvement. DMADV (Define, Measure, Analyze, Design, Verify) is an improvement system used to develop new processes at Six Sigma quality level. Below, the steps for DMAIC and DMADV are outlined.

Table 1 - Outline of the steps used in DMAIC (improve an existing process)

D	**Define** the goals of the improvement activity. The most important goals are obtained from customers. At the top level the goals will be the strategic objectives of the organization, such as greater customer loyalty, increased market share, or greater employee satisfaction. At the operations level, a goal might be to increase the throughput of a production department. At the project level goals might be to reduce the defect level and increase throughput for a particular process. Obtain goals from direct communication with customers, shareholders, and employees.
M	**Measure** the existing system. Establish valid and reliable metrics to help monitor progress towards the goal(s) defined at the previous step. Begin by determining the current baseline. Use exploratory and descriptive data analysis to help you understand the data.
A	**Analyze** the system to identify ways to eliminate the gap between the current performance of the system or process and the desired goal. Use statistical tools to guide the analysis.
I	**Improve** the system. Be creative in finding new ways to do things better, cheaper, or faster. Use project management and other planning and management tools to implement the new approach. Use statistical methods to validate the improvement.
C	**Control** the new system. Institutionalize the improved system by modifying compensation and incentive systems, policies, procedures, budgets, operating instructions and other management systems. You may wish to utilize standardization such as ISO 9000 to assure that documentation is correct. Use statistical tools to monitor stability of the new systems.

Table 2 - Outline of the steps used in DMADV (develop a new process)

D	**Define** the goals of the design activity. What is being designed? Why? Use QFD or Analytic Hierarchical Process to assure that the goals are consistent with customer demands and enterprise strategy.
M	**Measure.** Determine Critical to Stakeholder metrics. Translate customer requirements into project goals.
A	**Analyze** the options available for meeting the goals. Determine the performance of similar best-in-class designs.
D	**Design** the new product, service or process. Use predictive models, simulation, prototypes, pilot runs, etc. to validate the design concept's effectiveness in meeting goals.
V	**Verify** the design's effectiveness in the real-world.

All charts and descriptions referenced from Pyzdek, Thomas, "DMAIC and DMADV Explained" http://www.pyzdek.com/DMAICDMADV.htm , 2003

Statistics and Six Sigma

You are probably already familiar with the concepts of average, standard deviation, and Gaussian distribution. However, they are very important concepts in Six Sigma, so they are reviewed in this section. This section also discusses analysis methods used in Six Sigma.

Average

The equation for calculating an **average** is shown below.

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{x_1 + x_2 + \cdots + x_N}{N}$$

where

- $\bar{x}$ is the average
- xi is the measurement for trial i
- N is the number of measurements

This equation relates to Six Sigma because it is the value that you aim for when you are creating your product. After millions of products made, you will have a very good idea of what your exact average product specification is. The average is combined with the specification limits, which are the limits that determine if your product is in or out of spec. The wider the specification limits are, the more room for deviation from the average there is for your product. A product specification would be written like this:

$$10 \pm 2\,mm$$

Where the first number (10) represents the average and the second number (2) represents the amount of error allowable from the average without violating 4.5 standard deviations (on a long term scale). Thus, your product can range from 8 to 12 mm for this example.

The average is a good representation of the a set of data. However, the main problem it experiences is the how it is strongly influenced by outlier values. For example, say a town had a population of 50 people with an average income of $30,000 per person. Now say that a person moves into the town with an income of $1,000,000. This would move the average up to approximately $50,000 per person. However, this is not a good representation of the average income of the town. Hence, outlier values must be taken into account when analyzing data. In contrast to the mean, sometimes the median can be a good representation of a set of data. The median is defined as the middle value of a set of data are arranged in order. The median is immune to outlier values as it basically one value and is not calculated it any way so it cannot be influenced by a wide range of numbers. Both the mean and median can be taken into account when analyzing a set of data points.

Standard Deviation

The equation for **standard deviation** is shown below.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

where σ = standard deviation, and the other variables are as defined for the average.

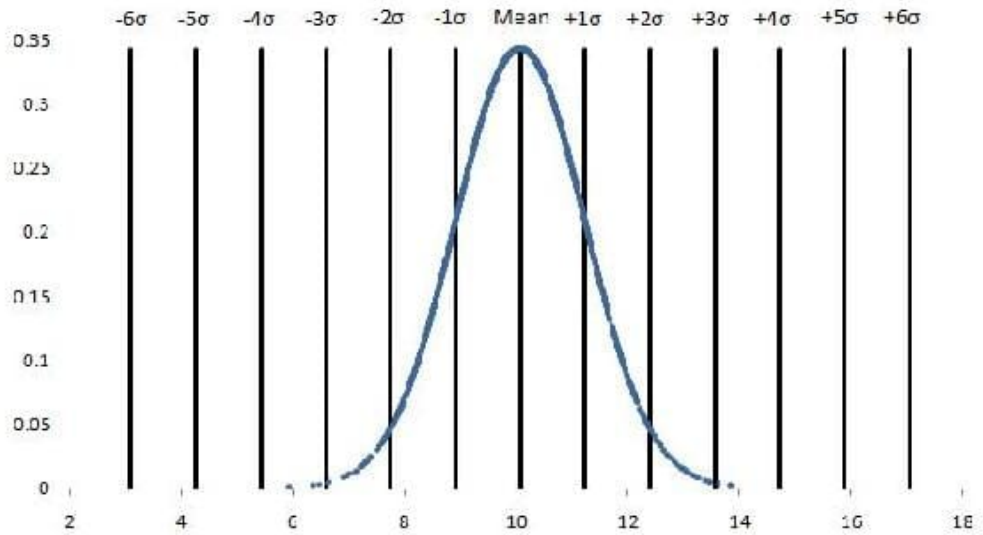
For each measurement, the difference between the measured value and the average is calculated. This difference is called the residual. The sum of the squared residuals is calculated and divided by the number of samples minus 1. Finally, the square root is taken.

The standard deviation is the basis of Six Sigma. The number of standard deviations that can fit within the boundaries set by your process represent Six Sigma. If you can fit 4.5 standard deviations within your process specifications then you have obtained a Six Sigma process for a long term scale. However, the number of errors that you can have for your process as you move out each standard deviation continues to decrease. The table below shows the percentage of data that falls within the standard deviations and the amount of defects per sigma, in terms of "Defects Per Million Opportunities" or DPMO. The percentage of errors that you are allowed is one minus the percentage encompassed by the percent of the total.

Data from: www.isixsigma.com/library/content/c020813a.asp

# of Standard Deviations	% of Total	DPMO
1	68.27	690,000
2	95.45	308,537
3	99.73	66,807
4	99.9937	6,210
5	99.99994	233
6	99.9999998	3.4

The image below shows an example data set (see #Gaussian Distribution below) with lines marking 1 to 6 standard deviations from the mean. In this example, the mean is approximately 10 and the standard deviation is 1.16.



Gaussian Distribution

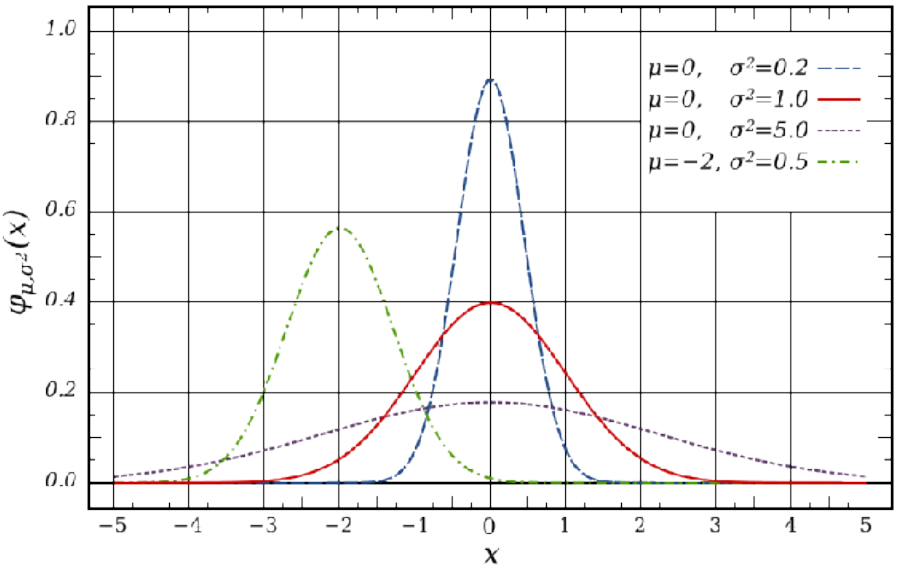
The normal, or Gaussian, distribution is a family of continuous probability distributions. The normal distribution function was first introduced in 1733, and since has become the most widely used family of distribution functions in statistical analysis. These distribution functions are defined by two parameters: a location (most commonly the "mean",μ), and a scale (most commonly the "variance", σ^2). The use of normal distributions in statistics can be seen as advantageous due to its ability to maximize information entropy around a given mean and variance. Information entropy is the measure of uncertainty associated with a random variable.

For the purposes of this class and Chemical Engineering applications, you can model a given set of data using the probability density function which is shown below.

$$\varphi = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

where $\varphi) is the probability density, \backslash(\mu$ is the average and the other variables are as defined in the above two sections.

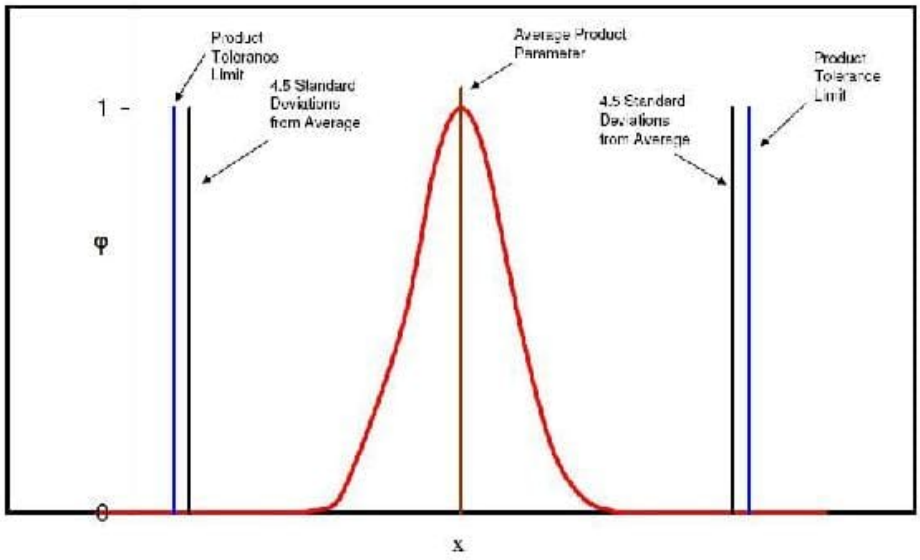
The above PDF equation gives graphs, depending on mean and variance, similar to those below.



Above are four examples of different distributions given different values for mean and standard deviation. An important case is the standard normal distribution shown as the red line. The standard normal distribution is the normal distribution with a mean of 0 and a variance of 1. It is also important to note some special properties of probability density functions:

- symmetry about the mean, μ
- the mode and mean are both equal to the mean
- the inflection points always occur one standard deviation away from the mean, at $\mu - \sigma$ and $\mu + \sigma$

Suppose we have a process where we make a product of a certain concentration and we have good control over the process. After analyzing a set of data from a time period we see that we have a standard deviation of only 0.01 and our product concentration is required to be within 0.05. In order to say our product is essentially defect-free, 4.5 standard deviations away from the average must be less than our required product tolerance (± 0.05). In this case 4.5 standard deviations is equal to 0.045 and our product tolerance is 0.05. This is more easily seen graphically, as seen in the figure below.



As you can see 4.5 standard deviations on either side of the averages falls just slightly inside of the required tolerances for our product. This means that 99.997% of our product will fall inside of our product tolerances.

Analysis Methods

Two ways to analyze data for similarities (percent confidence) would be running a regression analysis or an analysis of variance (ANOVA). The difference between these two methods of analysis is that the ANOVA is for comparing independent groups such as 4 different reactor compositions that are supposed to make the same product (in reality this doesn't happen). A linear regression would be used for data that is dependent on each other, such as looking at the difference between slurry product permeabilities when a one-reactant composition in the process is changed. The percent confidence is the same as the student's t-test, where you would want a t value (error) that is less than 0.05 to have statistically similar data.

One statistical program that is widely used in Six Sigma projects is MINITAB. MINITAB is similar to Microsoft Excel but has much greater statistical analysis capabilities. This program allows you to run linear regressions and ANOVAs by the click of an option. It also graphs not only in 2-D but 3-D as well and runs statistical analysis methods that are more in depth than those offered in Excel. MINITAB graphical and analytical tools can provide insight and information throughout the DMAIC process. MINITAB can be used to:

- **Identify** - visualize sources of variation and relationships among variables
- **Verify** - having statistical confidence behind conclusions

Key Tool Bar Descriptions on MINITAB

- File - typical of most programs - save, open, print, etc.
- Data - can subset data easily (stratification), transpose data, stach and unstack, convert text to numberic and back, code data
- Calc - calculator function - one time calc (doesn't update as you update your data)
- Stat - t-tests, regression, ANOVA, DOE, control charts, quality tools (Pareto and Gage R&R)
- Graph - plots, histogram, matrix plot
- Editor - find/replace (if you need to find items)
- Tools - options and links to some external tools
- Window - session window shows detains on what you have open

MINITAB is a powerful tool to help us understand the language of measurement and the different uses of measurement system. Some unique features of MINITAB include: Measurement System Analysis (MSA) and the Gage Reproducibility and Repeatability Study (Gage R&R). These two complicated Six Sigma analysis can be accomplished within steps using MINITAB.

Statistical Process Control

Statistical process control (SPC) is an effective method of determining whether a system has deviated from its normal statistical behavior by the use of control charts. Walter Shewhart first pioneered the use of statistical techniques when studying variations in product quality and consistency at Bell Laboratories in the early 1920's. SPC utilizes statistical tools in the form of control charts (such as Shewhart charts) to observe the performance of the production process in order to predict significant deviations that may result. Dr. Shewhart created concepts for statistical control and the basis for control charts through a series of carefully designed experiments. While he primarily drew from pure mathematical statistical theories, he understood that data from physical processes seldom produced a normal distribution curve (a Gaussian distribution, also known as a bell curve). Shewhart determined that every process displays variation: some display controlled variation that is natural to the process, while others display uncontrolled variation that is not always present in the process causal system.

SPC, most often used for manufacturing processes, is used in conjunction with experimental designs, process capability analyses, and process improvement plans to monitor product quality and maintain process targets. Six Sigma programs often depend on statistical process controls to provide their data with supportive information. If a process falls outside its preferred operating range, the statistical control can help you determine where and why the process failed.

Benefits of the SPC Method:

- Provides surveillance and feedback for keeping processes in control
- Signals when a problem with the process has occurred
- Detects assignable causes of variation
- Accomplishes process characterization
- Reduces need for inspection
- Monitors process quality
- Provides mechanism to make process changes and track effects of those changes
- Once a process is stable (assignable causes of variation have been eliminated), provides process capability analysis with comparison to the product tolerance

Capabilities of the SPC Method:

- All forms of SPC control charts
- Variable and attribute charts
- Average (X), Range (R), standard deviation (s), Shewhart, CuSum, combined Shewhart-CuSum, exponentially weighted moving average (EWMA)
- Selection of measures for SPC
- Process and machine capability analysis (Cp and Cpk)
- Process characterization
- Variation reduction
- Experimental design
- Quality problem solving
- Cause and effect diagrams

SPC is used to monitor the consistency of processes used to manufacture a product as designed. It aims to get and keep processes under control. No matter how good or bad the design, SPC can ensure that the product is being manufactured as designed and intended. Thus, SPC will not improve a poorly designed product's reliability, but can be used to maintain the consistency of how the product is made and, therefore, of the manufactured product itself and its as-designed reliability.

Methods and Control Charts

Control charts are the basis of statistical process controls methodologies, and are used to distinguish between random/uncontrollable variations and controlled/correctable variations in a system. Control chart detects and monitors process variation over time. It also plays a role as a tool for ongoing control of a process. There are many types of SPC control charts, though centering charts are used most often. The remainder of this article will focus on the different types of centering and dispersion charts. Figure 1 displays single point centering charts and Figure 2 displays subgroup sampling charts. Each figure has a brief description of the chart type within each of the two chart families (centering and dispersion). In the centering chart, the centerline is given by the average of the data samples, where as in the dispersion chart the centerline is the frequency distribution from that average.

Chart Family	Chart Type	Description
Centering Chart	Individual	Uses individual point values to provide a trend.
	Geometric Mean	Uses the geometric mean to provide exponential smoothing.
	Moving Average	Uses the moving average to remove noise and indicate an overall trend.
Dispersion Chart	Range (R chart)	Commonly called an R chart. Plots the range which is the difference between the highest and lowest sample in the range. The number of samples for calculating the range is configured by the user.

Figure 1: Control chart types and descriptions for the single point sampling method(click to enlarge)[1].

Chart Family	Chart Type	Description
Centering Chart	Average (XBar)	Commonly called an XBar chart. Plots the average of data in the specified range. All subgroup samples are saved until the specified size is met. The average is then calculated, and the value is written to disk.
	Median	Plots the median value in the same manner as the average chart plots the average value.
Dispersion Chart	Range (R chart)	Commonly called an R chart. Plots the range which is the difference between the highest and lowest sample in the subgroup. Samples are collected until the subgroup is complete, and then sorted in order of size. The average or median is determined for the centering chart, and then the range is calculated.
	Range with Sigma	Plots the R values in the same manner as in the range chart, but the limits are calculated using a deviation value (sigma) that was entered during configuration.
	Root Mean Squared	Used for subgroups of 12 to 25 points. Uses the root mean square (RMS) to calculate the limits and centerline.
	Root Mean Squared with Sigma	Plot point in the same manner as the RMS chart, but the limits and the centerline are calculated using a deviation value (sigma) that was entered during configuration.
	Standard Deviation	Plots the standard deviation of the subgroup points.

Figure 2: Control chart types and descriptions for the subgroup sampling method(click to enlarge)[1].

A very useful type of centering chart is the Xbar chart, also known as the **Shewhart chart** mentioned earlier in the article. Figure 3 gives an example of an Xbar chart (on top), as well as an example of a dispersion chart (on bottom). The charts are used in combination with a system of rule based checks to determine whether the system is in statistical control. These rules can be seen in Figure 4 for both the single point centering and subgroup sampling methods. Using these rules, one can inspect a given chart over a period of time and determine whether the system is in control. If the system is not in control they must then check the controllable aspects of the system to see where the problem is occurring.

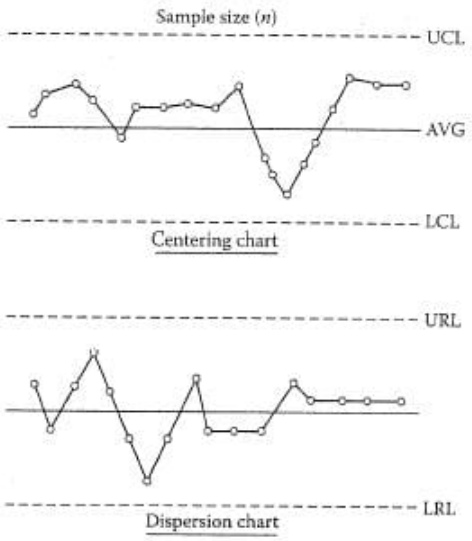


Figure 3: The centering Xbar chart (Shewhart Chart) on top is used to determine whether the centering of the process is stable. The dispersion (R) chart on bottom is used to determine whether the spread of the process is stable (click to enlarge)[1].

Centering Chart Type	Criteria which show that the system is not in statistical control
Single Point Values	1. Any of the points fall outside of the control limits. 2. Seven or more consecutive points fall on the same side of the centerline. 3. Ten of eleven consecutive points fall on the same side of the centerline. 4. Three or more consecutive points fall on the same side and all are located closer to the limit than the centerline.
Subgroup Sampling Method	1. Any single subgroup is more than 3 standard deviations away from the centerline (or set point). 2. Two consecutive subgroup values are more than two standard deviations away from the centerline. 3. Three consecutive subgroup values are more than one standard deviation away from the centerline and on the same side of the centerline. 4. Two out of three consecutive subgroup values are more than two standard deviations away from the centerline, with all on the same side of the centerline. 5. Five consecutive subgroup values are on the same side of the centerline.

Figure 4: Table showing the criteria that indicate if a system is not in statistical control for different types of centering charts (click to enlarge)[1].

Note: We only need to see if any one of the criteria objectives are achieved to claim that the system is statistically out of control.

As can be seen from the figures above, the primary tool of SPC is the control chart. Several different descriptive statistics are used in control charts along with several different types of control charts that test for different causes of failure. Control charts are also used with product measurements to analyze process capability and for continuous process improvement efforts.

Example 13.3.1

KLMM Industries has recent data on their widget making process. The data is shown below. The allowable range of widget width is 19.5-22.5 cm.

Part Number	Width (cm)
1	21.26
2	21.33
3	20.39
4	20.33
5	21.12
6	21.48
7	21.44
8	21.22
9	20.34
10	20.39
11	21.19
12	21.04
13	20.79
14	20.39
15	21.38
16	20.91
17	20.10
18	21.31
19	20.36
20	20.35

a. Please from this data calculate the mean and write the mean with its allowable range?
b. Next calculate the standard deviation for this data.
c. Does this process fall within 4.5 standard deviations?
d. 6 standard deviations?

Solution

a

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{x_1 + x_2 + \cdots + x_N}{N}$$

This is also calculated using the average function in excel, from this the average is calculated to be **20.9945**

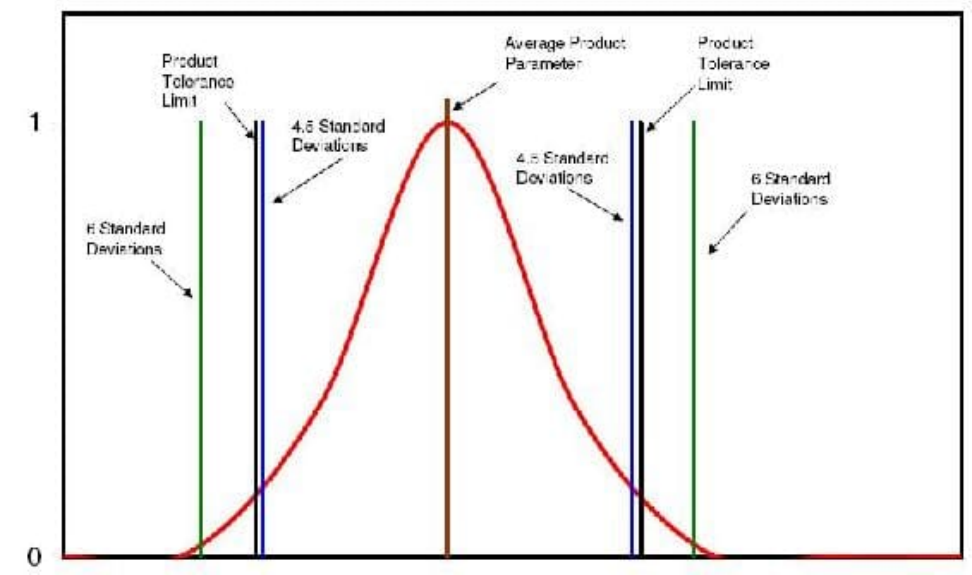
b

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - \bar{x})^2}$$

This is also calculated using the STDEV function in excel, from this the standard deviation is calculated to be **0.3206**

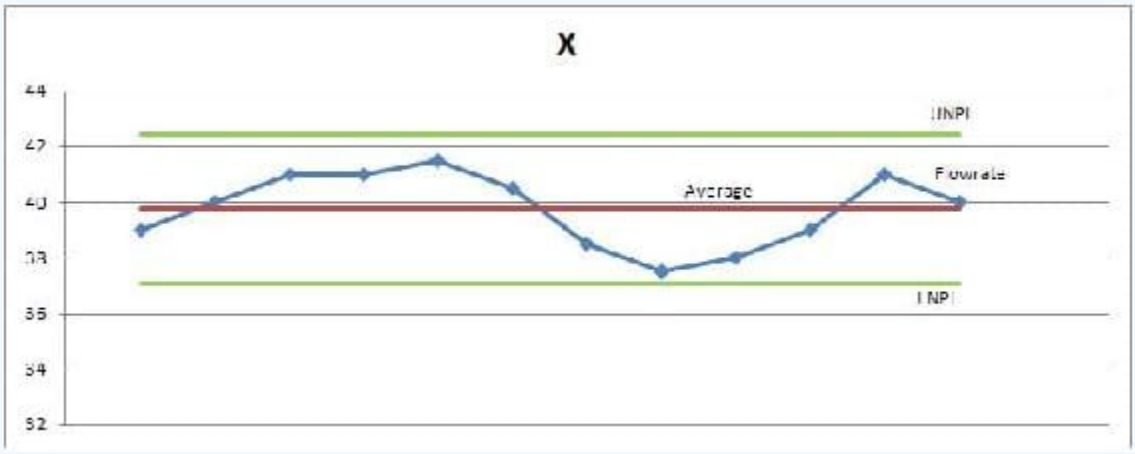
3) Yes

4) No



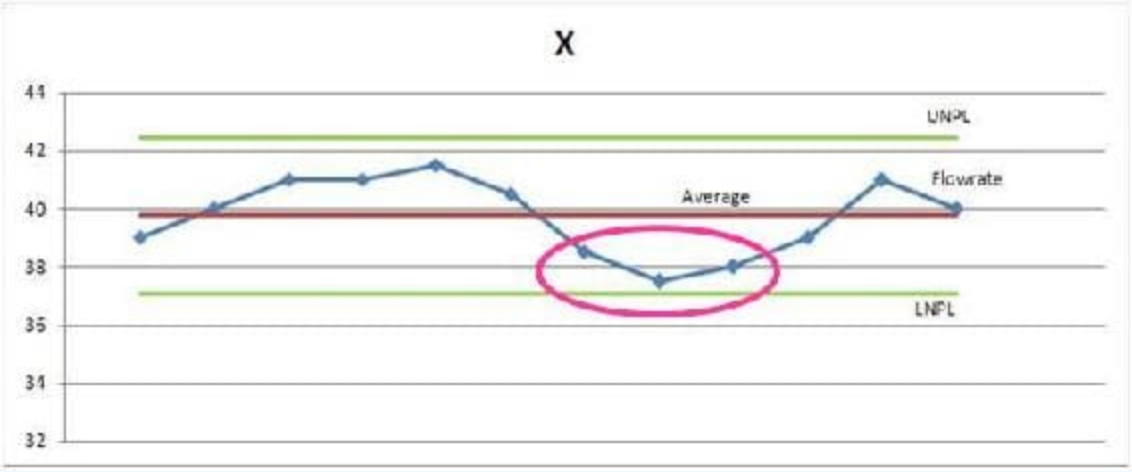
Example 13.3.2

A chemical engineer at Cool Processes Inc. is in charge of monitoring the flowrate of the cooling water for a reaction. He's new at the company and has asked you for your input. He says that the process is currently in statistical control and does not need to be changed. Below is the centering control chart he has made of the process. Do you agree or disagree with his opinion? Explain.



Solution

Even though the process appears to be within a reasonable range of oscillation, and none of the data points are exceeding the process limitations, the system is not in statistical control. Centering (X) charts show that a process is not in statistical control if "three or more consecutive points fall on the same side of the centerline and all are located closer to the control limit than to the centerline."(Liptak 2006). The data points circled in the plot below violate that rule.



Example 13.3.3

This example will also serve as a small tutorial for using Minitab.

ABC Production is coming out with a new product. One of the chemical engineers has hypothesized that the product thickness is dependent on the process temperature and the concentration of "compound Z." There are currently controls regulating the temperature between 77 and 80°C, and the concentration of Z between 0.0066 and .0073 g/L. Product specifications require that the thickness limits are 1.94 to 2.10 cm. You have been assigned the task of determining if these are good parameters to control. Use the data from random samples and use Minitab with linear regression to make your conclusion.

Sample No.	Thickness (cm)	Temperature (°C)	Concentration (g/L)
1	2.1740228	82	0.066
2	1.8774501	77	0.071
3	1.8774704	77	0.072
4	1.9762727	79	0.069
5	2.0266303	80	0.071
6	2.0994529	81	0.066
7	1.9468132	78	0.067
8	1.8972298	77	0.071
9	1.9169798	77	0.07
10	2.0692626	80	0.066
11	2.1292363	82	0.067
12	2.0479427	80	0.067
13	2.0479598	80	0.069
14	1.8972463	77	0.071
15	1.8774795	77	0.066

Solution

Conclusion: Temperature is much more influential on the thickness than the concentration, so it may be more worthwhile to refine the temperature control than to focus on the concentration controllers. Another option would be to investigate other system parameters (i.e., time in reactor, pressure, etc.) to see if they hold more influence on the thickness.

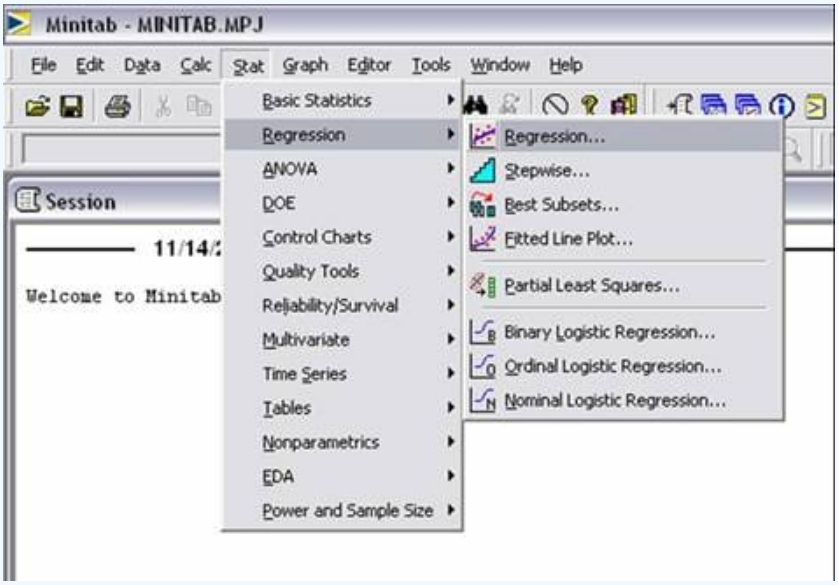
Using Minitab

1) Open Minitab and input data

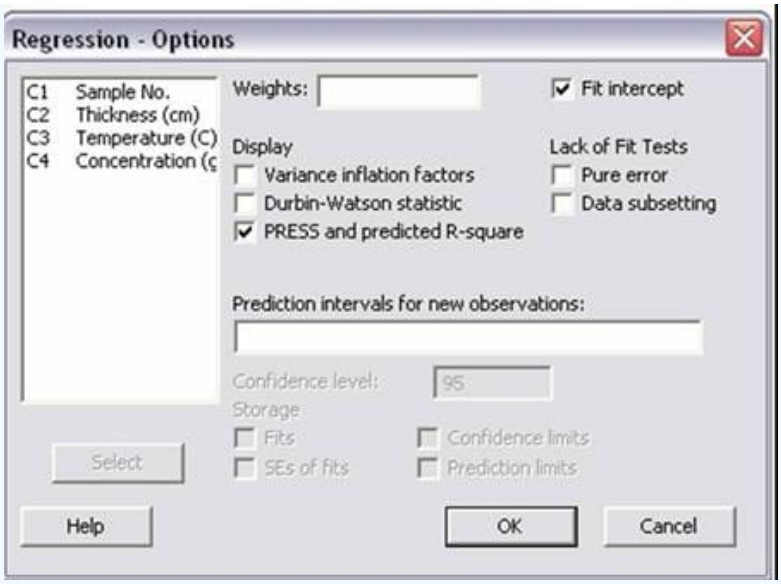
Worksheet 1 ***

	C1	C2	C3	C4	C5
	Sample No.	Thickness (cm)	Temperature (C)	Concentration (g/L)	
1	1	2.17402	82	0.066	
2	2	1.87745	77	0.071	
3	3	1.87747	77	0.072	
4	4	1.97627	79	0.069	
5	5	2.02663	80	0.071	
6	6	2.09945	81	0.066	
7	7	1.94681	78	0.067	
8	8	1.89723	77	0.071	
9	9	1.91698	77	0.070	
10	10	2.06926	80	0.066	
11	11	2.12924	82	0.067	
12	12	2.04794	80	0.067	
13	13	2.04796	80	0.069	
14	14	1.89725	77	0.071	
15	15	1.87748	77	0.066	
16					

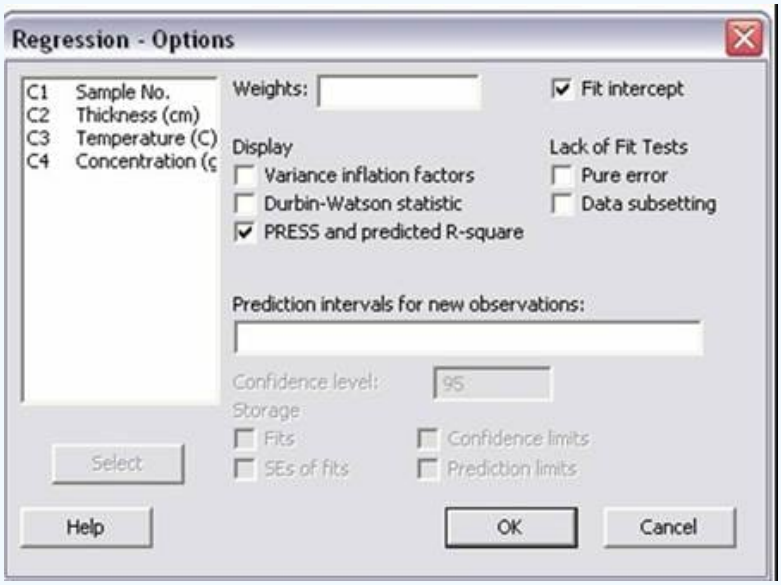
2) Select the “Regression” option from the “Stat” menu and the “Regression” tab



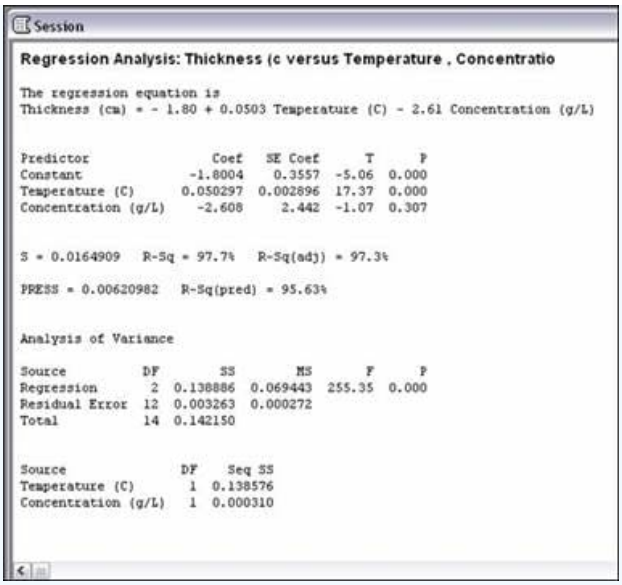
3)The regression window will open; add the “thickness” column for the response, and the Temperature and Concentration columns to the predictors.



4)Click the “Options” button and select “PRESS and predicted R-squared,” then click “OK” on both dialog boxes.



5)The results will appear in the session window:



Session

Regression Analysis: Thickness (c versus Temperature , Concentratio

The regression equation is
Thickness (cm) = - 1.80 + 0.0503 Temperature (C) - 2.61 Concentration (g/L)

Predictor	Coef	SE Coef	T	P
Constant	-1.8004	0.3557	-5.06	0.000
Temperature (C)	0.050297	0.002896	17.37	0.000
Concentration (g/L)	-2.608	2.442	-1.07	0.307

S = 0.0164909 R-Sq = 97.7% R-Sq(adj) = 97.3%

PRESS = 0.00620982 R-Sq(pred) = 95.63%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	0.138886	0.069443	255.35	0.000
Residual Error	12	0.003263	0.000272		
Total	14	0.142150			

Source	DF	Seq SS
Temperature (C)	1	0.138576
Concentration (g/L)	1	0.000310

Regression Analysis: Thickness (c versus Temperature , Concentratio

The regression equation is
Thickness (cm) = - 1.80 + 0.0503 Temperature (C) - 2.61 Concentration (g/L)

Predictor	Coef	SE Coef	T	P
Constant	-1.8004	0.3557	-5.06	0.000
Temperature (C)	0.050297	0.002896	17.37	0.000
Concentration (g/L)	-2.608	2.442	-1.07	0.307

S = 0.0164909 R-Sq = 97.7% R-Sq(adj) = 97.3%

PRESS = 0.00620982 R-Sq(pred) = 95.63%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	0.138886	0.069443	255.35	0.000
Residual Error	12	0.003263	0.000272		
Total	14	0.142150			

Source	DF	Seq SS
Temperature (C)	1	0.138576
Concentration (g/L)	1	0.000310

Interpreting the data

Some of the most important values in the returned data are the p values. These values test the null hypothesis that there is no effect by the inputs (concentration and temperature here) on the output (thickness). The lower the value of p, the more correlation there is between each variable and the result. In this case, the regression has a p value of 0 (under analysis of variance), so the model is significant and at least one of the coefficients is not zero. The p values in the predictor section show a zero value for Temperature (high correlation with thickness) but a 0.3 value for concentration (low correlation with thickness. This indicates that the concentration controller is not significant.

Next we will in spect the R2 values. The R-Sq value of 97.7% indicates that the predictors (concentration and temperature) explain that percent of the variance in thickness. The adjusted value (97.3%) accounts for the number of predictors (here, 2) in the model. Since both are high, the model that is given at the beginning of the readout fits the data well.

The predicted R value is also important because it measures the amount of overfit of the model. If this value were not close to the R-Sq value, then the model would only be applicable for this set of data points but would probably not be accurate for any other combination of data points. Since the “R-Sq(pred)” value is close to the R-Sq value, this model is good.

References

1. *Instrument Engineer*'s *Handbook:Process Control and Optimization* Vol.IV (2005) edited by Liptak, Bela G. New York: Taylor and Francis Group
2. Perlmutter, Barry (2005) *A Six Sigma Approach to Evaluating Vacuum Filtration Technologies* BHS-Filtration INC.
3. Chambers & Wheeler *Understanding Statistical Process Control* Tennessee: SPC Press

This page titled 13.3: Six Sigma- What is it and what does it mean? is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

13.4: Bayes Rule, Conditional Probability and Independence

Introduction

Probability is the likely percentage of times an event is *expected* to occur if the experiment is repeated for a large number of trials. The probability of rare event is close to zero percent and that of common event is close to 100%. Contrary to popular belief, it is not intended to accurately describe a single event, although people may often use it as such. For example, we all know that the probability of seeing the head side of a coin, if you were to randomly flip it, is 50%. However, many people misinterpret this as 1 in 2 times, 2 in 4 times, 5 in 10 times, etc. of seeing the head side of the coin appear. So if you were to flip a coin 4 times and it came up heads every single time, is this 50% probability incorrect? No! It's just that your sample size is small. If you were to flip a coin 10,000 times, you would begin to see a more even distribution of heads and tails. (And if you don't, you should probably get a different coin.) Now, as engineers, even if we know the probability of the system, we don't have time to perform 10,000 trials to verify it. However, you'd be surprised at the small number of trials that are required to get an accurate representation of the system. The following sections describe the relationship between events and their probabilities. These relationships will then be used to describe another probability theory known as Bayes' Theorem.

Types of Probability

Combination

Combinatorics is the study of all the possible orderings of a finite number of objects into distinct groups. If we use combinatorics to study the possible combinations made from ordering the letters A, B, and C we can begin by counting out all the orderings

$$\{(ABC),(ACB),(BCA),(BAC),(CBA),(CAB)\}$$

giving us a total of 6 possible combinations of 3 distinct objects. As you can imagine the counting method is simple when the number of objects is small yet, when the number of objects being analyzed increases the method of counting by hand becomes increasingly tedious. The way to do this mathematically is using factorials. If you have **n** distinct objects then you can order them into $n!$ groups. Breaking the factorial down for our first example we can say, first there are 3 objects to choose from, then 2, then 1 no matter which object we choose first. Multiplying the numbers together we get 3*2*1=3!. Now consider finding all the possible orderings using all the letters of the alphabet. Knowing there are 26 letters in the English alphabet, the number of possible outcomes is simply 26!, a number so large that counting would be difficult.

Now what if there are **n** objects and **m** that are equivalent and you wish to know the number of possible outcomes. By example imagine finding the number of distinct combinations from rearranging the letters of **PEPPER**. There are **6** letters, **2 Es** and **3 Ps** but only **1 R**. Starting with **6!** we need to divide by the repeat possible outcomes

$$\frac{6!}{3!\times 2!}=\frac{6\times 5\times 4\times 3!}{3!\times 2!}=\frac{6\times 5\times 4}{2}=6\times 5\times 2=60\,\text{possible arrangements}$$

where on the bottom, the $3!$ is for the repeated Ps and the $2!$ is for the repeated **Es**.

> **Note**
>
> You can cancel same integer factorials just like integers.

The next topic of importance is choosing objects from a finite set. For example, if four hockey teams are made from 60 different players, how many teams are possible? This is found using the following relation:

$$\frac{60\times 59\times 58\times 57}{4\times 3\times 2\times 1}=487,635\,\text{possible teams}$$

Generally, this type of problem can be solved using this relation:

$$(n, r) = \frac{n!}{(n-r)!r!} \tag{13.4.1}$$

called **n choose r** where n is the number of objects and r is the number of groups

Using Equation 13.4.1, for the example above the math would be:

$$\begin{aligned} \frac{60!}{56! \times 4!} &= \\ \frac{60 \times 59 \times 58 \times 57 \times 56!}{56! \times 4!} &= \\ \frac{60 \times 59 \times 58 \times 57}{4!} &= 487,635 \text{ possible teams} \end{aligned}$$

Joint Probability

Joint probability is the statistical measure where the likelihood of two events occurring together and at the same point in time are calculated. Because joint probability is the probability of two events occurring at the same time, it can only be applied to situations where more than one observation can be made at the same time. When looking at only two random variables, A and B , this is called **bivariate distribution**, however this can be applied to numerous events or random variables being measured at one time (multivariate distribution). The probability of two events, A and B, both occurring is expressed as:

$$P(A, B)$$

Joint probability can also be expressed as:

$$P(A \cap B)$$

This is read as the probability of the intersection of A and B.

If A, B, and C are independent random variables, then

$$P(A, B, C) = P(A)P(B)P(C)$$

Example 13.4.1

Two cards are selected randomly from a standard deck of cards (no jokers). Between each draw the card chosen is replaced back in the deck. What is the probability of choosing a four then a five? Let $P(A)$ denote the probability that the first card is a four and $P(B)$ denote the probability that the second card is a five.

Solution

If there are 52 cards in a standard card deck with 4 suits, then $P(A) = 4/52$ and $P(B) = 4/52$. Knowing that the events are independent, each probability is multiplied together to find the overall probability for the set of events. Therefore:

$$\begin{aligned} P(A, B) &= P(A) \times P(B) \\ &= \left(\frac{4}{52}\right)\left(\frac{4}{52}\right) \\ &= 1/169 \end{aligned}$$

The probability of choosing a four then a five from the deck with replacement is 1 out of 169.

Conditional Probability

Conditional probability is the probability of one event occurring, given that another event occurs. The following expression describes the conditional probability of event A given that event B has occurred:

$$P(A \mid B)$$

If the events A and B are dependent events, then the following expression can be used to describe the conditional probability of the events:

$$P(A \mid B) = \frac{P(A, B)}{P(B)}$$

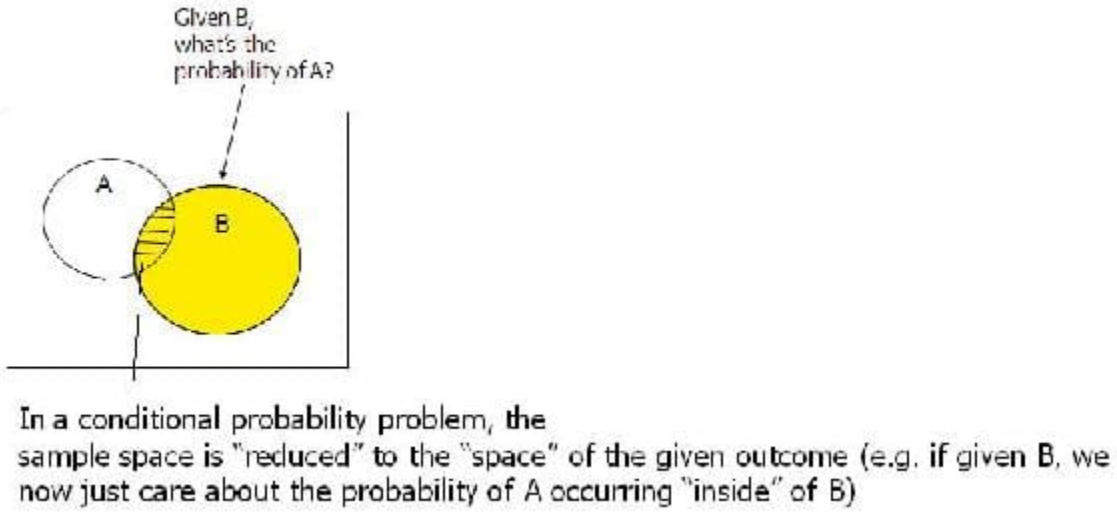
$$P(B \mid A) = \frac{P(A, B)}{P(A)}$$

This can be rearranged to give their joint probability relationship:

$$\begin{aligned} P(A, B) &= P(B \mid A) \times P(A) \\ &= P(A \mid B) \times P(B) \end{aligned}$$

This states that the probability of events A and B occurring is equal to the probability of B occurring given that A has occurred multiplied by the probability that A has occurred.

A graphical representation of conditional probability is shown below:



Conditional probability is often derived from tree diagrams or contingency tables. Suppose you manufacture 100 piston shafts. Event A: feature A is not defective Event B: feature B is not defective

		Feature B: (Roundness)	
		Defective	Not Defective
Feature A (Surface Finish)	Defective	80	9
	Not Defective	6	5

$$P(\text{A Not Def}|\text{B is Def}) = 6/(80+6) = 0.0698$$

$$P(\text{A Not Def}|\text{B Not Def}) = 5/(9+5) = 0.3571$$

Example 13.4.2

Two fair, or unbiased, dice are tossed. Some example outcomes are (1,6) and (3,2). The probability of each possible outcome of the dice is 1/36. When the first die is rolled it results in a value of 2. Once the second die is rolled, what it the probability that the sum of the dice is 7?

Solution

Since it is known that the first value is 2, the possible combination of the two die are as follows:

(2,1) (2,2) (2,3) (2,4) (2,5) (2,6)

This results in six outcomes with equal probabilities since the second die is fair. Therefore, the conditional probability of the outcomes above is 1/6. The conditional probability of the remaining 30 combinations is 0 since the first die is not a 2 in these cases. Finally, since only one of these six outcomes can sum up to 7, (2,5), the probability is 1/6 for rolling a sum of 7 given the value of the first die is a 2.

Example 13.4.3

The probability that a rare species of hamster will give birth to one male and one female is 1/3. The probability that the hamster will give birth to a male is 1/2. What is the probability that the hamster will give birth to a female knowing that the hamster has given birth to a male? Let A denote the probability of giving birth to a male and B denote the probability of giving birth to a female.

Solution

- $P(A$ is the probability of giving birth to a male
- $P(B \mid A$ is the probability of giving birth to a female given that birth of a male has already occurred
- $P(A, A$ is the probability of giving birth to one male and one female

These events are dependent so the following equation must be used:

$$P(A, B) = P(B \mid A) * P(A)$$

Rearranging this equation to find $P(B \mid A$ would give:

$$P(B \mid A) = \frac{P(A, B)}{P(A)}$$

Plugging in the known values would give:

$$P(B \mid A) = (1/3)/(1/2)$$

$$P(B \mid A) = 2/3$$

Therefore, the probability of giving birth to a female, given that birth of a male already occurred is 2/3.

Law of Iterative Expectation

An important application of conditional probability is called the "Law of Iterative Expectation".Given simply, it is: E[X]=E[E[X|Y]]. If the random variable distribution of X is unknown, but we are given the distribution of the conditional variable of X, then by finding the expection of the conditional variable twice, we can return to the expectation of the original random variable X. By looking at the expectation of the random variable X, we can deduce the distribution of the random variable X.

The Law of Iterative Expectation is quite useful in mathematics and often used to prove important relationships. Note the example below:

Example 13.4.1

Use the Law of Iterative Expection to find Var[X] given only X|Y.

Solution

E[Var(X|Y)] = E[X^2]-E[(E[X|Y])^2]

Var(E[X|Y])=E[(E[X|Y])^2]-(E[X])^2

E[Var(X|Y)]+ Var(E[X|Y]) = E[X^2]-E[(E[X|Y])^2] + E[(E[X|Y])^2]-(E[X])^2

Var[X]= E[(E[X|Y])^2] + E[(E[X|Y])^2] (by definition)

thus Var[X] = E[Var(X|Y)]+ Var(E[X|Y])

Marginal Probability

Marginal probability is the unconditional probability of one event; in other words, the probability of an event, regardless of whether another event occurs or not. Finding the marginal probability of an event involves summing all possible configurations of the other event to obtain a weighted average probability. The marginal probability of an event A is expressed as:

$$P(A) = \sum_{B} P(A, B) = \sum_{B} P(A \mid B) * P(B)$$

The marginal probability (of A) is obtained by summing all the joint probabilities. Marginal probability can be used whether the events are dependent or independent. If the events are independent then the marginal probability is simplified to simply the probability. The following example will clarify this computation.

Example 13.4.4

The joint probabilities for rain and wind are given in the following table

	No wind	Some wind	Strong wind	Storm
No rain	0.1	0.2	0.05	0.01
Light rain	0.05	0.1	0.15	0.04
Heavy rain	0.05	0.1	0.1	0.05

Marginal probability of no rain = sum of Joint probabilities = .1 + .2 + .05+.01 = .36

Similarly, marginal probability of light rain = .05+.1+.15+.04 = .34

Similarly, marginal probability of heavy rain = .3.

Marginalizing Out a Factor

In a system with two or more factors affecting the probability of the output of another factor, one of these initial factors can be marginalized out to simplify calculations if that factor is unknown.

For instance, consider a system where A and B both affect the output of C. If the condition of B is unknown but its probability is known, it can be marginalized out to put the system in terms of how only A affects C, using the equation below:

$$P(C \mid A) = \sum_i P(C \mid A, B_i)\, P(B_i)$$

Example Problem 2

The table below shows the probablitiy of having a large, small or no sand storm, if there is high, medium or no wind, and depending on if there is rain. The next table show the probability of rain.

		Sand Storm		
Wind Speed	Rain	P(Large)	P(small)	P(no)
High	yes	0.4	0.4	0.2
High	no	0.6	0.3	0.1
Meduim	yes	0.1	0.3	0.6
Meduim	no	0.15	0.35	0.5
Low	yes	0	0.1	0.9
Low	no	0	0.25	0.75

Rain	P
Yes	0.25
no	0.75

From this it is possible to calculate the probability of a large, small or no sand storm defendant just on the wind speed:

Similarly as to above:

P(Sandstorm Size|Wind Speed)= P(Sandstorm Size|Wind Speed, Rain)*P(Rain)+P(Sandstorm Size|Wind Speed, No Rain)*P(No Rain)

		Sand Storm	
Wind Speed	P(Large)	P(small)	P(no)
High	0.4*0.25+0.6*0.75=0.55	0.4*0.25+0.3*0.75=0.325	0.2*0.25+0.1*0.75=0.125
Medulm	0.1*0.25+0.15*0.75=0.1375	0.3*0.25+0.35*0.75=0.3375	0.6*0.25+0.5*0.75=0.525
Low	0*0.25+0*0.75=0	0.1*0.25+0.25*0.75=0.2125	0.9*0.25+0.75*0.75=0.7575

Relationships Between Events

Knowing whether two events are independent or dependent can help to determine which type of probability (joint, conditional, or marginal) can be calculated. In some cases, the probability of an event must be calculated without knowing whether or not related events have occurred. In these cases, marginal probability must be used to evaluate the probability of an event.

Independence

If the two events are considered independent, each can occur individually and the outcome of one event does not affect the outcome of the other event in any way.

Let's say that A and B are independent events. We'll examine what this means for each type of probability.

Independence in Conditional Probability

Independent events technically do not have a conditional probability, because in this case, A is not dependent on B and vice versa. Therefore, the probability of A given that B has already occurred is equal to the probability of A (and the probability of B given A is equal to the probability of B). This can be expressed as:

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

Independence in Joint Probability

Independent events can have a joint probability, even though one event does not rely on the other. For example, joint probability

is used to describe an event such as tossing a coin. The outcome when the coin is tossed the first time is not related to the outcome when the same coin tossed a second time. To calculate the joint probability of a set of events we take the product of the individual probabilities of each event in the set. This can be expressed as:

$$P(A,B) = P(A)P(B)$$

Independent events can also occur in a case where there are three events. For example, if two dice are being rolled with a sum of 6 (Event A). Let event B represent that the first die is a 1 and let event C represent that the second rolled die is a 5. To prove if these events are independent, the following relation is considered:

$$P(A,B,C) = P(A)P(B)P(C)$$

$$P(A,B) = P(A)P(B)$$

$$P(A,C) = P(A)P(C)$$

$$P(B,C) = P(B)P(C)$$

If any of these relations are false than the event is not independent. When considering events of more than three, the same relation would follow but with an additional relation to event D.

Dependence

If the two events are considered dependent, then the outcome of the second event depends on the probability of the first event. The probabilities of the individual events must be analyzed with conditional probability.

Let's now say that A and B are dependent events. We'll examine what this means for each type of probability.

Dependence in Conditional Probability

Conditional probability only applies to dependent events. In other words, A must depend on B in order to determine the probability of A occurring given that B has already occurred. Therefore, for dependent events A and B, one can just apply the equations as seen in the conditional probability section.

$$P(A \mid B) = \frac{P(A,B)}{P(B)}$$

and

$$P(B \mid A) = \frac{P(A,B)}{P(A)}$$

Dependence in Joint Probability

Joint probability can also be calculated for dependent events. For example, dependent joint probability can be used to describe cards being drawn from a deck without replacing the cards after each consecutive drawing. In this case, the probability of A and B happening is more complex, since the probability of the B happening depends on the probability of A happening. It can be expressed as:

$$P(A,B) = P(A)P(B \mid A)$$

Note that this equation is found by rearranging the conditional probability equation.

Bayes' Theorem

Most probability problems are not presented with the probability of an event "A," it is most often helpful to condition on an event A"." At other times, if we are given a desired outcome of an event, and we have several paths to reach that desired outcome, Baye's Theorem will demonstrate the different probabilities of the pathes reaching the desired outcome. Knowing each probability to reach the desired outcome allows us to pick the best path to follow. Thus, Baye's Theorem is most useful in a scenario of which when given a desired outcome, we can condition on the outcome to give us the separate probabilities of each condition that lead to the desired outcome.

The following is Bayes' Theorem:

$$P\left(B_j \mid A\right)=\frac{P\left(A \mid B_j\right) P\left(B_j\right)}{\sum_j P\left(A \mid B_j\right) P\left(B_j\right)} \tag{13.4.2}$$

where $P\left(A \mid B_j\right)$ is the probability of A conditioned on B_j and $\sum_j P\left(A \mid B_j\right) P\left(B_j\right)$ is the law of total probability.

Derivation of Bayes' Theorem

The derivation of Bayes' theorem is done using the third law of probability theory and the law of total probability.

Suppose there exists a series of events: B_1, B_2 ,..., B_n and they are **mutually exclusive**; that is, $B_1 \cap B_2 \cap \ldots \cap B_n=0$.

This means that only one event, B_j, can occur. Taking an event "A" from the same sample space as the series of B_i, we have:

$$A=\cup_j A B_j$$

Using the fact that the events AB_i are mutually exclusive and using the third law of probability theory:

$$P(A)=\sum_j P(A B_j)$$

Conditioning on the above probability, the result below is also called "the law of total probability"

$$P(A)=\sum_j P(A|B_j)P(B_j)$$

Using the definition of conditional probabilities:

$$P\left(B_j \mid A\right)=\frac{P\left(A \mid B_j\right) P\left(B_j\right)}{P(A)}$$

Putting the above two equations together, we have the **Bayes' Theorem** (Equation 13.4.2):

$$P\left(B_j \mid A\right)=\frac{P\left(A \mid B_j\right) P\left(B_j\right)}{\sum_j P\left(A \mid B_j\right) P\left(B_j\right)}$$

Real world/Chemical Applications

Bayes' rule can be used to predict the probability of a cause given the observed effects. For example, in the equation assume B represents an underlying model or hypothesis and A represents observable consequences or data. So,

$$P(\text{data} \mid \text{model})=\frac{P(\text{model} \mid \text{data}) * P(\text{data})}{P(\text{model})}$$

Where

$P(data|model)$: probability of obtaining observed data given certain model

$P(model|data)$: probability that certain model gave rise to observed data

$P(model)$: probability of occurence of the model prior to taking the data into account

Another application: Bayes' rule estimation is used to identify species in single molecule Fluorenscence microscopy. More information can be found on [1]

Underlying Principles and Significance of Bayes' Rule

As stated previously Bayes' Rule allows for changing the probability of an event based on new information about existing knowledge or expertise. For example, in the previous subsection the probability of the occurrence of the model, P(model), would be obtained through many trials conducted previously and thus be considered existing knowledge. The probability of the data, P(data), would then be considered new information. Bayes' Rule essentially uses this new information to upgrade the existing knowledge and then determine the probability of the new information based on the upgraded existing knowledge. Traditional, or Frequentist, statistics would differ from Bayesian statistics by comparing P(data) to P(model) and determine, with 95% confidence, if P(data) is statistically significant to P(model).

Bayesian theory is being used by many companies and institutions to better classify errors and calculate uncertainty. It has proven to perform better than averaging techniques and is used in safety systems as well as computing end states.

Example 13.4.1

A true-false question is posed to a team of two ChE students on a quiz show. Both the students will, independently, choose their answer. Team member A knows the correct answer, while team member B has the wrong answer. Use Marginal Probability to find the best strategy for the team?

- Strategy 1: Choose one of them and let that person answer the question without knowing what the other student thought was the correct answer.
- Strategy 2: Have both students consider the question and then give the common answer they agree on, or if they disagree, flip a coin to determine which answer to give.

Solution

Strategy 1

We will break down the different possibilities for this situation using the following variables:

- C is the event that they submit the correct answer
- A is the event that student A is chosen to answer
- B is the event that student B is chosen to answer

$$P(C) = P(C, A) + P(C, B) = P(C \mid A) * P(A) + P(C \mid B) * P(B)$$

$P(A) = P(B) = 1/2$ because there is an equal chance for either student to be chosen

$P(C|A) = 1$ because student A is correct

$P(C|B) = 0$ because student B is incorrect

$$P(C) = 1 * (1/2) + 0 * (1/2) = 1/2$$

$$P(C) = 1/2$$

Strategy 2

Since we know the students will disagree on their answer, they must flip a coin in order to decide. Therefore the probability is simply 1/2.

Both strategies provide a 1/2 chance in getting the answer correct, therefore they are equally accurate.

Example 13.4.2

A biologist is studying human cells that are virally infected by a deadly disease, which is present among 0.01% of the population. When preparing a sample, he mishandles a vial with infected cells and it ends up breaking, cutting through the protective gloves and cutting his hand. Even though he immediately washed his hands, he is worried that he has contracted the disease. Luckily for him, his laboratory has been developing a test for the disease and has performed trials with infected and non-infected patients alike. So, the test is administered and to the biologist's surprise, the test turns up negative. He becomes relieved, but then remembers that the test does not always work, so he decides to use Bayes' Rule to determine the probability that he has the disease. He goes into the laboratory database to get all of the data he needs to determine this.

Solution

He wants to determine the probability that he tested negative but is actually positive for the disease, or $P(\text{ DISpos} \mid \text{TESTneg})$.

From Bayes' Rule,

$$P(DISpos|TESTneg) = \frac{P(TESTneg|DISpos) * P(DISpos)}{P(TESTneg)}$$

To find the overall probability of a negative test, you need to use marginal probability and sum over all events, that is both having the disease and not having the disease:

$$P(TESTneg) = P(TESTneg|DISpos) * P(DISpos) + P(TESTneg|DISneg) * P(DISneg)$$

Assuming that all people without the virus do not test positive, and that 99% of infected patients test postive, what is the probability that the biologist is actually positive with the disease?

$$P(TESTneg) = .01 * .0001 + 1 * .9999 = 0.999901$$

$$P(DISpos|TESTneg) = \frac{(0.01 * 0.0001)}{.999901} = 1.0001E-06$$

The chance that the biologist has the disease is roughly one in a million.

Example 13.4.3: Independence in joint probability

If there are 52 cards in a standard card deck with 4 suits, then the probably of picking a card of a particular numerical (5 of spades, 5 of clubs, 5 of diamonds, 5 of hearts) value is P(A) = 1/13. The probability associated with any subsequent draw of a card is only dependent upon the remaining cards in the deck. If the same card is taken from the deck four times in a row and not returned (resulting in no fives in the deck), what is the probability of picking a four, returning it to the deck and then picking a 6.

Solution

The probability of each independent event is given below.

$$P(4) = 4/48$$
$$P(6) = 4/48$$

To find the probability of both events happening, one must multiply the probability of each independent event.

$$P = P(4) * P(6) = 1/144.$$

Example 13.4.4

A Gambler has two coins in his pocket, a fair coin and a two-headed coin. He picks one at random from his pocket, flips it and gets heads.

a. What is the probability that he flipped the fair coin?
b. If he flips the same coin a second time and again gets heads, what is the probability that he flipped the fair coin?

Solution

Call F the event he picked the fair coin and B the event he picked the biased coin. There are two coins in his pocket so the probability of picking the either coin is 0.50. Let H_1 be the event "the first flip is heads" and H_2 be the event "the second flip is heads".

For (a)

$$P(F \mid H_1) = P(H_1 \mid F) * \frac{P(F)}{P(H_1)} = \frac{P(H_1 \mid F)\,P(F)}{P(H_1 \mid F)\,P(F) + P(H_1 \mid B)\,P(B)}$$

Where $P(H_1|F) = 1/2$ and $P(H_1|B) = 1$, so

For (b)

This question is solved in the same way as (a)

Example 13.4.5

Consider the situation where you're going on a bus tour of the local refinery and the person you like has asked you to find them because they're going to save a seat for you. There are three buses that are taking the ChemE students on the tour and you need to immediately decide which is the best one to get on. The first bus has room for 25 students, the second has space for 45 and the third has space for 55. To decide which is the best bus, find the **expected value** (or expected number) of passengers for the bus your friend is on, given the following expression:

$$\mathrm{E}(X) = \sum_i x_i p(x_i)$$

Solution

The expected value of X is the sum of the products of x_i multiplied by p(x_i)

$$E(passengers) = (25 * (1/3) - 45 * (1/3) + 55 * (1/3i) = (25 + 45 + 55)/3 = 135/3 = 4.5$$

Therefore, the expected number of passengers for the bus your friend is on equals the sum of the passengers on each bus divided by the total number of buses, assuming equal probability of your friend being on any of the 3 buses. This shows that you should get on the second bus in order to find your friend.

Alternative method Now consider that you don't want to give each bus an equal chance that your friend is on it. The objective in this alternative method is to find the probability that your friend is on each bus.

$$P(\text{ Bus 1}) = 25/135 = 0.185185$$
$$P(\text{ Bus2 }) = 45/135 = 0.333333$$
$$P(\text{ Bus3 }) = 55/135 = 0.407407$$

Now using the new probability we can find a more accurate expected number of passengers on the bus your friend is on.

$$E'(\text{ passengers }) = 25 * (0.18518) + 45 * (0.333333) + 55 * (0.407407) = 42.037$$

This shows a similar result, you should get on the second bus, but the probabilities associated to each bus are adjusted to account for the number of passengers on each one.

Exercise 13.4.1

What is conditional probability?

a. A way of calculating probability under varying conditions
b. The probability of an event, given that another event has occurred
c. Probability that can not be determined
d. Probability that is always very small

Exercise 13.4.1

In Bayes' rule, what is P(A)?

a. The probability of A
b. The probability of A, given B
c. The probability of B, given A
d. The marginal probability of A

Answer available at the end of this section.

Exercise 13.4.3

Which type of probability does this statement infer to: What is the probability that a temperature sensor fails given that a flow sensor has failed?

a. Joint Probability
b. Conditional Probability
c. Marginal Probability

Answer available at the end of this section.

References

- Ross, Sheldon, A First Course in Probability: Seventh Edition, 2006, Pearson Education, Inc.
- Woolf, Keating, Burge, Yaffe,Statistics and probability primer for computational biologists, spring 2004, MIT.
- [2]University of Oslo
- [3], Wikipedia

Answers

Exercise 13.4.1

A

Exercise 13.4.3

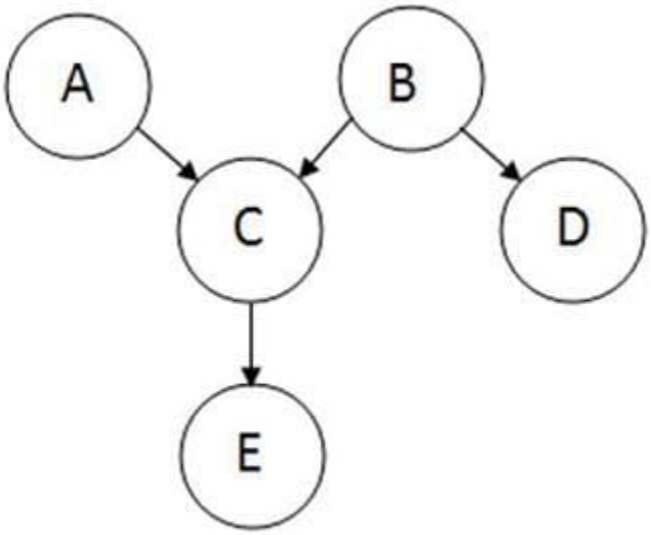
B

This page titled 13.4: Bayes Rule, Conditional Probability and Independence is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Ashwini Miryala, Brandon Kountz, Kyle Scarlett, & Zachary Zell via source content that was edited to the style and standards of the LibreTexts platform.

13.5: Bayesian Network Theory

Introduction

Bayesian network theory can be thought of as a fusion of incidence diagrams and Bayes' theorem. A **Bayesian network**, or **belief network**, shows conditional probability and causality relationships between variables. The probability of an event occurring given that another event has already occurred is called a conditional probability. The probabilistic model is described qualitatively by a **directed acyclic graph**, or DAG. The vertices of the graph, which represent variables, are called nodes. The nodes are represented as circles containing the variable name. The connections between the nodes are called arcs, or edges. The edges are drawn as arrows between the nodes, and represent dependence between the variables. Therefore, for any pairs of nodes indicate that one node is the parent of the other so there are no independence assumptions. Independence assumptions are implied in Bayesian networks by the absence of a link. Here is a sample DAG:



The node where the arc originates is called the parent, while the node where the arc ends is called the child. In this case, A is a parent of C, and C is a child of A. Nodes that can be reached from other nodes are called descendents. Nodes that lead a path to a specific node are called ancestors. For example, C and E are descendents of A, and A and C are ancestors of E. There are no loops in Bayesian networks, since no child can be its own ancestor or descendent. Bayesian networks will generally also include a set of probability tables, stating the probabilities for the true/false values of the variables. The main point of Bayesian Networks is to allow for probabilistic inference to be performed. This means that the probability of each value of a node in the Bayesian network can be computed when the values of the other variables are known. Also, because independence among the variables is easy to recognize since conditional relationships are clearly defined by a graph edge, not all joint probabilities in the Bayesian system need to be calculated in order to make a decision.

Joint Probability Distributions

Joint probability is defined as the probability that a series of events will happen concurrently. The joint probability of several variables can be calculated from the product of individual probabilities of the nodes:

$$\mathrm{P}\left(X_1, \ldots, X_n\right)=\prod_{i=1}^{n} \mathrm{P}\left(X_i \mid \text { parents } \left(X_i\right)\right)$$

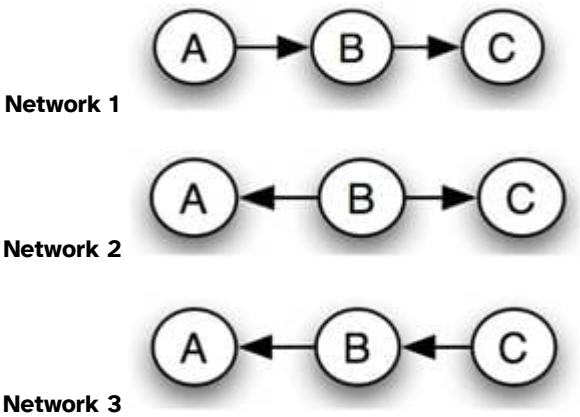
Using the sample graph from the introduction, the joint probability distribution is:

$$P(A, B, C, D, E)=P(A) P(B) P(C \mid A, B) P(D \mid B) P(E \mid C)$$

If a node does not have a parent, like node A, its probability distribution is described as unconditional. Otherwise, the local probability distribution of the node is conditional on other nodes.

Equivalence Classes

Each Bayesian network belongs to a group of Bayesian networks known as an equivalence class. In a given equivalence class, all of the Bayesian networks can be described by the same joint probability statement. As an example, the following set of Bayesian networks comprises an equivalence class:



The causality implied by each of these networks is different, but the same joint probability statement describes them all. The following equations demonstrate how each network can be created from the same original joint probability statement:

Network 1

$$P(A,B,C) = P(A)P(B \mid A)P(C \mid B)$$

Network 2

$$\begin{aligned} P(A,B,C) &= P(A)P(B \mid A)P(C \mid B) \\ &= P(A)\frac{P(A \mid B)P(B)}{P(A)}P(C \mid B) \\ &= P(A \mid B)P(B)P(C \mid B) \end{aligned}$$

Network 3

Starting now from the statement for Network 2

$$\begin{aligned} P(A,B,C) &= P(A \mid B)P(B)P(C \mid B) \\ &= P(A \mid B)P(B)\frac{P(B \mid C)P(C)}{P(B)} \\ &= P(A \mid B)P(B \mid C)P(C) \end{aligned}$$

All substitutions are based on Bayes rule.

The existence of equivalence classes demonstrates that causality cannot be determined from random observations. A controlled study – which holds some variables constant while varying others to determine each one's effect – is necessary to determine the exact causal relationship, or Bayesian network, of a set of variables.

Bayes' Theorem

Bayes' Theorem, developed by the Rev. Thomas Bayes, an 18th century mathematician and theologian, it is expressed as:

$$P(H \mid E, c) = \frac{P(H \mid c) \cdot P(E \mid H, c)}{P(E \mid c)}$$

where we can update our belief in hypothesis H given the additional evidence E and the background information c. The left-hand term, $P(H|E,c)$ is known as the "posterior probability," or the probability of H after considering the effect of E given c. The term $P(H|c)$ is called the "prior probability" of H given c alone. The term $P(E|H,c)$ is called the "likelihood" and gives the probability of the evidence assuming the hypothesis H and the background information c is true. Finally, the last term $P(E|c)$ is called the "expectedness", or how expected the evidence is given only c. It is independent of H and can be regarded as a marginalizing or scaling factor.

It can be rewritten as

$$P(E \mid c) = \sum_i P\left(E \mid H_i, c\right) \cdot P\left(H_i \mid c\right)$$

where *i* denotes a specific hypothesis H_i, and the summation is taken over a set of hypotheses which are mutually exclusive and exhaustive (their prior probabilities sum to 1).

It is important to note that all of these probabilities are conditional. They specify the degree of belief in some proposition or propositions based on the assumption that some other propositions are true. As such, the theory has no meaning without prior determination of the probability of these previous propositions.

Bayes' Factor

In cases when you are unsure about the causal relationships between the variables and the outcome when building a model, you can use **Bayes Factor** to verify which model describes your data better and hence determine the extent to which a parameter affects the outcome of the probability. After using Bayes' Theorem to build two models with different variable dependence relationships and evaluating the probability of the models based on the data, one can calculate Bayes' Factor using the general equation below:

$$BF = \frac{p(\text{model1} \mid \text{data})}{p(\text{model2} \mid \text{data})} = \frac{\frac{p(\text{data}|\text{model1})p(\text{model1})}{p(\text{data})}}{\frac{p(\text{data}|\text{model 2})p(\text{model 2})}{p(\text{data})}} = \frac{p(\text{data} \mid \text{model1})}{p(\text{data} \mid \text{model 2})}$$

The basic intuition is that prior and posterior information are combined in a ratio that provides evidence in favor of one model verses the other. The two models in the Bayes' Factor equation represent two different states of the variables which influence the data. For example, if the data being studied are temperature measurements taken from multiple sensors, Model 1 could be the probability that all sensors are functioning normally, and Model 2 the probability that all sensors have failed. Bayes' Factors are very flexible, allowing multiple hypotheses to be compared simultaneously.

BF values near 1, indicate that the two models are nearly identical and BF values far from 1 indicate that the probability of one model occurring is greater than the other. Specifically, if BF is > 1, model 1 describes your data better than model 2. IF BF is < 1, model 2 describes the data better than model 1. In our example, if a Bayes' factor of 5 would indicate that given the temperature data, the probability of the sensors functioning normally is five times greater than the probability that the senors failed. A table showing the scale of evidence using Bayes Factor can be found below:

scale of Evidence for Bayes factors

(13.5.1)

Although Bayes' Factors are rather intuitive and easy to understand, as a practical matter they are often quite difficult to calculate. There are alternatives to Bayes Factor for model assessment such as the **Bayesian Information Criterion (BIC)**.

The formula for the BIC is:

$$-2 \cdot \ln p(x \mid k) \approx \mathrm{BIC} = -2 \cdot \ln L + k \ln(n).$$

- x = the observed data; n = the number of data points in x, the number of observations, or equivalently, the sample size;
- k = the number of free parameters to be estimated. If the estimated model is a linear regression, k is the number of regressors, including the constant; p(x|k) = the likelihood of the observed data given the number of parameters;
- L = the maximized value of the likelihood function for the estimated model.

This statistic can also be used for non-nested models. For further information on Bayesian Information Criterion, please refer to:

Advantages and Limitations of Bayesian Networks

The advantages of Bayesian Networks are as follows:

- Bayesian Networks visually represent all the relationships between the variables in the system with connecting arcs.
- It is easy to recognize the dependence and independence between various nodes.
- Bayesian networks can handle situations where the data set is incomplete since the model accounts for dependencies between all variables.
- Bayesian networks can maps scenarios where it is not feasible/practical to measure all variables due to system constraints (costs, not enough sensors, etc.)
- Help to model noisy systems.
- Can be used for any system model - from all known parameters to no known parameters.

The limitations of Bayesian Networks are as follows:

- All branches must be calculated in order to calculate the probability of any one branch.
- The quality of the results of the network depends on the quality of the prior beliefs or model. A variable is only a part of a Bayesian network if you believe that the system depends on it.
- Calculation of the network is NP-hard (nondeterministic polynomial-time hard), so it is very difficult and possibly costly.
- Calculations and probabilities using Baye's rule and marginalization can become complex and are often characterized by subtle wording, and care must be taken to calculate them properly.

Inference

Inference is defined as the process of deriving logical conclusions based on premises known or assumed to be true. One strength of Bayesian networks is the ability for inference, which in this case involves the probabilities of unobserved variables in the system. When observed variables are known to be in one state, probabilities of other variables will have different values than the generic case. Let us take a simple example system, a television. The probability of a television being on while people are home is much higher than the probability of that television being on when no one is home. If the current state of the television is known, the probability of people being home can be calculated based on this information. This is difficult to do by hand, but software programs that use Bayesian networks incorporate inference. One such software program, Genie, is introduced in Learning and analyzing Bayesian networks with Genie.

Marginalization

Marginalization of a parameter in a system may be necessary in a few instances:

- If the data for one parameter (P1) depends on another, and data for the independent parameter is not provided.
- If a probability table is given in which P1 is dependent upon two other system parameters, but you are only interested in the effect of one of the parameters on P1.

Imagine a system in which a certain reactant (R) is mixed in a CSTR with a catalyst (C) and results in a certain product yield (Y). Three reactant concentrations are being tested (A, B, and C) with two different catalysts (1 and 2) to determine which combination will give the best product yield. The conditional probability statement looks as such:

$$P(R, C, Y) = P(R)P(C)P(Y \mid R, C)$$

The probability table is set up such that the probability of certain product yield is dependent upon the reactant concentration *and* the catalyst type. You want to predict the probability of a certain product yield given only data you have for catalyst type. The concentration of reactant must be marginalized out of *P(Y|R,C)* to determine the probability of the product yield without knowing the reactant concentration. Thus, you need to determine *P(Y|C)*. The marginalization equation is shown below:

$$P(Y \mid C) = \sum_i P(Y \mid R_i, C)\, P(R_i)$$

where the summation is taken over reactant concentrations A, B, and C.

The following table describes the probability that a sample is tested with reactant concentration A, B, or C:

Reactant	P(Reactant)
A	0.2
B	0.5
C	0.3

This next table describes the probability of observing a yield - High (H), Medium (M), or Low (L) - given the reactant concentration and catalyst type:

Reactant	Catalyst	P(Y=H)	P(Y=M)	P(Y=L)
A	1	0.51	0.08	0.41
A	2	0.3	0.2	0.5
B	1	0.71	0.09	0.2
B	2	0.92	0.05	0.03
C	1	0.21	0.4	0.39
C	2	0.12	0.57	0.31

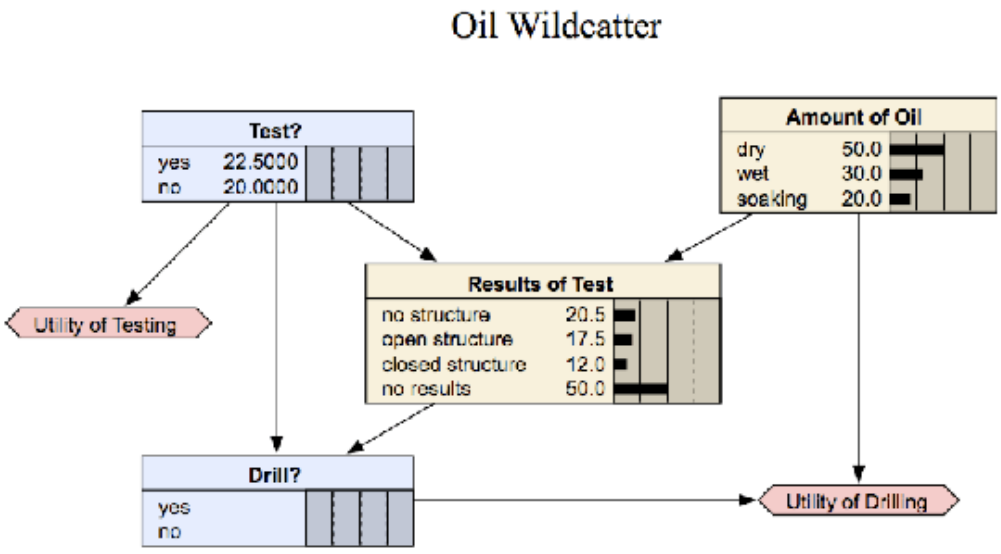
The final two tables show the calculation for the marginalized probabilities of yield given a catalyst type using the marginalization equation:

Catalyst	P(Y=H)	P(Y=M)	P(Y=L)
1	0.2*0.51+0.5*0.71+0.3*.21	0.2*0.08+0.5*0.09+0.3*0.4	0.2*0.41+0.5*0.2+0.3*0.39
2	0.2*0.3+0.5*0.92+0.3*0.12	0.2*0.2+0.5*0.05+0.3*0.57	0.2*0.5+0.5*0.03+0.3*0.31

Catalyst	P(Y=H)	P(Y=M)	P(Y=L)
1	0.520	0.181	0.299
2	0.556	0.236	0.208

Dynamic Bayesian Networks

The static Bayesian network only works with variable results from a single slice of time. As a result, a static Bayesian network does not work for analyzing an evolving system that changes over time. Below is an example of a static Bayesian network for an oil wildcatter:

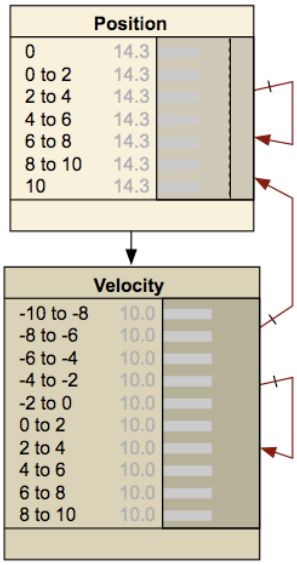


www.norsys.com/netlibrary/index.htm

An oil wildcatter must decide either to drill or not. However, he needs to determine if the hole is dry, wet or soaking. The wildcatter could take seismic soundings, which help determine the geological structure at the site. The soundings will disclose whether the terrain below has no structure, which is bad, or open structure that's okay, or closed structure, which is really good. As you can see this example does not depend on time.

Dynamic Bayesian Network (DBN) is an extension of Bayesian Network. It is used to describe how variables influence each other over time based on the model derived from past data. A DBN can be thought as a Markov chain model with many states or a discrete time approximation of a differential equation with time steps.

An example of a DBN, which is shown below, is a frictionless ball bouncing between two barriers. At each time step the position and velocity changes.



www.norsys.com/netlibrary/index.htm

An important distinction must be made between DBNs and Markov chains. A DBN shows how variables affect each other over time, whereas a Markov chain shows how the state of the entire system evolves over time. Thus, a DBN will illustrate the probabilities of one variable changing another, and how each of the individual variables will change over time. A Markov chain looks at the state

of a system, which incorporates the state of each individual variable making up the system, and shows the probabilities of the system changing states over time. A Markov chain therefore incorporates all of the variables present in the system when looking at how said system evolves over time. Markov chains can be derived from DBNs, but each network represents different values and probabilities.

There are several advantages to creating a DBN. Once the network has been established between the time steps, a model can be developed based on this data. This model can then be used to predict future responses by the system. The ability to predict future responses can also be used to explore different alternatives for the system and determine which alternative gives the desired results. DBN's also provide a suitable environment for model predictive controllers and can be useful in creating the controller. Another advantage of DBN's is that they can be used to create a general network that does not depend on time. Once the DBN has been established for the different time steps, the network can be collapsed to remove the time component and show the general relationships between the variables.

A DBN is made up with interconnected time slices of static Bayesian networks. The nodes at certain time can affect the nodes at a future time slice, but the nodes in the future can not affect the nodes in the previous time slice. The causal links across the time slices are referred to as temporal links, the benefit of this is that it gives DBN an unambiguous direction of causality.

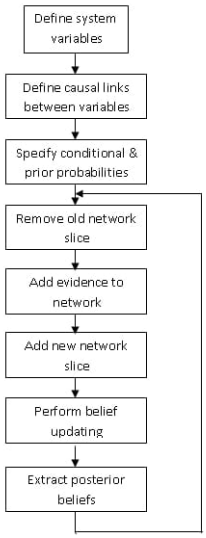
For the convenience of computation, the variables in DBN are assumed to have a finite number of states that the variable can have. Based on this, conditional probability tables can be constructed to express the probabilities of each child node derived from conditions of its parent nodes.

Node C from the sample DAG above would have a conditional probability table specifying the conditional distribution P(C|A,B). Since A and B have no parents, so it only require probability distributions P(A) and P(B). Assuming all the variables are binary, means variable A can only take on A1 and A2, variable B can only take on B1 and B2, and variable C can only take on C1 and C2. Below is an example of a conditional probability table of node C.

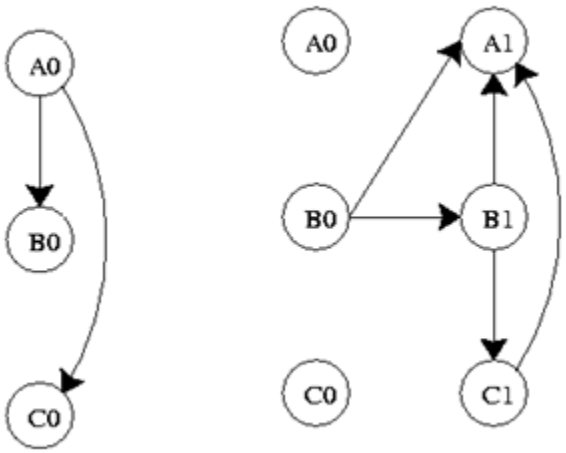
		C	
A	B	C1	C2
A1	B1	P(C=C1 | A=A1,B=B1)	P(C=C2 | A=A1,B=B1)
A1	B2	P(C=C1 | A=A1,B=B2)	P(C=C2 | A=A1,B=B2)
A2	B1	P(C=C1 | A=A2,B=B1)	P(C=C2 | A=A2,B=B1)
A2	B2	P(C=C1 | A=A2,B=B2)	P(C=C2 | A=A2,B=B2)

The conditional probabilities between observation nodes are defined using a sensor node. This sensor node gives conditional probability distribution of the sensor reading given the actual state of system. It embodies the accuracy of the system.

The nature of DBN usually results in a large and complex network. Thus to calculate a DBN, the outcome old time slice is summarized into probabilities that is used for the later slice. This provides a moving time frame and forms a DBN. When creating a DBN, temporal relationships between slices must be taken into account. Below is an implementation chart for DBN.



The graph below is a representation of a DBN. It represents the variables at two different time steps, t-1 and t. t-1, shown on the left, is the initial distribution of the variables. The next time step, t, is dependent on time step t-1. It is important to note that some of these variables could be hidden.

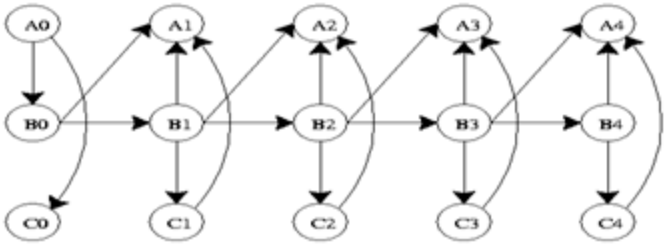


Where Ao, Bo, Co are initial states and Ai, Bi, Ci are future states where i=1,2,3,...,n.

The probability distribution for this DBN at time t is...

$$P\left(Z_t \mid Z_{t-1}\right)=\prod_{i=1}^{N} P\left(Z_t^i \mid \pi\left(Z_t^i\right)\right)$$

If the process continues for a larger number of time steps, the graph will take the shape below.

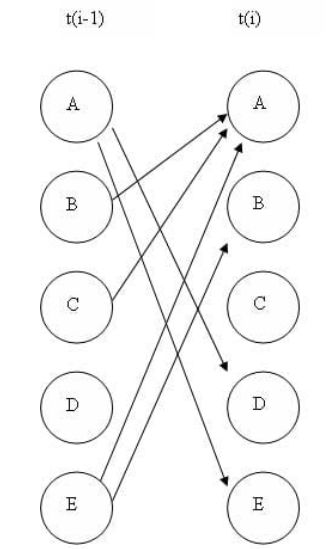


Its joint probability distribution will be...

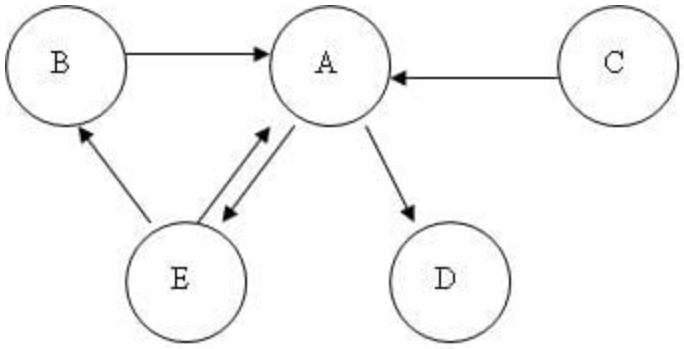
$$P(Z_{1:T}) = \prod_{t=1}^{T} \prod_{i=1}^{N} P\left(Z_t^i \mid \pi\left(Z_t^i\right)\right)$$

DBN's are useful in industry because they can model processes where information is incomplete, or there is uncertainty. Limitations of DBN's are that they do not always accurately predict outcomes and they can have long computational times.

The above illustrations are all examples of "unrolled" networks. An unrolled dynamic Bayesian network shows how each variable at one time step affects the variables at the next time step. A helpful way to think of unrolled networks is as visual representations of numerical solutions to differential equations. If you know the states of the variables at one point in time, and you know how the variables change with time, then you can predict what the state of the variables will be at any point in time, similar to using Euler's method to solve a differential equation. A dynamic Bayesian network can also be represented as a "rolled" network. A rolled network, unlike an unrolled network, shows each variables' effect on each other variable in one chart. For example, if you had an unrolled network of the form:



then you could represent that same network in a rolled form as:



If you examine each network, you will see that each one provides the exact same information as to how the variables all affect each other.

Applications

Bayesian networks are used when the probability that one event will occur depends on the probability that a previous event occurred. This is very important in industry because in many processes, variables have conditional relationships, meaning they are not independent of each other. Bayesian networks are used to model processes in a wide variety of applications. Some of these include...

1. Gene regulatory networks
2. Protein structure
3. Diagnosis of illness
4. Document classification
5. Image processing
6. Data fusion
7. Decision support systems
8. Gathering data for deep space exploration
9. Artificial Intelligence
10. Prediction of weather
11. On a more familiar basis, Bayesian networks are used by the friendly Microsoft office assistant to elicit better search results.\
12. Another use of Bayesian networks arises in the credit industry where an individual may be assigned a credit score based on age, salary, credit history, etc. This is fed to a Bayesian network which allows credit card companies to decide whether the person's credit score merits a favorable application.

Summary: A General Solution Algorithm for the Perplexed

Given a Bayesian network problem and no idea where to start, just relax and try following the steps outlined below.

Step 1: What does my network look like? Which nodes are parents (and are they conditional or unconditional) and which are children? How would I model this as an incidence diagram and what conditional probability statement defines it?

Step 2: Given my network connectivity, how do I tabulate the probabilities for each state of my node(s) of interest? For a single column of probabilities (parent node), does the column sum to 1? For an array of probabilities (child node) with multiple possible states defined by the given combination of parent node states, do the rows sum to 1?

Step 3: Given a set of observed data (usually states of a child node of interest), and probability tables (aka truth tables), what problem am I solving?

- Probability of observing the particular configuration of data, order unimportant

Solution: Apply multinomial distribution

- Probability of observing the particular configuration of data in that particular order

Solution: Compute the probability of each individual observation, then take the product of these

- Probability of observing the data in a child node defined by 2 (or n) parents given only 1 (or n-1) of the parent nodes

Solution: Apply marginalization to eliminate other parent node

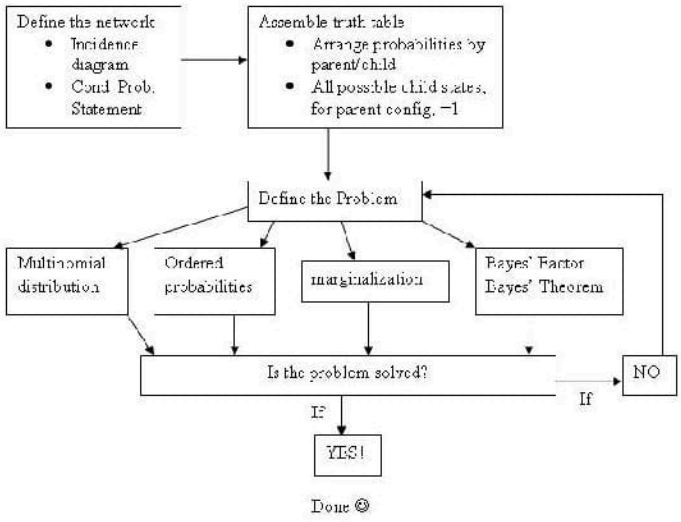
- Probability of a parent node being a particular state given data in the form of observed states of the child node

Solution: Apply Bayes' Theorem
Solve for Bayes' Factor to remove incalculable denominator terms generated by applying Bayes' Theorem, and to compare the parent node state of interest to a base case, yielding a more meaningful data point

Step 4: Have I solved the problem? Or is there another level of complexity? Is the problem a combination of the problem variations listed in step 3?

- If problem is solved, call it a day and go take a baklava break
- If problem is not solved, return to step 3

Graphically:

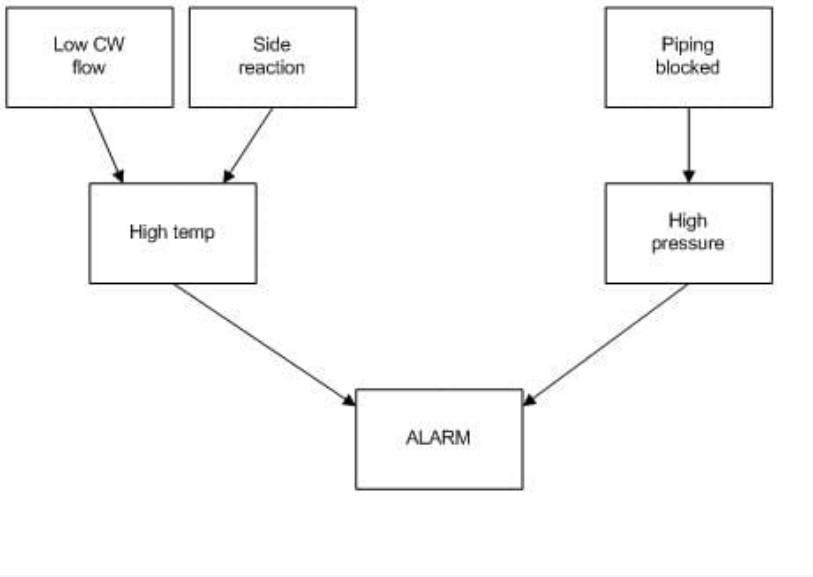


Example 13.5.1

A multipurpose alarm in a plant can be tripped in 2 ways. The alarm goes off if the reactor temperature is too high or the pressure in a storage tank is too high. The reactor temperature may be too high because of a low cooling water flow (1% probability), or because of an unknown side reaction (5% probability). The storage tank pressure might be too high because of a blockage in the outlet piping (2% probability). If the cooling water flow is low and there is a side reaction, then there is a 99% probability that a high temperature will occur. If the cooling water flow is normal and there is no side reaction, there is only a 3% probability a high temperature will occur. If there is a pipe blockage, a high pressure will always occur. If there is no pipe blockage, a high pressure will occur only 2% of the time.

Create a DAG for the situation above, and set up the probability tables needed to model this system. All the values required to fill in these tables are not given, so fill in what is possible and then indicate what further values need to be found.

Solution



The following probability tables describe the system, where CFL = Cold water flow is low, SR = Side reaction present, PB = Pipe is blocked, HT = High temperature, HP = High pressure, A = Alarm. T stands for true, or the event did occur. F stands for false, or the event did not occur. A blank space in a table indicates an area where further information is needed.

P(CFL=T)	0.01
P(CFL=F)	0.99
P(SR=T)	0.05
P(SR=F)	0.95
P(PB=T)	0.02
P(PB=F)	0.98

CFL	SR	P(HT=T)	P(HT=F)
T	T	0.99	0.01
T	F		
F	F	0.03	0.97
F	T		

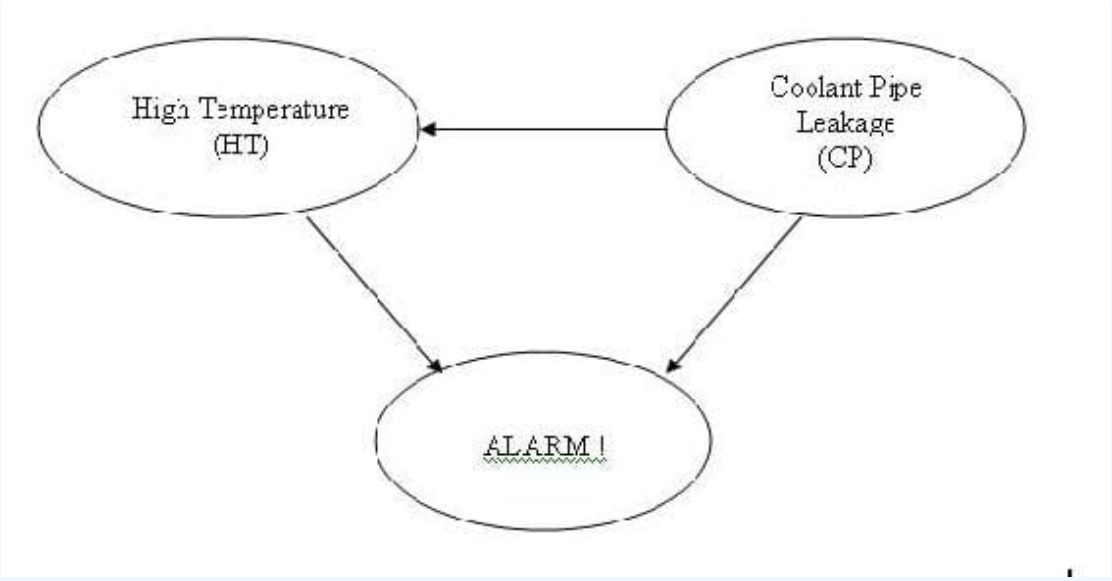
PB	P(HP=T)	P(HP=F)
F	0.02	0.98
T	1.00	0.00

HT	HP	P(A=T)	P(A=F)
T	T		
T	F		
F	F		
F	T		

An advantage of using DAGs becomes apparent. For example, you can see that there is only a 3% chance that there is a high temperature situation given that both cold water flow is not low and that there is no side reaction.However, as soon as the cold water becomes low, you have at least a 94% chance of a high temperature alarm, regardless of whether or not a side reaction occurs. Conversely, the presence of a side reaction here only creates a 90% chance of alarm trigger. From the above probability calculations, one can estimate relative dominance of cause-and-effect triggers. For example you could now reasonably conjecture that the cold water being low is a more serious event than a side reaction.

Example 13.5.2

The DAG given below depicts a different model in which the alarm will ring when activated by high temperature and/or coolant water pipe leakage in the reactor.



The table below shows the truth table and probabilities with regards to the different situations that might occur in this model.

CP	
T	F
0.3	0.7

	HT	
CP	T	F
T	0.9	0.1
F	0	1

		A	
HT	CP	T	F
F	F	0	1
F	T	0.7	0.3
T	F	0.9	0.1
T	T	0.99	0.01

The joint probability function is:

$$P(A, HT, CP) = P(A|HT, CP)P(HT|CP)P(CP)$$

A great feature of using the Bayesian network is that the probability of any situation can be calculated. In this example, write the statement that will describe the probability that the temperature is high in the reactor given that the alarm sounded.

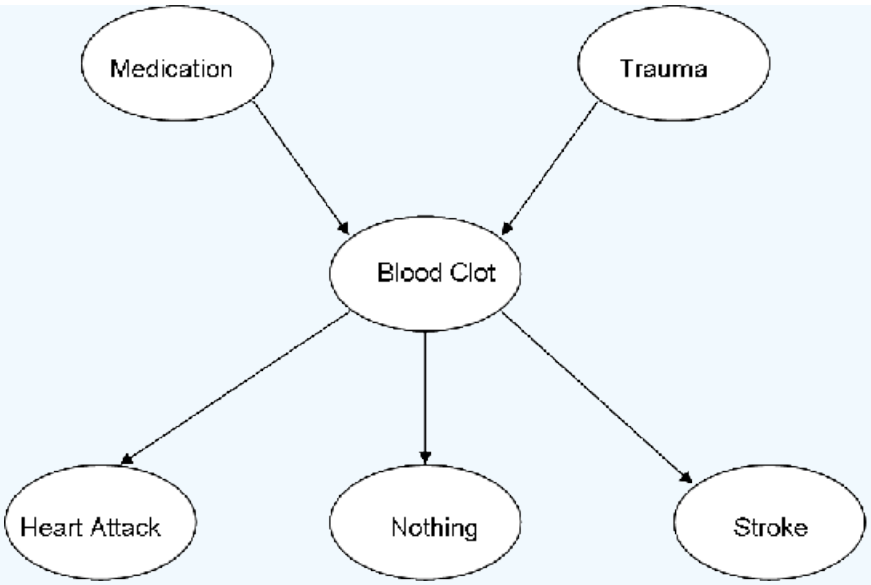
Solution

$$\mathrm{P}(CP = T \mid \Delta = T) = \frac{\mathrm{P}(A = T, CP = T)}{\mathrm{P}(A = 1)} [= \frac{\sum_{HTC[T,P]} P(A - T, HT, CP - T)}{\sum_{HT,CPC[T,P]} P(A = T, HT, CP)}$$

Example 13.5.3

Certain medications and traumas can both cause blood clots. A blood clot can lead to a stroke, heart attack, or it could simply dissolve on its own and have no health implications. Create a DAG that represents this situation.

Solution



b. The following probability information is given where M = medication, T = trauma, BC = blood clot, HA = heart attack, N = nothing, and S = stroke. T stands for true, or this event did occur. F stands for false, or this event did not occur.

P(M=T)	0.2
P(M=F)	0.8
P(T=T)	0.05
P(T=F)	0.95

M	T	P(BC=T)	P(BC=F)
T	T	0.95	0.05
T	F	0.3	0.7
F	T	0.6	0.4
F	F	0.9	0.1

BC	P(HA=T)	P(HA=F)	P(S=T)	P(S=F)	P(N=T)	P(N=F)
T	0.4	0.6	0.35	0.65	0.25	0.75
F	0.15	0.85	0.1	0.9	0.75	0.25

What is the probability that a person will develop a blood clot as a result of both medication and trauma, and then have no medical implications?

Answer

$$\mathrm{P(N, BC, M, T) = P(N \mid BC)P(BC \mid M, T)P(M)P(T) = (0.25)(0.95)(0.2)(0.05) = 0.2375\%}$$

Example 13.5.4

Suppose you were given the following data.

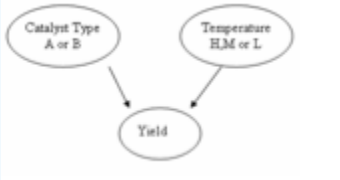
Catalyst	p(Catalyst)
A	0.40
B	0.60

Temperature	Catalyst	P(Yield = H)	P(Yield = M)	P(Yield = L)
H	A	0.51	0.08	0.41
H	B	0.30	0.20	0.50
M	A	0.71	0.09	0.20
M	B	0.92	0.05	0.03
L	A	0.21	0.40	0.39
L	B	0.12	0.57	0.31

How would you use this data to find *p(yield|temp)* for 9 observations with the following descriptions?

# Times Observed	Temperature	Yield
4x	H	H
2x	M	L
3x	L	H

An DAG of this system is below:



Solution

Marginalization! The state of the catalyst can be marginalized out using the following equation:

$p(yield \mid temp) =$	$\sum$	$p(yield \mid temp,cat_i)p(cat_i)$
	$i = A,B$	

The two tables above can be merged to form a new table with marginalization:

Temperature	P(Yield = H)	P(Yield = M)	P(Yield = L)
H	0.51*0.4 + 0.3*0.6 = 0.384	0.08*0.4 + 0.2*0.6 = 0.152	0.41*0.4 + 0.5*0.6 = 0.464
M	0.71*0.4 + 0.92*0.6 = 0.836	0.09*0.4 + 0.05*0.6 = 0.066	0.20*0.4 + 0.03*0.6 = 0.098
L	0.21*0.4 + 0.12*0.6 = 0.156	0.40*0.4 + 0.57*0.6 = 0.502	0.39*0.4 + 0.31*0.6 = 0.342

$$p(\text{yield} \mid \text{temp}) = \frac{9!}{4!2!3!} * \left(0.384^4 * 0.098^2 * 0.156^3\right) = 0.0009989$$

Example 13.5.5

A very useful use of Bayesian networks is determining if a sensor is more likely to be working or broken based on current readings using the Bayesian Factor discussed earlier. Suppose there is a large vat in your process with large turbulent flow that makes it difficult to accurately measure the level within the vat. To help you use two different level sensors positioned around the tank that read whether the level is high, normal, or low. When you first set up the sensor system you obtained the following probabilities describing the noise of a sensor operating normally.

Tank Level (L)	p(S=High)	p(S=Normal)	p(S=Low)
Above Operating Level Range	0.80	0.15	0.05
Within Operating Level Range	0.15	0.75	0.10
Below Operating Level Range	0.10	0.20	0.70

When the sensor fails there is an equal chance of the sensor reporting high, normal, or low regardless of the actual state of the tank. The conditional probability table for a fail sensor then looks like:

Tank Level (L)	p(S=High)	p(S=Normal)	p(S=Low)
Above Operating Level Range	0.33	0.33	0.33
Within Operating Level Range	0.33	0.33	0.33
Below Operating Level Range	0.33	0.33	0.33

From previous data you have determined that when the process is acting normally, as you believe it is now, the tank will be operating above the level range 10% of the time, within the level range 85% of the time, and below the level range 5% of the time. Looking at the last 10 observations (shown below) you suspect that sensor 1 may be broken. Use Bayesian factors to determine the probability of sensor 1 being broken compared to both sensors working.

Sensor 1	Sensor 2
High	Normal
Normal	Normal
Normal	Normal
High	High
Low	Normal
Low	Normal
Low	Low
High	Normal
High	High
Normal	Normal

From the definition of the Bayesian Factor we get.

$$BF = \frac{p(model1|data)}{p(model2|data)} = \frac{\frac{p(data|model1)p(model1)}{p(data)}}{\frac{p(data|model2)p(model2)}{p(data)}} = \frac{p(data|model1)}{p(data|model2)}$$

For this set we will use the probability that we get the data given based on the model.

$$\frac{p(data|model1)}{p(data|model2)}$$

If we consider model 1 both sensors working and model 2 sensor 2 being broken we can find the BF for this rather easily.

p(data | model 1) = p(s1 data | model 1)*p(s2 data | model 1)

For both sensors working properly:

The probability of the sensor giving each reading has to be calculated first, which can be found by summing the probability the tank will be at each level and multiplying by probability of getting a specific reading at that level for each level.

p(s1 = high | model 1) = [(.10)*(.80) + (.85)*(.15) + (.05)*(.10) = 0.2125

p(s1 = normal | model 1) = [(.10)*(.15) + (.85)*(.75) + (.05)*(.20) = 0.6625

p(s1 = low | model 1) = [(.10)*(.05) + (.85)*(.10) + (.05)*(.70) = 0.125

Probability of getting sensor 1's readings (assuming working normally)

$p(s1data \mid model1) = (.2125)^4 * (.6625)^3 * (.125)^3 = 5.450 * 10^{-6}$

The probability of getting each reading for sensor 2 will be the same since it is also working normally

$p(s2data \mid model1) = (.2125)^2 * (.6625)^7 * (.125)^1 = 3.162 * 10^{-4}$

$p(data \mid model1) = (5.450 * 10^{-6}) * (3.162 * 10^{-4}) = 1.723 * 10^{-9}$

For sensor 1 being broken:

The probability of getting each reading now for sensor one will be 0.33.

$p(s1data \mid model2) = (0.33)^4 * (0.33)^3 * (0.33)^3 = 1.532 * 10^{-5}$

The probability of getting the readings for sensor 2 will be the same as model 1, since both models assume sensor 2 is acing normally.

$p(data \mid model2) = (1.532 * 10^{-5}) * (3.162 * 10^{-4}) = 4.844 * 10^{-9}$

$$BF = \frac{1.723 * 10^{-9}}{4.844 * 10^{-9}} = 0.356$$

A BF factor between 1/3 and 1 means there is weak evidence that model 2 is correct.

True or False?

1. Is the other name for Bayesian Network the "Believer" Network?

2. The nodes in a Bayesian network can represent any kind of variable (latent variable, measured parameter, hypothesis..) and are not restricted to random variables.

3. Bayesian network theory is used in part of the modeling process for artificial intelligence.

Answers:

1. F

2. T

3. T

References

1. Aksoy, Selim. "Parametric Models Part IV: Bayesian Belief Networks." Spring 2007.
2. Ben-Gal, Irad. “BAYESIAN NETWORKS.” Department of Industrial Engineering. Tel-Aviv University. <http://www.eng.tau.ac.il/~bengal/BN.pdf>http://www.dcs.qmw.ac.uk/~norman/BBNs/BBNs.htm
3. Charniak, Eugene (1991). "Bayesian Networks without Tears", AI Magazine, p. 8.
4. Friedman, Nir, Linial, Michal, Nachman, Iftach, and Pe’er, Dana. “Using Bayesian Networks to Analyze Expression Data.” JOURNAL OF COMPUTATIONAL BIOLOGY, Vol. 7, # 3/4, 2000, Mary Ann Liebert, Inc. pp. 601–620
5. Guo, Haipeng. "Dynamic Bayesian Networks." August 2002.
6. Neil, Martin, Fenton, Norman, and Tailor, Manesh. “Using Bayesian Networks to Model Expected and Unexpected Operational Losses.” Risk Analysis, Vol. 25, No. 4, 2005 <http://www.dcs.qmul.ac.uk/~norman/papers/oprisk.pdf>
7. Niedermayer, Daryle. “An Introduction to Bayesian Networks and their Contemporary Applications.” December 1, 1998. < www.niedermayer.ca/papers/bayesian/bayes.html>
8. Seeley, Rich. "Bayesian networks made easy". Application Development Trends. December 4, 2007 .
9. http://en.Wikipedia.org/wiki/Bayesian_network

This page titled 13.5: Bayesian Network Theory is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Jamie Polan, Matthew Morabito, Sarah Hebert, & Valerie Lee via source content that was edited to the style and standards of the LibreTexts platform.

13.6: Learning and analyzing Bayesian networks with Genie

Introduction

Genie (Graphical network interface) is a software tool developed at the University of Pittsburgh for Microsoft Windows and available free of charge at Genie. It is useful for decision analysis and for graphically representing the union of probability and networked occurrences. Particularly, Genie can be used for the analysis of Bayesian networks, or directed acylic graphs (i.e. occurrences in a web of happenings are conditionally independent of each other). Bayesian networks or Dynamic Bayesian Networks (DBNs) are relevant to engineering controls because modelling a process using a DBN allows for the inclusion of noisy data and uncertainty measures; they can be effectively used to predict the probabilities of related outcomes in a system. In Bayesian networks, the addition of more nodes and inferences greatly increases the complexity of the calculations involved and Genie allows for the analysis of these complicated systems. Additionally, the graphical interface facilitates visual understanding of the network (Charniak, 1991).

This link gives an example of a complex Bayesian network depicted in the graphical interface of Genie. As can be seen, Genie arranges the network of nodes and inferences in a topology that is easily visualized and is useful for both simple and extremely complex systems.

Using Genie to Construct and Analyze Dynamic Bayesian Networks

This section will include instructions for downloading and installing the Genie software as well as describe the procedure for using Genie to analyze simple Bayesian networks.

Downloading and Installing Genie

The Genie software download is available the Genie website, depicted below:

To download Genie, click on the “Downloads” link on the left menu of the page and highlighted in red in the figure above.

A registration page will appear which requires some basic user information. After this has been entered, click on the Register link.

A Downloads page with several download options will appear. Click on the GeNIe 2.0 link highlighted in red on the figure below. This will initiate the download.

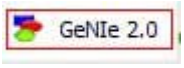
NOTE: The Genie Software download is only available for Windows.

Install the software by following the steps indicated in the Genie installation program.

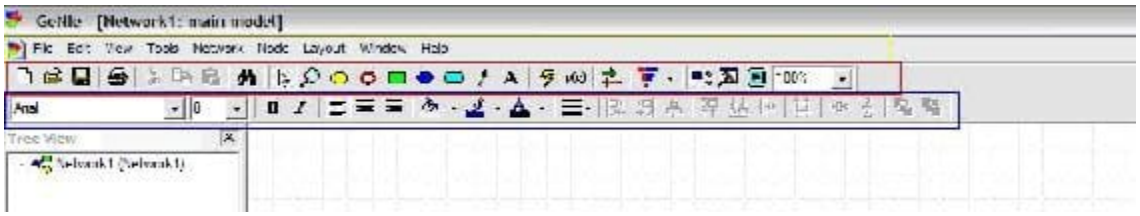
The installation of the Genie software is now complete. Please note the help section of the software features many tutorials describing how to use a wide array of functions. However, this article will be focused on the analysis of Dynamic Bayesian Networks.

- All informational materials used to create this tutorial come from GeNIe

Using GeNIe to Analyze Dynamic Bayesian Networks

If the program was installed in the default location given in the installation wizard a shortcut to GeNIe can be found in the start menu listed under All Programs. Click on the GeNIe 2.0 Icon as seen below: 

The main interface will appear as shown in the figure below:



Note the important tool bars highlighted on the figure above: the menu bar (yellow), the standard toolbar (red), and the format toolbar (blue).

Consider this simple illustration of how to build a Dynamic Bayesian Network using Genie:

Polly popular is the most popular girl on the UM campus who only dates other UM students. It is common knowledge that Polly only accepts 30% of the invitations she receives to go on a date. Of the invitations that she accepts, 40% of them come from guys with high GPAs, 40% of them come from guys with medium GPAs and 20% of them come from guys with low GPAs. Of the invitations that she rejects, 10% of them come from guys with high GPAs, 30% of them come from guys with medium GPAs, and 60% of them come from guys with low GPAs.

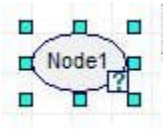
NOTE: In this example, the variable of GPA is discretized into categories of high, medium, and low. This is necessary for analysis

using Genie because otherwise there would exist infinite states for each node in the DBN (Charniak, 1991).

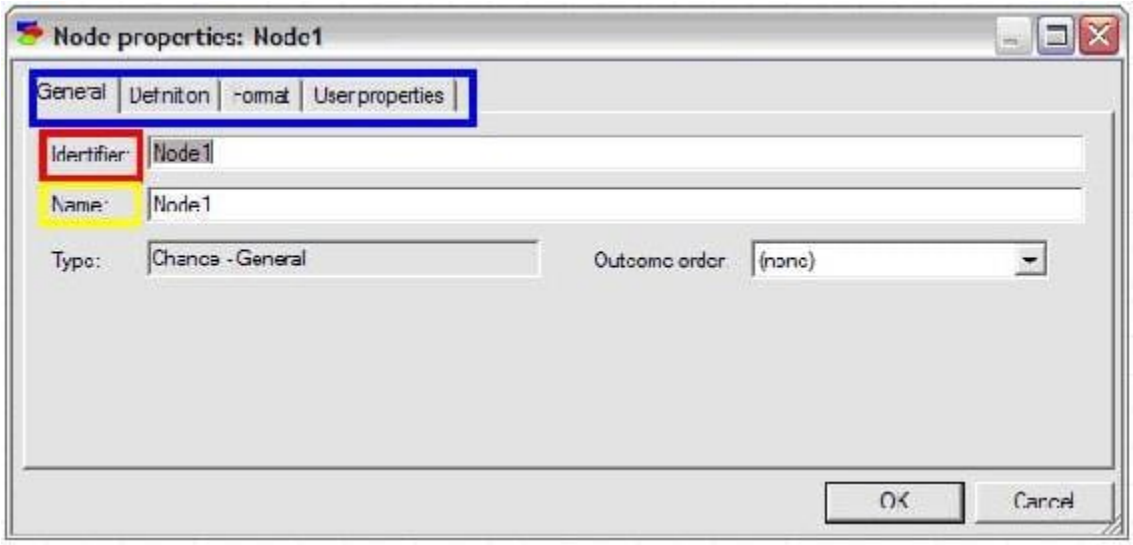
Creating a Bayesian network allows for determinations to be made about probabilities of Polly accepting or rejecting certain date invitations. Shown below is how to use Genie to find the probability of Polly accepting a date invitation from a guy if she knows the guy inviting her has a high GPA.

First, a node is created for the variable called acceptance of invitation. Select the “chance” node from the standard toolbar as is shown in the figure below highlighted in red.

Left-click on a clear part of the graph area of the screen. An oval will appear with “Node1” inside as seen in the figure below.

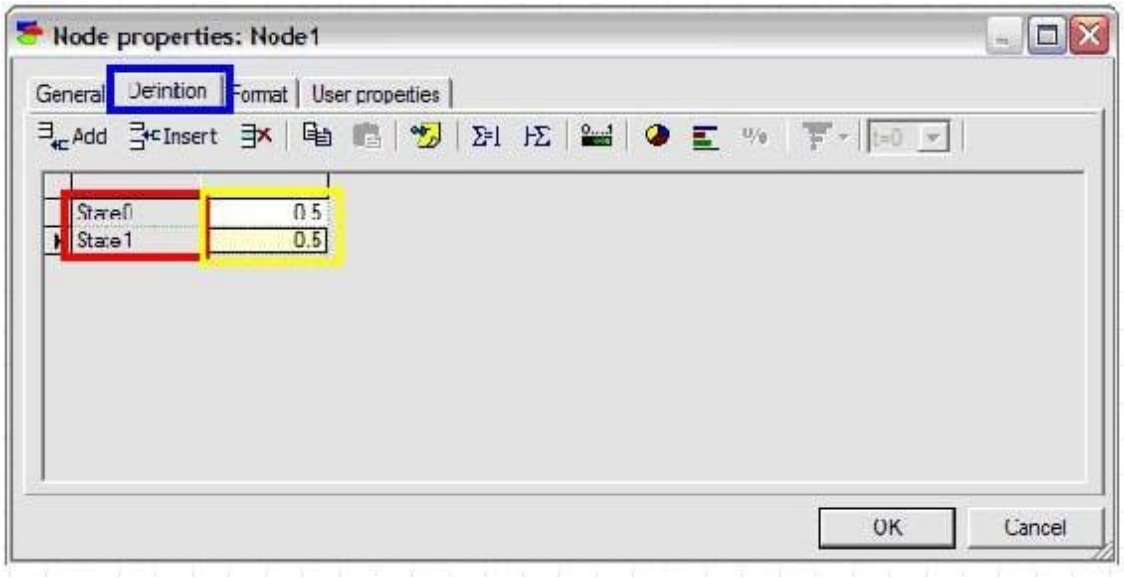


The edit mode for the node should come up automatically; if not,simply double-click on the node to pull up the edit screen as depicted in the figure below.



Enter an identifier for the node (must be a valid variable name in a programming language) and enter a name for the node (can be a string of any length). For this example, the identifier is entered as “Acceptance” and the name is entered as “Acceptance of Invitation”.

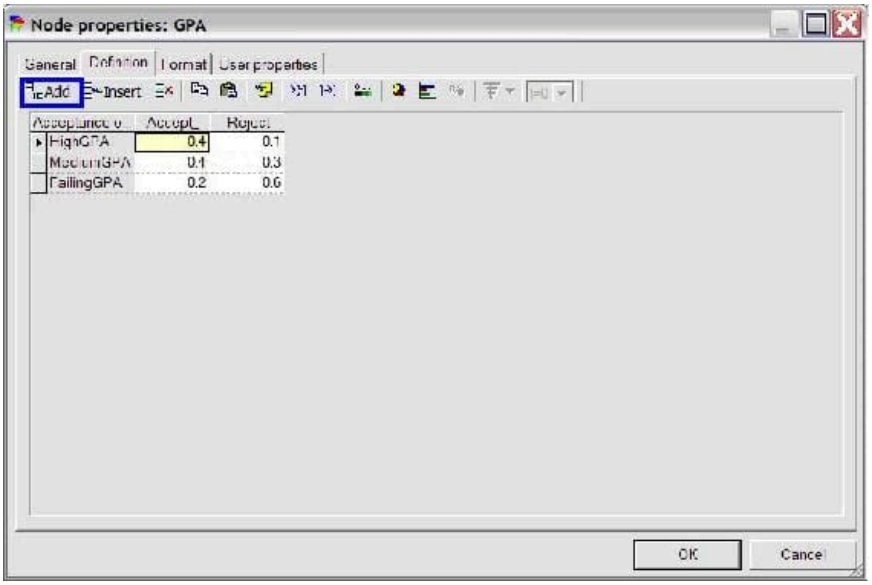
To define the probabilities of the node, click on the definition tab highlighted in blue as shown below. The names of the states can be edited by double clicking on the name fields (highlighted in red in the figure below) and the probabilities can be edited by clicking on the probability field (highlighted in yellow in the figure below).



In this example, “State0” is changed to “Accept” and “State1” is changed to “Reject”. The default probabilities are entered as 0.3 for Accept and 0.7 for Reject.

To create a second node for the variable GPA, simply click the change node again and place a second node under the “Acceptance” node. Define the name and identifier of this node as done with the “Acceptance” node. Then define the probabilities for the “GPA” node by first adding another outcome by clicking the “add outcome” button highlighted in blue in the figure below. Change state0, state1, and state2 to HighGPA, MediumGPA, and LowGPA respectively. To finish defining this node, fill in the probabilities listed in the problem statement (shown in the figure below) and press ok.

After the creation and definition of both nodes, connect these two nodes with an influence arc to represent that GPA affects how Polly accepts or rejects the invitations she receives. To do this, click on the Arc tool (found on the standard toolbar), , and click on the “Acceptance” node and drag the arrow to the new “GPA” node.



The Bayesian network describing this problem is now fully defined within the Genie program. To determine the probability of Polly accepting a date invitation from a guy if she knows the guy has a high GPA, first right click on the GPA node, scroll down to "Set Evidence", and select "HighGPA". This is shown in the figure below.

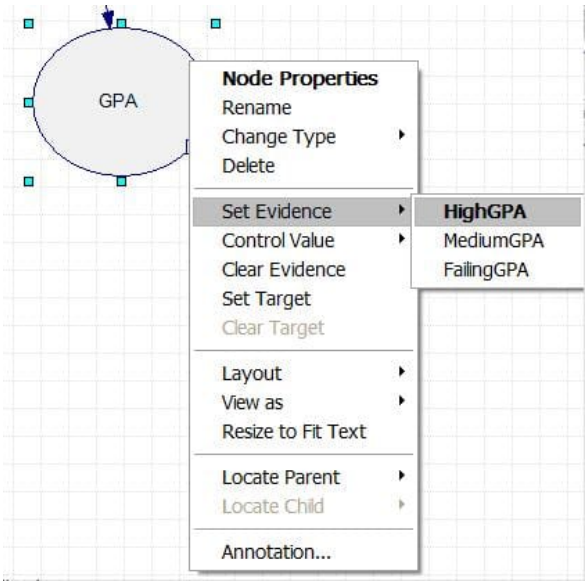
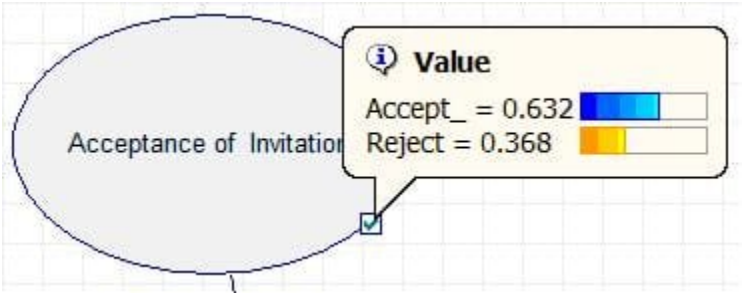
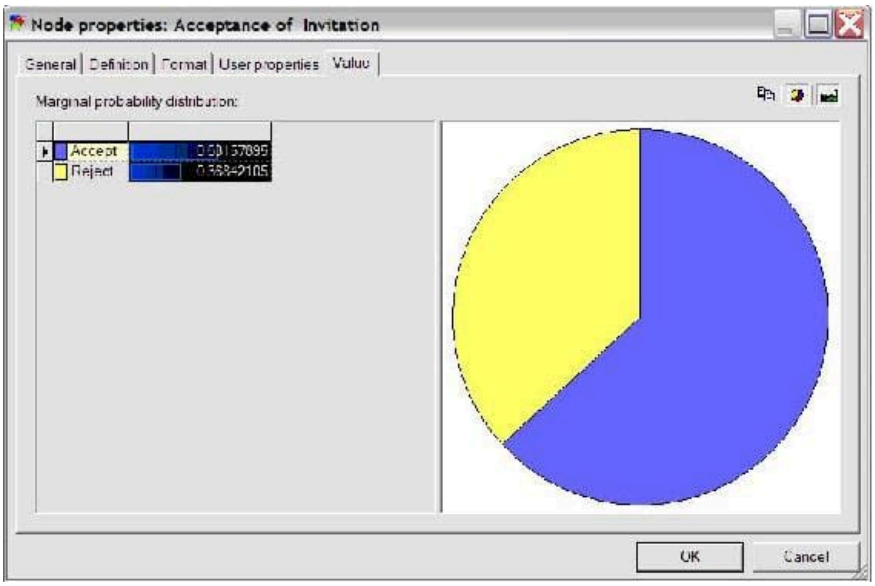


Figure below.



The results can also be accessed by double clicking the Acceptance node and selecting the value tab, shown in the figure below.



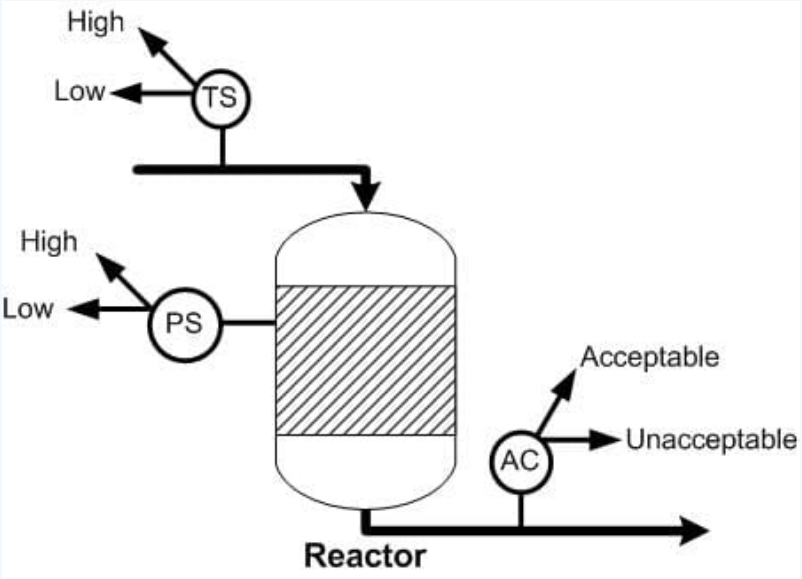
As can be seen in the figure above, Polly will accept the date invitation 63.16% of the time when she knows the guy inviting her has a high GPA. Other probabilities related to this example can be determined similarly.

The logic and procedure involved in this simple problem can be applied to complex systems with many interconnected nodes. Please see the Worked Out Examples sections for more examples of how to use Genie to analyze Bayesian networks.

- All informational materials used to create this tutorial come from GeNIe

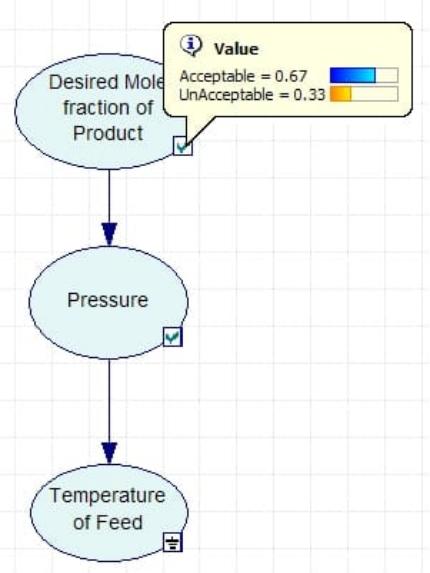
Example 13.6.1

For the reactor shown below, the probability that the effluent stream will contain the acceptable mole fraction of product is 0.5. For the same reactor, if the effluent stream contains the acceptable mole fraction, the probability that the pressure of the reactor is high is 0.7. If the effluent stream does not contain the acceptable fraction, the probability that the pressure is low is 0.85. If the pressure of the reactor is high, the probability that the temperature of the feed is high is 0.9 and if the pressure of the reactor is low, the probability that temperature of the feed is low is 0.75. Given that the temperature of the feed is low, what is the probability that the effluent stream contains the acceptable mole fraction of product?



Solution

The variables to be included in the Bayesian network are the acceptable mole fraction of product in the effluent stream, the pressure of the reactor, and the temperature of the reactor. The variables are connected in an acyclic network as shown in the image below. After the nodes were created and the incidence arcs were defined, the probability was calculated by updating the network and moving the pointer over the checkmark on the node. As can be seen in the figure, the probability of the effluent stream containing the acceptable mole fraction of product given that the feed temperature is low is 67%.



A GeNIe file containing the full solution of the problem is located here. In this file, the values entered into Genie for the probabilities and stages can be accessed by double-clicking on the nodes.

Example 13.6.2

The following example is complicated and contains many layers of information. It is intended to illustrate how many elements can be put together in a Bayesian Network and analyzed using Genie such that the problem is not impossibly complicated. Assume that detailed statistical experiments were conducted to find all the stated probabilities.

Consider a CSTR with a cooling jacket and an exothermic reaction with A --> B. The feed stream is pure A and the cooling stream in the jacket has a sufficient flow rate high such that the ambient temperature is constant.

The pumps transporting the reactor feed stream are old, and sometimes do not pump at their full capacity. 98% of the time the pumps work normally, but at other times the feed flow rate is slightly lower than normal. The preheating process for the feed stream is inefficient, and sometimes does not heat the reactants to 80°C. There is a 95% chance the feed stream is slightly cooler than desired. Finally, the pumps for the cooling water sometimes lose power, causing the ambient temperature for the reactor to climb higher than usual. There is a 1% chance of this power loss occurring.

The concentration of A in the reactor depends on the feed flow rate. If this flow rate is normal, then there is a 98% chance CA is normal, and a 1% chance each that it is slightly higher or lower than normal. If the flow rate is low, then there is a 40% chance that CA will be higher than normal and a 60% chance that CA will be normal.

The reactor temperature depends on the feed temperature and the ambient temperature. If both T0 and Ta are normal, there is a 98% chance T is normal, and a 1% chance each that it is slightly higher or lower than normal. If T0 is normal and Ta is high, there is a 90% chance T is normal and a 10% chance it is high. If Ta is normal and T0 is low, there is a 80% chance T is normal and a 20% T is low. Finally, if T0 is low and Ta is high, there is a 90% chance T is normal, a 8% chance T is low, and a 2% chance T is high.

The conversion depends on CA and T. If one of these variables is high and the other low, there is an equal chance that X will be low, normal, or high. If both CA and T are normal, there is a 98% chance X is normal, and a 1% chance each that it is slightly higher or lower than normal. If both CA and T are low, there is equal chance that X will be low or normal. If both CA and T are high, there is equal chance that X will be normal or high. If CA is normal and T is low, there is a 75% chance X is normal and a 25% chance X is low. If CA is normal and T is high, there is a 75% chance X is normal and a 25% chance X is high. Finally, if T is normal and CA is low or high, there is a 85% chance X is normal and a 15% chance that X is low or high, respectively.

Create a model in Genie for this system that can be used to determine probabilities of related events occuring.

Solution

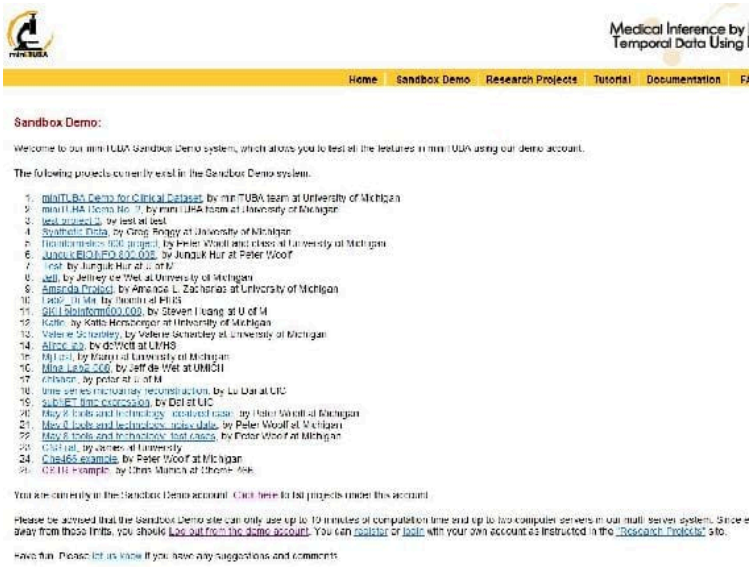
A GeNIe file containing the Bayesian Network of this problem is located here.

This model can be used to answer questions such as (1) if a composition sensor for B in the exit stream tells us that the conversion is slightly low, what is the probability that the feed temperature is normal and (2) if a temperature sensor in the reactor tells us that the reactor temperature is high what is the probability that the ambient temperature is normal?

miniTuba

MiniTuba is a program that allows the creation of Bayesain Networks thought time, with only the final data sets. This was created by Zuoshuang Xiang, Rebecca M. Minter, Xiaoming Bi, Peter Woolf, and Yongqun He, to analysis medical data, but can be used to create Bayesian Networks for any propose. In oder to use this program go to www.minituba.org/ and go to the Sandbox demo Tab. From here one can either start a new project or merly modify an old project. This wiki will talk though how to create a new project.

First Log in: to do this click the Log In link at the bottom of the page and enter demo@e.d.u as the user name and demo as the password this should bring up this screen:



and click My Projects and then Creat New Project.

The user needs to know how many variables (EX: Temperature, Pump speed, Yield, flowrate in ChemE), data sets (EX: Reactors, Tanks etc.) and the average number of time steps that will be analyzed.

In this example we will have a CSTR whose yield may be affected by the reactor Volume, the motor speed, the temperature or the concentration. This means that there are 5 variable. Lets say that we have 5 reactors running this experiment with differing either high or low Moter speed, temperature, and yield and either high medium or low volume and concentration. In this example Yield is only affected by the motor speed and Temperature with the following probability

	yeild	
Motor/Temp (high or low)	H	L
M+T+	0.96	0.04
M+T-	0.3	0.7
M-T+	0.6	0.4
M-T-	0	100

Enter the requested data in miniTuba and then click open project from the list. Then click "load/updata data"

this example will use the following data:

Media:CSTRExample.xls

To insert the data, simply copy and paste in the data to open box, then click LOAD DATA. In Order to run the analysis click "Start DBN botton"

Here you can slect which data series and variables you would like to analize. Now miniTuba gives several different options for the data. The data can be analyises and a child a parent or both. A Child means that the data isnt effected by the other datathat , Parent means the data doesnt affect the other data, and both means that the data can be affected and effect other variables. In our example, the motor speed is the only thing that we know is not affected by teh other variables, so we will select it as a parent.

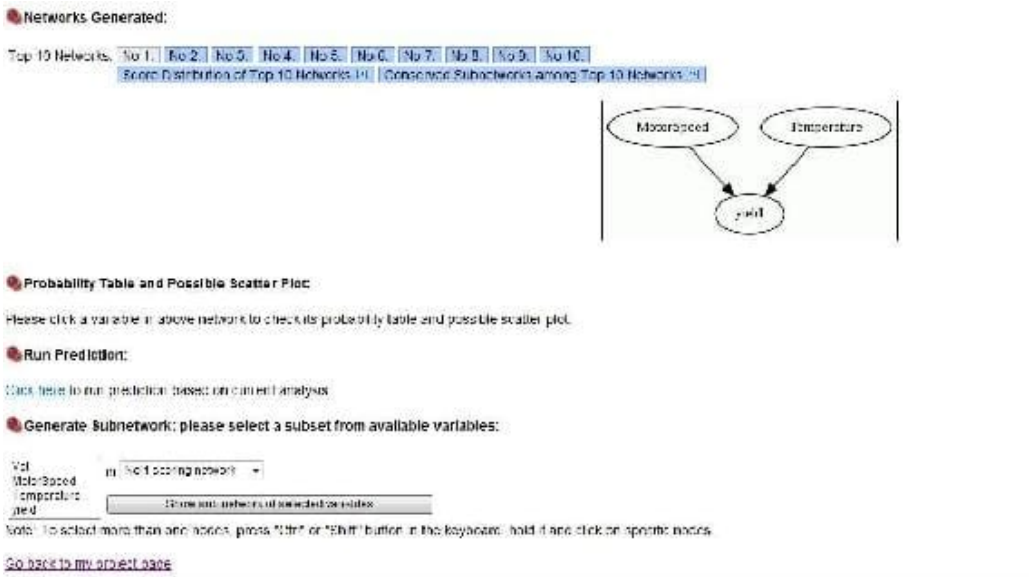
Discretization Policy tells MiniTuba how to bin the. In our case, as things are exactly high, med or low represented by 0,1,2 it is easy to desice how to descretize the data, but it is ok if the data isnt allready bin (ie in 0s 1s and 2s). Quantile spits the data into even chucks (ie if you select quantile 2, it will find the mean and then everything above that is in one bin and everything below that mean is in a different bin). Interval bins things that are in equal sized intervals around each other.

Select Natural fit for slipe fitting (this will allow you to have some data points missing ex: no volume reading in 1 data set) Lets select 2 intervals for moter speed, temperature and yeild adn then 3 intervals for volume and concentration.

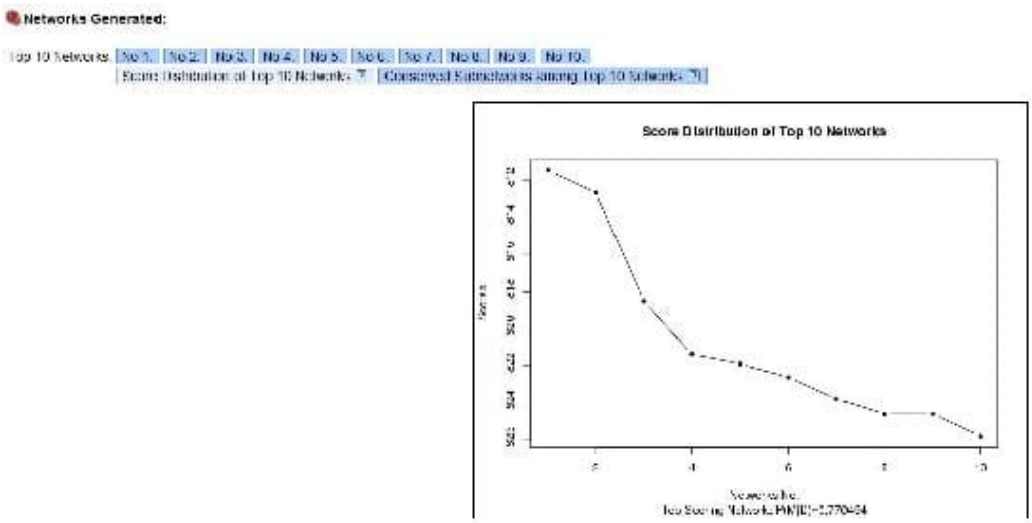
MiniTuba also give you the option to force it to come up with some relationships weather MiniTuba thinks they should or not. MiniTuba also allows you to choose the method used to solve this, we will select simulated anneling Number of Instances allows you to choose the number of computers used to solve the problem, we will just leave it at 1 In the Demo version, the max calculation time is 1 min.

When everything is done, click Run Bayesian Analysis. By clicking Check Progress teh following screen will show up. Click Updata every so often to check weather the solution has been computed.

This will output the following screen when it is done



As you can see, the yield depends on the motor speed and the temperature alone, just like we crated the module to show. Top Ten networks, shows the 10 most likely networks. Click Score Distribution of Top Ten Networks shows the relative probability of each network, and the probability that the number 1 network is the correct network out of the top 10. In our example the best network has a probablitity of 0.77 as seen below.



References

- Charniak, Eugene (1991). "Bayesian Networks without Tears", *AI Magazine*, 50-62.
- GeNIe Helpfile at GeNIe & SMILE
- Murphy, Kevin. (2002). "Dynamic Bayesian Networks."
- Xiang Z., Minter R., Bi X., Woolf P., and He Y.(2007). "miniTUBA: medical inference by network integration of temporal data using Bayesian analysis," *Bioinformatics*, v. 23, 2423-2432, 2007.

Contributors and Attributions

- Authors: Kevin Dahlberg, Paul Jantzen, Genevieve Lampinen, Albert Sawalha, and David Toronto
- Stewards: Ross Bredeweg, Jessica Morga, Ryan Sekol, Ryan Wong

This page titled 13.6: Learning and analyzing Bayesian networks with Genie is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

13.7: Occasionally Dishonest Casino? Markov Chains and Hidden Markov Models

Introduction

Basic probability can be used to predict simple events, such as the likelihood a tossed coin will land on heads rather than tails. Instinctively, we could say that there is a 50% chance that the coin will land on heads.

But let's say that we've just watched the same coin land on heads five times in a row, would we still say that there is a 50% chance of the sixth coin toss will results in heads? No – intuitively, we sense that there is a very low probability (much less than 50%) that the coin will land on heads six times in a row.

What causes us to rethink the probability of the coin landing on heads? The same coin is being tossed, but now we are taking into account the results of previous tosses. Basic probability can be used to predict an isolated event. But more advanced probability concepts must be used to assess the likelihood of a certain sequence of events occurring. This section provides an introduction to concepts such as Bayes rule, Markov chains, and hidden Markov models, which are more useful for predicting real world scenarios (which are rarely singular, isolated events).

Bayes' Rule

Bayes' rule is a foundational concept for finding the probability of related series of events. It relates the probability of event A conditional to event B and the probability of event B conditional to event A, which are often not the same probabilities.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Another interpretation of Bayes' rule: it's a method for updating or revising beliefs (i.e. probability of landing on heads) in light of new evidence (knowing that the coin landed on heads the last five tosses).

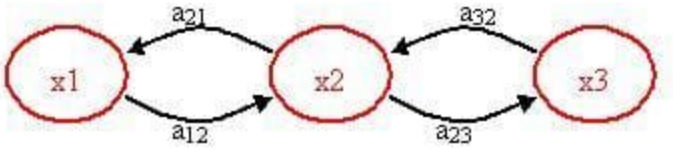
Markov Chains

A Markov chain is a particular way of modeling the probability of a series of events. A Markov chain is a sequence of random values whose probabilities at a time interval depend only upon the value of the number at the previous time. A Markov chain, named after Andrey Markov, is a stochastic process with the Markov property. This means that, given the present state, future states are independent of the past states. In other words, the description of the present state fully captures all the information that could influence the future evolution of the process, so that future states of the system will be reached through a probabilistic (determined from current data) process instead of a deterministic one (from past data).

At each point in time, the system could either change its state from the current state to a different state, or remain in the same state. This happens in accordance with the specified probability distribution. These changes of state are called transitions, and the probabilities associated with various state changes are called transition probabilities.

For example, let's say we want to know the probability of an elevator moving more than five floors up in a 30-floor building (including the lobby). This is an example of a Markov chain because the probability does not depend on whether the elevator was on the 6th floor at 7:00 am or on the 27th floor at 10:30 pm. Instead the probability is dependent on what floor the elevator is on right before it starts to move. More specifically, the fact that the elevator is on the top floor at t_1 results in a zero probability that the elevator will be five floors higher at t_2.

The figure below helps to visualize the concept of a Markov Model.



x_1, x_2, x_3 = states

a_{12}, a_{21}, a_{23}, a_{32} = transition probabilities

A transition probability is simply the probability of going from one state to another state. Please note that there may be other paths or transition probabilities for state x_1, say to state x_3 or back to itself. The sum of the transition probabilities must always equal to 1.

$$a_{1,i} = a_{11} + a_{12} + a_{13} + \ldots a_{1n}$$

$$\sum_i = 1$$

Transition Probability

The transition probabilities of each state for one system can be compiled into a transition matrix. The transition matrix for a system can represent a one-step transition or a multiple-step transition. A n-step transition matrix is simply a one-step transition matrix raised to the n-th power.

For state i to make a transition into state j, the transition probability P_{ij} must be greater than 0. so:

$$P_{i,j} \geq 0, i, j \geq 0$$

$$\sum_{j=0} P_{ij} = 1$$

with i = 0, 1, ...

Let **P** be the matrix of a one-step transition matrix,

P= $\begin{Vmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \cdots \\ P_{1,0} & P_{1,1} & P_{1,2} & \cdots \\ \vdots & \ddots & \vdots & \\ P_{i,0} & P_{i,1} & P_{i,2} & \cdots \\ \vdots & \vdots & \vdots & \end{Vmatrix}$

$P_{i,1}$ would denote the probability of going from state i to 1

Example 13.7.1: Will It Rain Tomorrow?

Suppose that whether or not it rains today depends on previous weather conditions through the last two days. Specifically, suppose that if it has rained for the past two days, then it will rain tomorrow with probability 0.7; if it rained today but not yesterday, then it will rain tomorrow with probability 0.5; if it rained yesterday but not today, then it will rain tomorrow with probability 0.4; if it has not rained in the past two days, then it will rain tomorrow with probability 0.2.

Solution

The transition matrix for the above system would be:

- state 0: if it rained both today and yesterday
- state 1: if it rained today but not yesterday
- state 2: if it rained yesterday but not today
- state 3: if it did not rain either yesterday or today

Missing \left or extra \right

Applications of Markov Chains

Since this concept is quite theoretical, examples of applications of this theory are necessary to explain the power this theory has. Although these following applications are not chemical control related, they have merit in explaining the diversity of operations in which Markov chains can be used. Hopefully, by seeing these examples, the reader will be able to use Markov chains in new, creative approaches to solve chemical controls problems.

Queuing Theory

Markov chains can be applied to model various elements in something called queuing theory. Queuing theory is simply the mathematical study of waiting lines (or queues).

This theory allows for the analysis of elements arriving at the back of the line, waiting in the line, and being processed (or served) at the front of the line. This method uses the variables including the average waiting time in the queue or the system, the expected number waiting or receiving service and the probability of encountering the system in certain states, such as empty, full, having an available server or having to wait a certain time to be served. These variables are all unique to the system at hand. A link that explains this theory in more detail can be found below:

Queueing Theory

Although Queuing Theory is generally used in customer service, logistics and operations applications, it can be tailored to chemical processing. For instance, it is essential to model how long a particular product will be in an assembly line before it makes it to the desired end process. Thus, this theory is robust for all kinds of uses.

Example 13.7.1: Queuing Problem

On a sunny Tuesday morning in the Computer Engineering Building, a group of Chemical Engineering students are finishing their engineering laboratory reports and have stayed up all night writing ambivalent fluid mechanics questions, they need some prestigious coffee for refreshment instead of Red Bull. Coffee shop called "Dreamcaster's Cafè" was opened by a University of Michigan Chemical Engineering graduate. Business is going very well for "Dreamcaster's Cafè" and there are many people, *m* people, waiting in line in front of the Chemical Engineering students. The service time for the first person in line (not necessarily a Chemical Engineering student) is an exponential distribution with an average waiting time of $\frac{1}{\mu}$. Since the Chemical Engineering students have to return to the Computer Engineering Building to finish their report, they have a finite patience waiting in line and their patience is an exponential distribution with an average of $\frac{1}{\theta}$. The group of Chemical Engineering students does not care how long the service time is. Find the probability that the *ith* Chemical Engineering student to wait long enough to get coffee.

Solution

Add text here.

Google's Page Ranking System

The PageRank of a webpage as used by Google is defined by a Markov chain. PageRank is the probability to be at page i in the following Markov chain on all (known) webpages. If N is the number of known webpages, and a page i has k_i links then it has transition probability

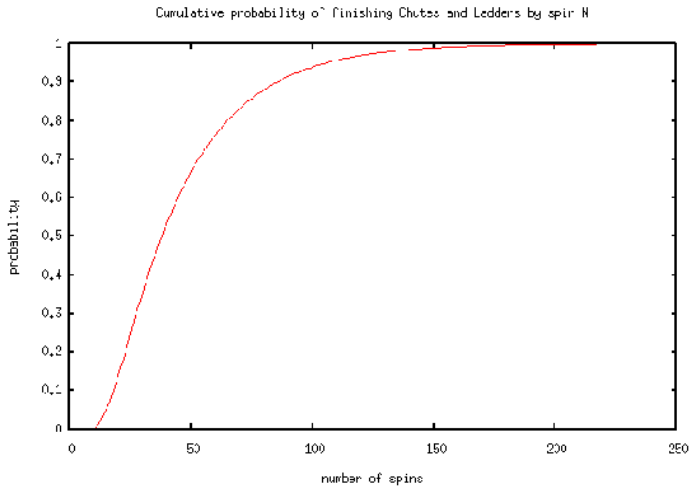
$$\frac{(1-q)}{ki} + \frac{q}{N}$$

for all pages that are linked to and q/N for all pages that are not linked. The constant q is around 0.15.

Children's Games

Markov chains are used in the children's games "Chutes and Ladders," "Candy Land" and "Hi Ho! Cherry-O!" These games utilize the Markov concept that each player starts in a given position on the game board at the beginning of each turn. From there, the player has fixed odds of moving to other positions.

To go into further detail, any version of Snakes and Ladders can be represented exactly as a Markov chain, since from any square the odds of moving to any other square are fixed and independent of any previous game history. The Milton Bradley version of Chutes and Ladders has 100 squares, with 19 chutes and ladders. A player will need an average of 45.6 spins to move from the starting point, which is off the board, to square 100.



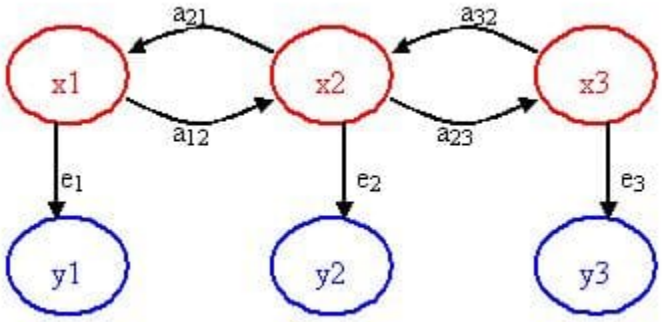
One last game of interest is Monopoly. This can be modeled using Markov chains. Check out the following link to see a simulation of the probability of landing on a particular property in the Monopoly game:

Monopoly Simulation

Hidden Markov Models

Building further on the concept of a Markov chain is the hidden Markov model (HMM). A hidden Markov model is a statistical model where the system being modeled is assumed to be a Markov process with unknown parameters or casual events. HMMs are primarily helpful in determining the hidden parameters from the observable parameters.

Again, the figure below may help visualize the Hidden Markov Model concept.



x_1, x_2, x_3 = hidden/unknown states

a_{12}, a_{21}, a_{23}, a_{32} = transition probabilities

y_1, y_2, y_3 = observable states

e_1, e_2, e_3 = emission probabilities

The key point of the figure above is that only the blue circles are seen, but you suspect or know that these *observable states* are directly related or dependent on some *hidden state*. These hidden states (the red circles) are what actually dictates the outcome of the observable states. The challenge is to figure out the hidden states, the emission probabilities and transition probabilities.

For example, say you have a data acquisition program that continuously records the outlet temperature of your product stream which you are heating up through a heat exchanger. The heat exchanger provides a constant supply of heat. You observe that the temperature has some step disturbances and is not always constant. The temperature is the observable parameter, but the cause of its deviations/states is actually the stream's flowrate. When the flowrate fluctuates, the outlet temperature also fluctuates since the heat exchanger is providing the same amount of heat to the stream. So, the stream flowrate is the hidden parameter (since you do not have access to this measured variable) and the probability that the flowrates switches from say 100 to 150 to 300 L/h are the transition probabilities. The emission probability is the probability that the observed outlet temperature is directly caused by the flowrate. This emission probability is not necessarily 1 since temperature variations could also be due to noise, etc.

Another common scenario used to teach the concept of a hidden Markov model is the “Occasionally Dishonest Casino”. If a casino uses a fair die, each number has a 1/6 probability of being landed on. This type of probability that represents a known parameter is known as emission probability in HMMs. Sometimes a casino can tip the odds by using a loaded die, where one number is favored over the other five numbers and therefore the side that is favored has a probability that is higher than 1/6.

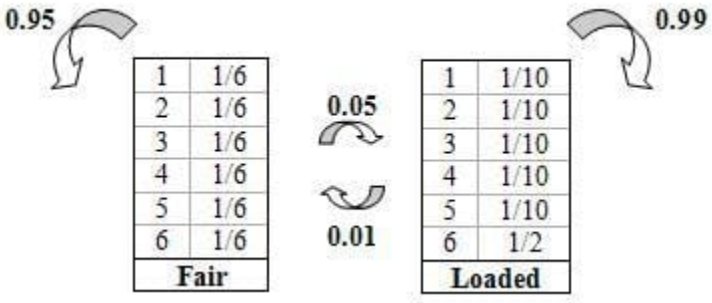
But how does one know that the casino is being dishonest by using a loaded die? Pretend that the die is loaded in such a way that it favors the number 6. It would be difficult to differentiate between a fair die and loaded die after watching only one roll. You may be able to get a better idea after watching a few tosses and see how many times the die landed on 6. Let’s say that you saw the following sequence of die tosses:

465136

It is still difficult to say with certainty whether the die is loaded or not. The above sequence is a feasible for both a fair die and a loaded die. In this particular case, the above numbers that represent what was rolled are the observable parameters. The hidden parameter is the type of die used just because we do not know which type produced the above sequence of numbers.

Instead of relying on a sneaking suspicion that the casino is being dishonest, one can use a hidden Markov model to prove that a loaded die is being used occasionally. Please note that if a loaded die was used all the time it would be more blatantly obvious. In order to get away with this “slight” unfairness, the casino will switch the fair die with a loaded one every so often. Let’s say that if a fair die is being used there is a 5% chance that it will be switched to a loaded die and a 95% chance that it will remain fair. These are known as transition probabilities because they represent the likelihood of a causal event changing to another causal event.

Below is a picture representing the transition probabilities (probability of staying or changing a type of die) as well as the emission probabilities (probability of landing on a number).



Using the basic diagram above that show the emission and transition properties, conditioned statements in Excel can be used to model the probability that a loaded die is being used. Please refer to example two for the hidden Markov model Excel sheet.

HHMs can be modeled using Matlab. Follow the link to the site which contains information in installing the HHM toolbox in Matlab:

HHMs in Matlab

Example 13.7.1: "What should I wear?"

The weather affects important issues such as, choice of clothing, shoes, washing your hair, and so many more. Considering the bizarre weather in Ann Arbor, you wish to be able to calculate the probability that it will rain given the data of the last three consecutive days. Applying your chemical engineering skills and empowered by the knowledge from this Wiki article, you sit down and decide to be the next Weather-Person.

You know the following information since you have been an avid fan of the Weather Channel since age 5.

Weather	P(rain|weather)	P(blizzard|weather)
Sunny	0.15	0.05
Blizzard	0.25	0.15
Rainy	0.6	0.1

You then record the weather for the past 3 days.

Day	Weather
1	Rain
2	Rain
3	Blizzard
4	?

Solution

Since the last day was a blizzard you use a matrix expressing a 100% chance in the past, and multiply this by the probabilities of each weather type.

$$\begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 8 & 15 & 05 \\ .6 & 25 & .15 \\ 3 & .6 & .1 \end{bmatrix} = \begin{bmatrix} .6 & .25 & .15 \end{bmatrix}$$

This means that there is a 60% chance of sun, a 25% chance of rain and a 15% chance of a blizzard. This is because a Markov process works based on the probability of the immediately previous event. Thus, you decide that is a low enough risk and happily decide to wear your best new outfit AND wash your hair.

Example 13.7.2: "What should I wear for the weekend?"

Turns out it does rain the next day and you get your favorite clothes soaked. However, you are leaving for a weekend with your family and must decide what to wear for the next two days. What is the prediction for the next two days?

Solution

Predicting for the first day is like the previous example substituting rain in for a blizzard.

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .8 & .15 & .05 \\ .6 & .25 & .15 \\ .3 & .6 & .1 \end{bmatrix} = \begin{bmatrix} .3 & .6 & .1 \end{bmatrix}$$

In order to predict for the second day we must use our previous data to predict how the weather will act. Since we predict two days ahead we multiply the probabilities twice, effectively squaring them.

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 8 & 15 & 05 \\ 6 & 25 & .15 \\ 3 & 6 & .1 \end{bmatrix}^2 = \begin{bmatrix} .63 & .255 & .115 \end{bmatrix}$$

Alternatively, we can take the prediction from the previous day and multiply that by the probability of weather.

$$\begin{matrix} \text{L. } 3 & .6 & 1\rfloor & .0 & .15 & .05 \\ .6 & .25 & .15 & = .63 & .255 & .115 \end{matrix}$$

Both ways give you the same answer. Therefore, you decide to pack for both sunny and rainy weather.

Example 13.7.3: OSEH Example

Now, back to REAL chemical engineering applications, the probability theories described above can be used to check if a piece of equipment is working. For example, you are an OSEH inspector and you have to analyze a set of data recording the pollutant levels of a plant. If they use the scrubber they have, the probability of pollutant levels being high is low. However, if the scrubber is not used, the probability of pollutant levels being high increases. You have gotten some inside information that the operators use the scrubbers only 70% off the time. Your task is to create a model based on the data recordings and using probability theories to *predict* WHEN they switch their scrubbers ON and OFF, so that you can catch them red-handed in the act!

There are 6 possible pollutant levels: none, low, med-low, med, med-high, and high. If the level = high, then this is above the allowable limit. Every other level is considered allowable. If the scrubber is ON, there is equal chance that any of the states will occur. If the scrubber is OFF, there is a 50% chance that a *high* pollutant level will be recorded.

Given the following data:

P(E_on) = 0.7	probability of equipment is operating
P(E_off) = 0.3	probability of equipment is NOT operating
six possible states of pollutant quantity: none, low, med-low, med, med-high, high,	
[illegible] probability of high, [illegible] allowed limit by OSEH	
Readings are taken every hour by a concentration analyzer	
Scrubber on	
P(<limit| E_on)	probability of pollutant quantity being less than the allowable limit i.e. none, low, med-low, med, med-high
P(<limit| E_on)	(5/6)
P(=limit| E_on)	probability of pollutant quantity being equal or more than the allowable limit i.e. high
P(=limit| E_on)	(1/6)
Scrubber off	
P(<limit| E_off)	0.5
P(=limit| E_off)	0.5

Solution

Formulas used are listed below:

(1) $$P\left(Data|E_{on}\right)=P\left(<limit|E_{on}\right)^{n-1}P\left(limit|E_{on}\right)^{n}=\left(\frac{5}{6}\right)^{n-1}\left(\frac{1}{6}\right)^{n}$$

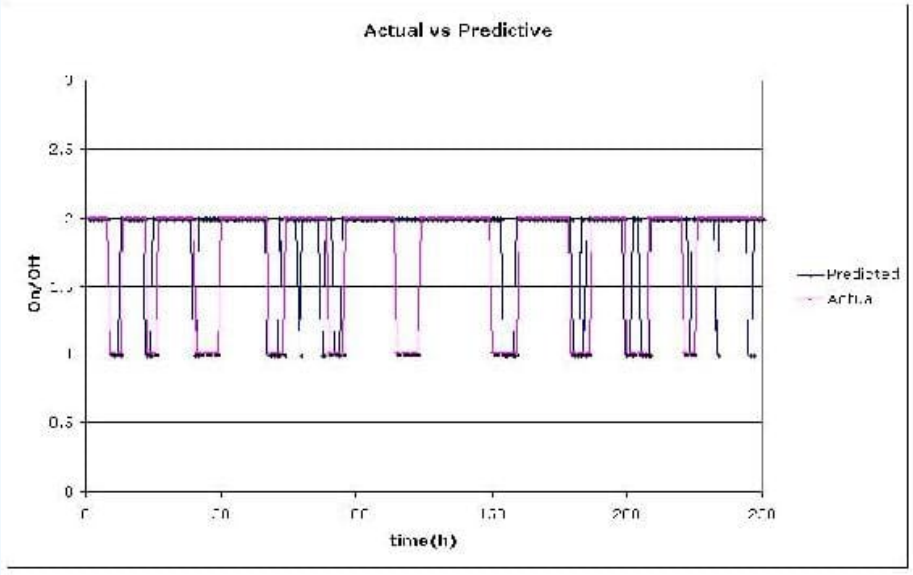
(2) $$P\left(Data|E_{off}\right)=P\left(<limit|E_{off}\right)^{n-1}P\left(limit|E_{off}\right)^{n}=\left(\frac{1}{2}\right)^{n-1}\left(\frac{1}{2}\right)^{n}$$

(3) $$P\left(Data\right)=P\left(Data|E_{on}\right)*P\left(E_{on}\right)+P\left(Data|E_{off}\right)*P\left(E_{off}\right)$$

(4) $$P\left(E_{on}|Data\right)=\frac{P\left(Data|E_{on}\right)*P\left(E_{on}\right)}{P\left(Data\right)}$$

Please refer to the attached Excel File for the predictive model. OSEHExample

From our results in the Excel file, we graphed the predicted vs actual times when the scrubber was turned off. The value of "1" was chosen arbitrarily to represent when the scrubber was OFF and "2" when it was "ON".



Multiple Choice Question 1

What is the key property of a Markov chain?

A) Probability of each symbol depends on FUTURE symbols.

B) Probability of each symbol depends on ALL preceding symbols.

C) Probability of each symbol depends on NONE of the preceding symbols.

D) Probability of each symbol depends only on the value of the PREVIOUS symbol.

Multiple Choice Question 2

What kind of parameters does emission probability represent?

A) unknown parameters

B) Markov parameters

C) known parameters

D) Bayes' parameters

References

- Woolf P., Keating A., Burge C., and Michael Y. "Statistics and Probability Primer for Computational Biologists". Massachusetts Institute of Technology, BE 490/ Bio7.91, Spring 2004
- Smith W. and Gonic L. "Cartoon Guide to Statistics". Harper Perennial, 1993.
- Kemeny J., Snell J. "Finite Markov Chains". Springer, 1976. ISBN 0387901922
- en.Wikipedia.org/wiki/Snakes_and_Ladders
- Page, Lawrence; Brin, Sergey; Motwani, Rajeev and Winograd, Terry (1999). "The PageRank citation ranking: Bringing order to the Web"

- Use the Harvard Referencing style for references in the document.
- For more information on when to reference see the following Wikipedia entry.
- Ross, Sheldon M. "Introduction to Probability Models". Academic Press, 2007.

Contributors

Authors: Nicole Blan, Jessica Nunn, Pamela Roxas, Cynthia Sequerah
Stewards: Kyle Goszyk, So Hyun Ahn, Sam Seo, Mike Peters

This page titled 13.7: Occasionally Dishonest Casino? Markov Chains and Hidden Markov Models is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

13.8: Continuous Distributions- normal and exponential

Authors: Navin Raj Bora, Dallas Burkholder, Nina Mohan, Sarah Tschirhart

Introduction

Random variables whose spaces are not composed of a countable number of points but are intervals or a union of intervals are said to be of the continuous type. Continuous distributions are probability models used to describe variables that do not occur in discrete intervals, or when a sample size is too large to treat each individual event in a discrete manner (please see Discrete Distributions for more details on discrete distributions). The main difference between continuous and discrete distributions is that continuous distributions deal with a sample size so large that its random variable values are treated on a continuum (from negative infinity to positive infinity), while discrete distributions deal with smaller sample populations and thus cannot be treated as if they are on a continuum. This leads to a difference in the methods used to analyze these two types of distributions: continuous and discrete distributions is continuous distributions are analyzed using calculus, while discrete distributions are analyzed using arithmetic. There are many different types of continuous distributions including some such as Beta, Cauchy, Log, Pareto, and Weibull. In this wiki, though, we will only cover the two most relevant types of continuous distributions for chemical engineers: Normal (Gaussian) distributions and Exponential distributions.

In chemical engineering, analysis of continuous distributions is used in a number of applications. For example in error analysis, given a set of data or distribution function, it is possible to estimate the probability that a measurement (temperature, pressure, flow rate) will fall within a desired range, and hence determine how reliable an instrument or piece of equipment is. Also, one can calibrate an instrument (eg. temperature sensor) from the manufacturer on a regular basis and use a distribution function to see of the variance in the instruments' measurements increases or decreases over time.

Normal Distributions

What is a Gaussian (normal) distribution curve?

A Gaussian distribution can be used to model the error in a system where the error is caused by relatively small and unrelated events.

This distribution is a curve which is symmetric about the mean (i.e. a Bell shaped curve) and has a range measured by standard deviations above and below the mean of the data set (please see Basic Statistics for a further discussion on these statistical parameters). To better explain, consider that a certain percentage of all data points will fall within one standard deviation of the mean. Likewise, more data points will fall within two standard deviations of the mean, and so on. However, under this model it would require an infinite range to capture ALL the data points thus presenting a minor difficulty in this appraoch.

The figure below is a possible Gaussian distribution, where the mean (μ) is 10 and standard deviation (σ) is 2. F(x) is the number of times a certain value of x occurs in the population. The mean is simply the numerical average of all the samples in the population, and the standard deviation is the measure of how far from the mean the samples tend to deviate. The following sections explain how and why a normal distribution curve is used in control and what it signifies about sets of data.

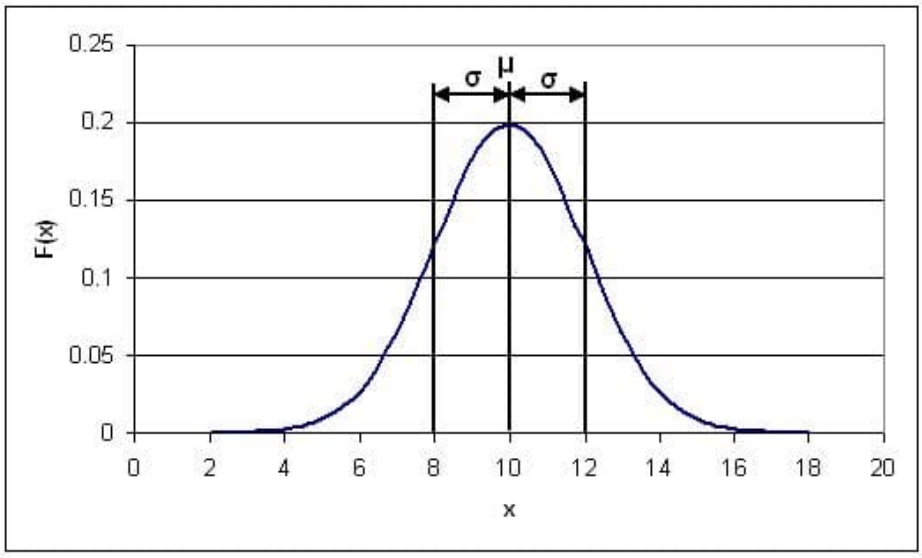


Figure 1: Generic Normal Distribution curve

The Probability Density Function (PDF) for a normal distribution

As was mentioned in the Introduction section, distribution curves can be used to determine the probability, P(x), of a certain event occurring. When this is done, the distribution curve is known as a Probability Density Function (**PDF**). In the figure shown above, the x-axis represents the range of possible events (eg. the range of ages in a sample population or the magnitude of noise generated by a temperature sensor). The y-axis represents the number of times a certain x value occurs in a population. The PDF can be described mathematically as follows:

In some cases, it might not be necessary to know the probability of just one event occurring. Rather, you may want to know the probability of a range of events (eg. what is the probability that the noise generated by my temperature sensor will fall in the range of 5-10 Hz?). When this happens, you must integrate the above PDF over the desired range, in the following manner:

where k_1 and k_2 are the limits of your desired range. This integral results in the following expression:

$$P(k_1 < x < k_2) = \frac{1}{2}(Erf[\frac{k_2 - \mu}{\sigma\sqrt{2}}] - Erf[\frac{k_1 - \mu}{\sigma\sqrt{2}}])$$

The *Erf* function can be found in most scientific calculators and can also be calculated using tables of *Erf[]* values. For example, its use here parallels what we learned in ChE 342, Heat and Mass Transfer. Determine the value inside the brackets of the erf function through simple arithmetic, then take this value and find the corresponding Erf number from a table. Finally use this value in the calculation to determine the probability of a certain point, x, falling within a range bound from k1 to k2.

Sample Calculation

Given a data set with an average of 20 and a standard deviation of 2, what is the probability that a randomly selected data point will fall between 20 and 23?

Solution 1

To solve simply substitue the values into the equation above. This yeilds the following equation:

$$P(20 < x < 23) = \frac{1}{2}(Erf[\frac{23-20}{2\sqrt{2}}] - Erf[\frac{20-20}{2\sqrt{2}}])$$

$$P(20 < x < 23) = \frac{1}{2}(Erf[1.061] - Erf[0])$$

These Erf values must be looked up in a table and substituted into the equation. Doing this yeilds

$$P(20 < x < 23) = \frac{1}{2}(0.933255 - 0.5) = 0.433255$$

Thus there is a 43.3255% chance of randomly selecting a number from the data set with a value between 20 and 23.

Graphically speaking, the PDF is just the area under the normal distribution curve between k_1 and k_2. So, for the distribution shown above, the PDF for 8

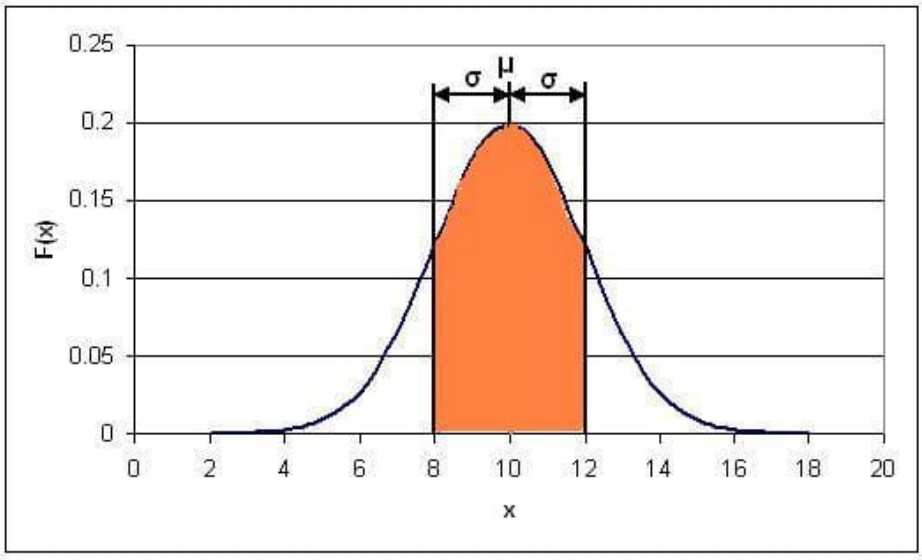


Figure 2: PDF for 8

Solution 2

Alternatively, rather than using the error function, Mathematica's built-in probability density function can be used to solve this problem. This probability density function can be applied to the normal distribution using the syntax shown below.

NIntegrate[PDF[NormalDistribution[μ,σ],x],{x,x1,x2}]

In this syntax, μ represents the mean of the distribution, σ represents the standard deviation of the distribution, and x1 and x2 represent the limits of the range.

This function can be applied to this problem as shown below.

```
In[1]:= NIntegrate[PDF[NormalDistribution[20, 2], x], {x, 20, 23}]
Out[1]= 0.433193
```

This shows that the probability of a randomly selected data point falling between 20 and 23 is 0.4332, or 43.32%. As expected, this value calculated using the built-in probability density function in Mathematica matches the value calculated from using the error function. Mathematica provides a faster solution to this problem.

An important point to note about the PDF is that when it is integrated from negative infinity to positive infinity, the integral will always be equal to one, regardless of the mean and standard deviation of the data set. This implies that there is a 100% chance that your random variable x will fall between negative infinity and positive infinity. Likewise, the integral between negative infinity and the mean is 0.5, or there is a 50% chance of finding a value in this region due to the symmetric nature of the distribution.

The Cumulative Density Function (CDF) for a normal distribution

To transition our arguement to a more holistic perspective regarding the probability density function for a normal distribution, we present the cumulative density function, which represents the integral (area under the curve) of the PDF from negative infinity to some given value on the y- axis of the graph incrementally accoring to the x axis, which remains the same as before. Because of this type of definition, we may circumvent the rigorous error function analysis presented above by simply subtracting one CDF points from another. For example, if engineers desire to determine the probability of a certain value of x falling within the range defined by k1 to k2 and posses a chart feauturing data of the relevant CDF, they may simply find CDF(k2)- CDF(k1) to find the relevant probability.

The Cumulative Density Function (**CDF**) is simply the probability of the random variable, x, falling at or below a given value. For example, this type of function would be used if you wanted to know the probability of your temperature sensor noise being less than or equal to 5 Hz. The CDF for a normal distribution is described using the following expression:

$$P(-\infty < x < k) = \frac{1}{2}(Erf[\frac{k-\mu}{\sigma\sqrt{2}}] + 1)$$

where k is the maximum allowable value for x.

The main difference between the PDF and CDF is that the PDF gives the probability of your variable x falling within a definite range, where the CDF gives the probability of your variable x falling at or below a certain limit, k. The following figure is the CDF for a normal distribution. You'll notice that as x approaches infinity, the CDF approaches 1. This implies that the probability of x falling between negative and positive infinity is equal to 1.

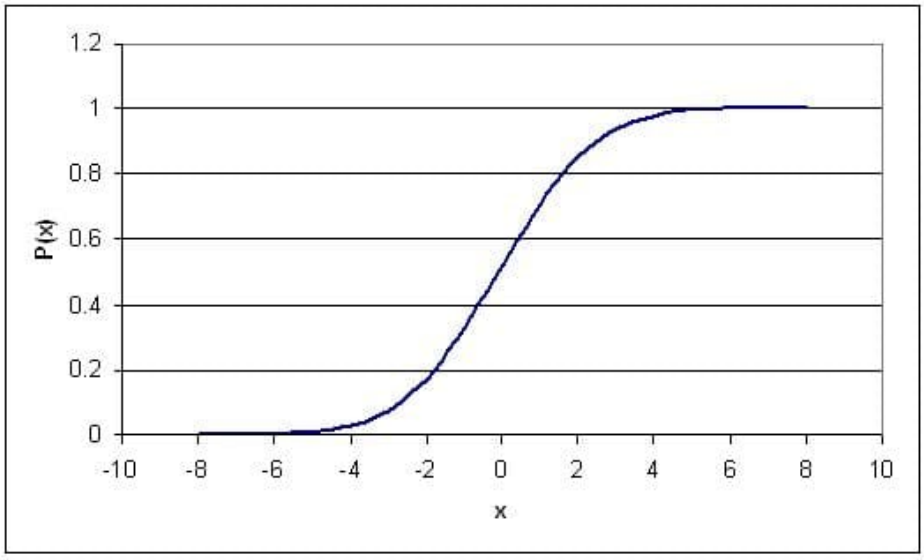


Figure 3: CDF plot for a Normal Distribution

Standard Normal Distribution

A special type of probability distribution curve is called the **Standard Normal Distribution**, which has a mean (μ) equal to 0 and a standard deviation (σ) equal to 1.

This simplified model of distribution typically assists engineers, statisticians, business strategists, economists, and other

interested professionals to model process conditions, and to associate the attention and time needed to adress particular issues (i.e. higher probability for a failed condition necessitate additional attention, etc.). Also, our grades in many of the courses here at the U of M, both in and outside of the college of engineering, are based either strictly or loosely off of this type of distribution. For example, the B range typically falls within +/- of one standard deviation.

The following figure is a Standard Normal Distribution curve:

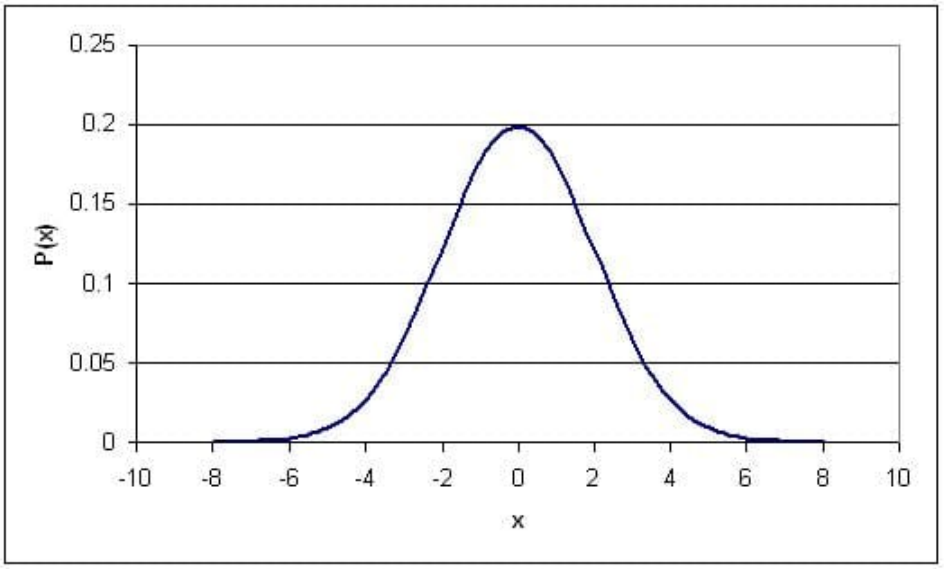


Figure 4: Standard Normal Distribution Curve

The benefit of the standard normal distribution is it can be used in place of the *Erf[]* function, if you do not have access to a scientific calculator or *Erf[]* tables. To use the Standard Normal Distribution curve, the following procedure must be followed:

1.Perform a z-transform. This is a transformation which essentially normalizes any normal distribution into a standard normal distribution. It is done using the following relationship:

$$z = \frac{x - \mu}{\sigma}$$

Mathematically speaking, the z transform normalizes the data by changing all the raw data points into data points that dictate how many standard deviations they fall away from the mean. So, regardless of the magnitude of the raw data points, the standardization allows multiple sets of data to be compared to each other.

2.Use a standard normal table to find the p-value. A standard normal table has values for z and corresponding values for F(x), where F(x) is known as the p-value and is just the cumulative probability of getting a particular x value (like a CDF). A standard normal table may look like the following table (it should be noted that if you are dealing with a Standard Normal Table that contains only positive z values, the following property can be used to convert your negative z values into positive ones: F(-z)=1-F(z)):

Table 1: Standard Normal Table

z	F(x)	z	F(x)	z	F(x)
-2.5	0.006	-1	0.159	0.5	0.691
-2.4	0.008	-0.9	0.184	0.6	0.726
-2.3	0.011	-0.8	0.212	0.7	0.758
-2.2	0.014	-0.7	0.242	0.8	0.788
-2.1	0.018	-0.6	0.274	0.9	0.816
-2	0.023	-0.5	0.309	1	0.841
-1.9	0.029	-0.4	0.345	1.1	0.864
-1.8	0.036	-0.3	0.382	1.2	0.885
-1.7	0.045	-0.2	0.421	1.3	0.903
-1.6	0.055	-0.1	0.46	1.4	0.919
-1.5	0.067	0	0.5	1.5	0.933
-1.4	0.081	0.1	0.54	1.6	0.945
-1.3	0.097	0.2	0.579	1.7	0.955
-1.2	0.115	0.3	0.618	1.8	0.964
-1.1	0.136	0.4	0.655	1.9	0.971

A more detailed standard normal table can be found here (Note: This table is the same table used in the 'Basic Statistics' wiki).

3. What if I want the probability of z falling between a range of x=a and x=b?. First, find your two z values that correspond to a and b. So these would be $z_a = \frac{a-\mu}{\sigma}$ and $z_b = \frac{b-\mu}{\sigma}$, respectively. The probability of x falling in between a and b is just: F(zb) – F(za), where F(zb) and F(za) are found from the standard normal tables.

Sample Calculation

Lets take the same scenario as used above, where you have a data set with an average of 20 and standard deviation of 3 and calculate the probability of a randomly selected data point being between 20 and 23.

Solution

To do this simply subtract the two Z scores:

$$z_a = \frac{23-20}{2} = 1.5$$

$$z_= \frac{0-20}{2} = 1$$

These Z scores correspond to probabilities of 0.933 and 0.5 respectively. Their difference, 0.433, is the probability of a randomly selected value falling between the two. Notice that this is almost identical to the answer obtained using the Erf method. Note that obtaining this answer also required much less effort. This is the main advantage of using Z scores.

Properties of a Normal Distribution

There are several properties for normal distributions that become useful in transformations.

1 If *X* is a normal with mean μ and σ^2 often noted $X \sim N(\mu, \sigma^2).$ then the transform of a data set to the form of *aX* + *b* follows a $\sim N(a*\mu+b, a^2*\sigma^2),$.

2 A normal distribution can be used to approximate a binomial distribution (n trials with probability p of success) with parameters μ = *np* and σ^2 = *np*(1 – *p*). This is derived using the limiting results of the central limit theorem.

3 A normal distribution can be used to approximate a Poisson distribution (parameter λ) with parameters $\mu = \sigma^2 = \lambda$. This is derived using the central limit theorem.

4 The a sum of random variables following the same normal distribution ~$N(n\mu,n\sigma^2)$. This result shows how the sample mean $\sim N(\mu, \frac{\sigma^2}{n})$.

5 The square of a standard normal variable is a chi-squared variable with one degree of freedom.

6 Combining 4 and 5 yields the sum of n squared standard normal variables is a chi-squared variable with n degrees of freedom. This has application for chi-square testing as seen other sections of this text.

Exponential Distribution

The exponential distribution can be thought of as a continuous version of the geometric distribution without any memory. It is often used to model the time for a process to occur at a constant average rate. Events that occur with a known probability for a given x value build the theory developed previously (i.e. F(x) vs. x, F(x) in this case is different than before), where x may indicate distance between mutations on a DNA strand., for one example application, rate of instrument failure, or time required for a radioactive particle to decay (the decay RATE) for the below applications.

However, do remember that the assumption of a constant rate rarely holds as valid in actuality. The rate of incoming phone calls differs according to the time of day. But if we focus on a time interval during which the rate is roughly constant, such as from 2 to 4 p.m. during work days, the exponential distribution can be used as a good approximate model for the time until the next phone call arrives.

One can implement the exponential distribution function into Mathematica using the command: ExponentialDistribution[lambda].

Or, for a more grass-roots understanding of the function reference the following website, detailing the number of sharks seen in area one square mile in different one hour time periods. Use the "Fish" button to run the applet. To change the parameter lambda, type in the value and hit the "Clear" button.

www.math.csusb.edu/faculty/stanton/m262/poisson_distribution/Poisson_old.html

The Probability Density Function (PDF)

A few notes are worth mentioning when differentiating the PDF from the two-parameter Exponential Distribution function. As λ is decreased in value, the distribution is stretched out to the right, and as λ is increased, the distribution is pushed toward the origin. This distribution has no shape parameter as it has only one shape, i.e. the exponential, and the only parameter it has is the failure rate, λ. The distribution starts at T = 0 at the level of f(T = 0) = λ and decreases thereafter exponentially and monotonically as T increases, and is convex. The PDF for an exponential distribution is given in the form below, where λ is the rate parameter and x is random variable:

$$P(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

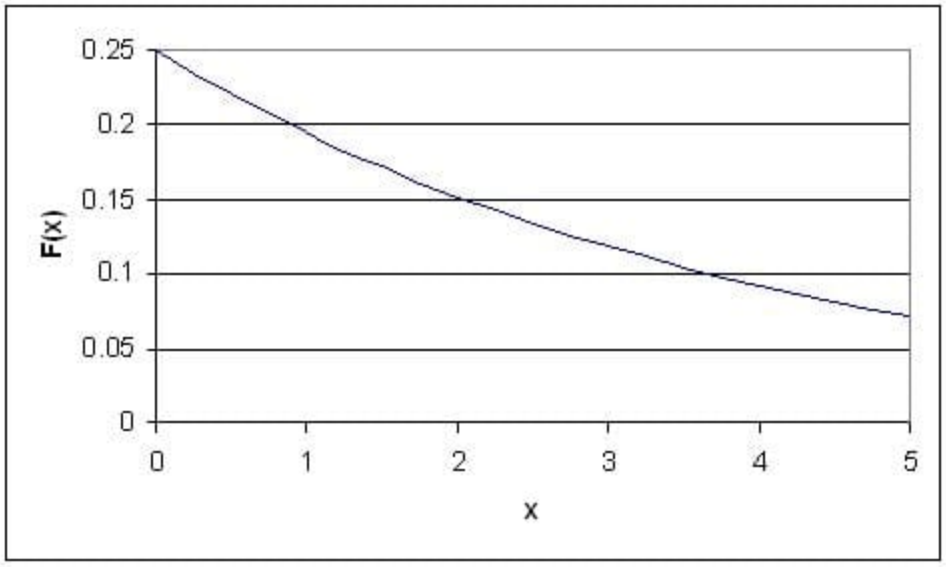


Figure 5: PDF (λ = 0.25) for Exponential Distribution

Here, x could represent time while the rate parameter could be the rate in which decay occurs. The rate parameter must be constant and greater than 0. The PDF decreases continuously in this diagram because of its definition as a decay example. Exponential decay typically models radioactive particles which lose mass per unit of time. Thus F(x) represents the mass of the particle with x equalling elapsed time since the start of the decay. As time passes, the mass falls due to radioactive decay (particle emits radiation as a form of energy release, any energy that results subtracts from mass, $E = MC^2$).

The Cumulative Distribution Function (CDF)

The CDF for an exponential distribution is expressed using the following:

$$P(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

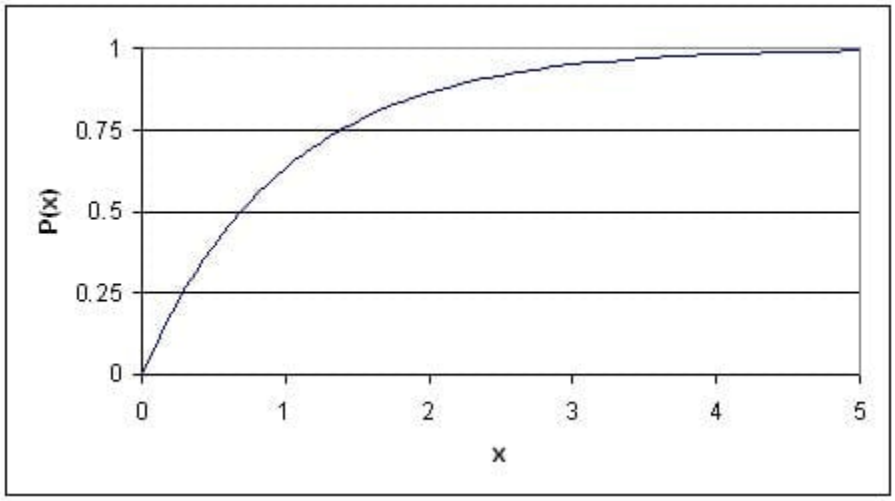


Figure 6: CDF (λ = 1) for Exponential Distribution

Following the example given above, this graph describes the probability of the particle decaying in a certain amount of time (x).

Properties of the Exponential Distribution

Among the distribution functions, the exponential distribution funtion has two unique properties, they are the memoryless property and a constant hazard rate.

--**The Memoryless Property**--

If a random variable, *X*, survived for "t" units of time, then the probability of *X* surviving an additional "s" units of time is the same as the probability of *X* suriving "s" units of time. The random variable has "forgotten" that it has survived for "t" units of time, thus this property is called the "memoryless" property.

A random variable, *X*, is *memoryless* if

$P\{X > s+t \mid X > t\} = P\{X > s\}$ for all $s,t \geq 0$ (1)

From using (1) above and the definition of conditional probability:

\frac{P(X s, X > t)}{P (X > t)} = P(x > s)" src="/@api/deki/files/19503/image-794.png">(2)

and so:

$P\{X > s+t\} = P\{X > s\}P\{X > t\}$ (3)

to prove the above property, we suppose that there is a function such that $F(x)=P\{X > x\}$.

by equation 3 above, we have

$F(x+t) = F(s)F(t)$ (4)

By elementary calculus, the only continuous solution for this kind of functiona equation is

$F(x) = e^{-\lambda x}$ (5)

It is thus proven the only distribution that can solve equation (4), or the memoryless property, is an exponential function.

--**The Constant Hazard Rate Property**--

Suppose the random variable, *X*, survived for "t" units of time, and r(t) represents the conditional probability that the t-year-old variable will fail. Mathematically the hazard rate, or the failure rate, is defined as:

$$r(t) = \frac{f(t)}{1 - F(t)} \quad (6)$$

For the exponential distribution, the hazard rate is constant since

$f(t) = e^{-\lambda x}$,

and

$$r(t) = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}}$$

$r(t) = \lambda$

the above proved that r(t) is constant for exponential distribution.

Standard Exponential Distribution

When an exponential distribution has λ = 1, this is called the standard exponential distribution. The equation and figure for this function is described below:

$$P(x \geq 0) = e^{-x}$$

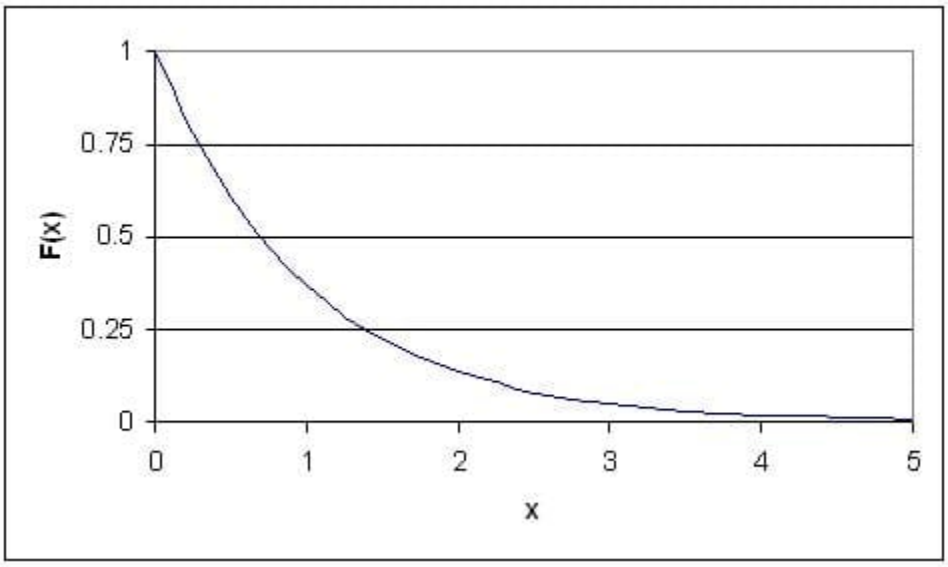


Figure 7: Standard Exponential Distribution Curve

This type of distribution is a way of standardizing your graph. This parallels our previous example of standard normal distribution, however, since time is now the x variable, it may not be negative (as an assumption of our hypothetical scenario). Thus, decay occurs only for positive time (as the radioactive particle emits mass).

Worked out Example 1

You are the lead process engineer for a leading pharmaceutical company, and your newest responsibility is to characterize one of your bioreactors. Part of this characterization is to determine the temperature variance within the reactor during the last step in an isothermal reaction. After taking data during this reaction for two weeks straight, you determine that the average temperature of the reaction is 30 C, with a standard deviation of your set of data being 3 C. Also, the data fits an normal distribution (ie. it is not skewed either side of the average) The higher-ups, ie. the ones who give you the raises, decide that this reactor needs to operate within 6 degrees of its average operating temperature at least 60% of the time. Given the data that you have taken, does the reactor meet these standards, or will you be pulling some 80 hour weeks in the near future trying to fix the process?

Solution

The information given in this problem is as follows:

- Average: 30
- Standard Deviation: 3
- We want to find the probability of the temperature falling between 24 and 36 C

Using the PDF for a Normal Distribution, we can say that:

$$P(24 < T < 36) - \frac{1}{2}(Erf[\frac{36-30}{3\sqrt{2}}] - Erf[\frac{24-30}{3\sqrt{2}}]) - \frac{1}{2}(Erf(1.4142) - Erf(-1.4142))$$

This expression can easily be solved using a table of Erf[] values as well as the property of the Erf function that states Erf(-x)=-Erf(x). Another simple approach would be to use Mathematica with the syntax shown below.

```
In[4]:= NIntegrate[PDF[NormalDistribution[30, 3], x], {x, 24, 36}]
Out[4]= 0.9545
```

Solving for P(24

Worked out Example 2

You are an engineer at a plant producing a volatile fuel. In your first week on the job both the primary control and redundancy control fail in the same day, necessitating the significant inconvenience of shutting down the reactor in order to prevent a much more inconvenient explosion. Your boss wants you to decide whether to simply replace both sensors or add an additional sensor to ensure that such problems are avoided in the future. While adding another sensor would certainly help solve the problem there is an extra cost associated with purchasing and installing the sensor. A very helpful fellow engineer with a penchant for statistics helps you out by informing you that these dual instrument failures occur at the rate of one per 3.8 * 10^8 hours. What is the probability that this type of failure will happen again in the next 50 years (the estimated life of the plant)? What is your recommendation to your boss?

Solution

Because you are given the probability of the event occuring and are interested in the amount of time separating two events using an exponential distribution is appropriate. To solve this you use the CDF expression:

$$CDF(x) = 1 - e^{-\lambda x}$$

In this case λ is given in hours and x in years, so you must first convert one of the units. Also the rate give is in events per hour. λ must be entered into the equation as the inverse (probability per time unit). Entering the appropriate numbers yeilds the following equation:

$$CDF(50) = 1 - e^{-50 years * \frac{1}{4.34 * 10^4 years}}$$

There is a 0.115% chance of both sensors failing in the same day sometime in the next 50 years. Based in this data you can conclude that you are probably safe simply replacing the two sensors and not adding a third.

Worked Out Example 3

You work at a chemical plant that uses the famous reaction A + B -> C to produce C. As part of your process you operate a mixing tank to dilute reactant A before sending it to the reactor. Reactant B is a byproduct from another process in the plant and comes in at a variable rate. This causes fluctuations in the level of the mixing tank for reactant A as the flow in and out of the tank has to be adjusted. In the past this has caused overflows in mixing tank A that your boss has tasked you with avoiding. You've decided the easiest way to avoid overflows is to simply buy a larger mixing tank. Currently the mixing tank is 50 gallons. After some experimentation you find the normal operating conditions of the tank are 45 gallons of reactant A with a 3 gallon standard deviation. However, to minimize costs the tank needs to be as small as possible. What is smallest tank you can buy that avoids overflows 99.99% of the time?

Solution:

Known values:
μ = 45*gal*.
σ = 3*gal*.

Using a z-table you find that the smallest z-value that gives an F(x) of 0.9999 is 3.71. You then rearrange the z equation given below to solve for x and plug in all known values.

$$z = \frac{x - \mu}{\sigma}$$

$x = z * \sigma + \mu$

$x = 3.71 * 3 + 45 = 56.1$

In order to prevent overflows 99.99% of the time the smallest tank that can be used is 56.1 gal.

Multiple Choice Question 1

Please briefly summarize the primary difference between continuous and discrete distributions, and their respective methods of analysis:

a. Continuous distributions and discrete distributions are always both analyzed through calculus methods only. Continuous distributions are distinct from discrete distributions in that the sample size is too small to treat each individual event in a discrete manner.

b. Continuous distributions are analyzed by applying calculus methods whereas discrete distributions are commonly addressed via arithmetic methods. Continuous distributions are distinct from discrete distributions in that the sample size is too large to treat each individual event in a discrete manner.

c. By performing a standard z- transform, the interested student may begin the process of characterizing the probability of z falling between the specific range from x = ‘a’ and x = ‘b’.

d. Error functions allow us to determine the likelihood of a data point falling within certain bounds without having to complete the rigorous of the probability distribution function. Continuous distributions may be addressed in this manner since their sample size is too confusing to be addressed in a discrete manner.

Multiple Choice Question 2

Please explain the primary philosophical difference between the parameters of the probability distribution function (PDF) and the cumulative distribution function (CDF).

a. Who uses the word philosophical in a chemical engineering problem? The author of this question is most likely confused and trying too hard, so I will protest by not answering.

b. The PDF and CDF are identical in application, and as such, must also have identical parameters.

c. The PDF and CDF both commonly measure the probability of sensor noise being at or below a specific value. They are different in the fact that the CDF contains a specific range of use, ranging from k1 to k2, whereas the PDF does not, as it rises from negative infinity to a value, k.

d. The major difference the PDF and CDF is that the PDF provides the probability of the variable x falling within a defined range. Alternatively, CDF gives the probability of the variable x falling at or below a certain limit, k.

Sage's Corner

Therapeutic Insulin Production
www.youtube.com/v/b3aDqdOzP5I

A Brief Problem on Continuous Exponential Distribution Functions
video.google.com/googleplayer...13000402551106

A copy of the slides can be found here:Slides without narration

Normal Distribution Functions in Excel

If YouTube doesn't work, click here Normal Distribution Functions in Excel

References

- Gernie, L. and W. Smith "Cartoon Guide to Statistics" Harper Perennial, c. 1993

This page titled 13.8: Continuous Distributions- normal and exponential is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

13.9: Discrete Distributions - Hypergeometric, Binomial, and Poisson

What are Discrete Distributions?

In order to explain the meaning and utility of dicrete distributions, it is necessary to first introduce the topic of **Random Variables**. The term *random* in Random Variable refers to the idea that each individual outcome has an equal chance of occuring. Therefore, each outcome is randomly determined. For example, a fair coin has an equal probability of flipping up heads as it does flipping up tails. Hence, the random variable is determined by the outcome of flipping a coin. However, it is frequently the case in chemical engineering that when a process is performed we are only interested in some function of the outcome rather than the outcome itself. For example, in a chemical reactor we might want to know the probability that two molecules collide and react to form a product at a certain temperature, however it is more useful to know the sum of the these collisions that form product rather than the individual occurences. This is analgous to rolling two dice and wanting to know the sum of the dice rather than the individual numbers rolled. These discrete values defined on some sample space compose what are called **Random Variables**.

Random Variable Example

Suppose that our experiment consists of tossing 4 fair coins. If we define a variable X and let it denote the number of heads flipped, then X is a random variable with possible discrete values of 0,1,2,3,4 with respective probabilities:

$$P\{X=0\} = P\{T,T,T,T\} = \frac{1}{16}$$

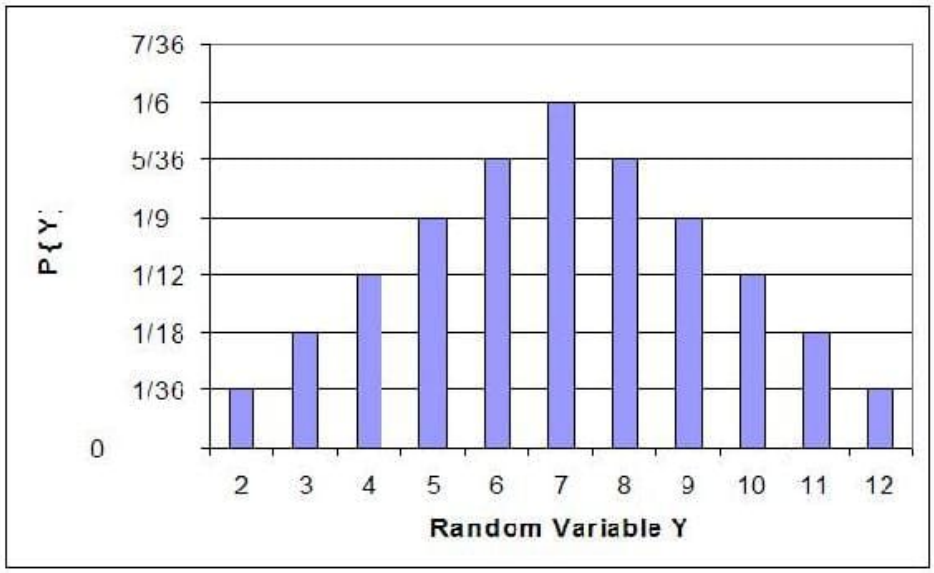
$$P\{X=1\} = P\{T,T,T,H\}, P\{T,T,H,T\}, P\{T,H,T,T\}, P\{H,T,T,T\} = \frac{1}{4}$$

$$P\{X=2\} = P\{T,T,H,H\}, P\{T,H,H,T\}, P\{H,T,H,T\}, P\{H,T,T,H\}, P\{T,H,T,H\}, P\{H,H,T,T\} = \frac{3}{8}$$

$$P\{X=3\} = P\{T,H,H,H\}, P\{H,T,H,H\}, P\{H,H,T,H\}, P\{H,H,H,T\} = \frac{1}{4}$$

$$P\{X=4\} = P\{H,H,H,H\} = \frac{1}{16}$$

We can do the same for the probability of rolling sums of two dice. If we let Y denote the sum, then it is a random variable that takes on the values of 2, 3, 4, 5, 6, 7, 8, 9, 10, & 12. Rather than writing out the probability statements we can represent Y graphically:



The graph plots the probability of the Y for each possible random value in the sample space (y-axis) versus the random value in the sample space (x-axis). From the graph one can infer that the sum with the greatest probability is $Y = 7$.

These are just two ways one can describe a random variable. What this leads into is representing these random variables as functions of probabilities. These are called the discrete distributions or probability mass functions. Furthermore, independent random events with known probabilities can be lumped into a discrete Random Variable. The Random Variable is defined by certain criteria, such as flipping up heads a certain number of times using a fair coin. The probability of a certain random variable equaling a discrete value can then be described by a discrete distribution. Therefore, a discrete distribution is useful in determining the probability of an outcome value without having to perform the actual trials. For example, if we wanted to know the probability of rolling a six 100 times out of 1000 rolls a distribution can be used rather than actually rolling the dice 1000 times.

Note: Here is a good way to think of the difference between discrete and continuous values or distributions. If there are two continuous values in a set, then there exists an infinite number of other continuous values between them. However, discrete values can only take on specific values rather than infinite divisions between them. For example, a valve that can only be completely open or completely closed is analogous to a discrete distribution, while a valve that can change the degree of openness is analogous to a continuous distribution.

The three discrete distributions we discuss in this article are the binomial distribution, hypergeometric distribution, and poisson distribution.

Binomial Distribution

This type of discrete distribution is used only when both of the following conditions are met:

1. The test has only two possible outcomes
2. The sample must be random

If both of the above conditions are met, then one is able to use this distribution function to predict the probability of a desired result. Common applications of this distribution range from scientific and engineering applications to military and medical ones. For example, a binomial distribution may be used in determining whether a new drug being tested has or has not contributed ("yes" or "no" labels) to alleviating symptoms of a disease.

Since only two outcomes are possible, they can be denoted as M_S and M_F for the number of successes and number of failures respectively. The term p is the frequency with which the desired number, M_S or M_F will occur. The probability of a collection of the two outcomes is determined by the following equation.

$$P\left(M_S, M_F\right) = kp^{M_S}(1-p)^{M_F} \tag{13.9.1}$$

k in the above equation is simply a proportionality constant. For the binomial distribution it can be defined as the number of different combinations possible

$$k = C\left(\frac{M_S + M_F}{M_S}\right) = \frac{(M_S + M_F)!}{M_S! M_F!} \tag{13.9.2}$$

! is the factorial operator. For example, 4! = 4 * 3 * 2 * 1 and *x*! = *x* * (*x* – 1) * (*x* – 2) * ... * 2 * 1. In the above equation, the term (M_S + M_F)! represents the number of ways one could arrange the total number of M_S and M_F terms and the denominator, $N_S!M_F!$, represents the number of ways one could arrange results containing M_S successes and M_F failures. Therefore, the total probability of a collection of the two outcomes can be described by combining the two above equations to produce the binomial distribution function.

$$P(M_S, M_F) = \frac{(M_S + M_F)!}{M_S! M_F!} p^{M_S} (1-p)^{M_F} \tag{13.9.3}$$

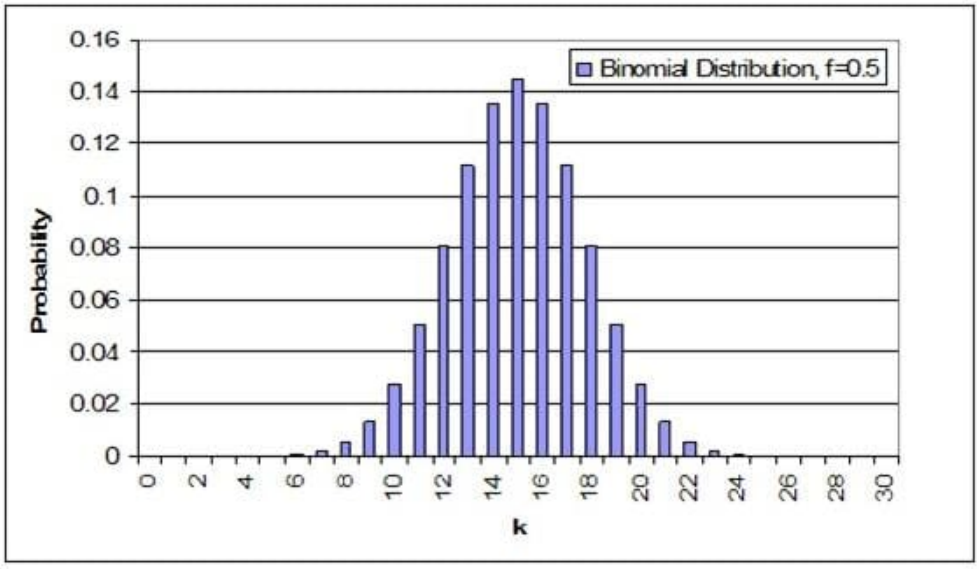
This can be simplified as the equation

$$\Gamma\{X - k\} - \frac{n!}{(n-k)!k!} p^k (1-p)^n \tag{13.9.4}$$

for $k = k = -0, 1, 2, 3 \ldots$.

In the above equation, *n* represents the total number of possibilities, or M_S + M_F, and *k* represents the number of desired outcomes, or M_S. These equations are valid for all non-negative integers of M_S , M_F , *n* , and *k* and also for *p* values between 0 and 1.

Below is a sample binomial distribution for 30 random samples with a frequency of occurrence being 0.5 for either result. This example is synonymous to flipping a coin 30 times with *k* being equal to the number of flips resulting in heads.



An important note about these distributions is that the area under the curve, or the probability of each integer added together, always sums up to 1.

One is able to determine the mean and standard deviation, which is described in the Basic Statistics article, as shown below.

$$\mu = np \qquad (5)$$

$$\sigma = \sqrt{np(1-p)} \qquad (6)$$

An example of a common question that may arise when dealing with binomial distributions is “What is the probability of a certain result occurring 33 – 66% of the time?” In the above case, it would be synonymous to the likelihood of heads turning up between 10 and 20 of the trials. This can be determined by summing the individual probabilities of each integer from 10 to 20.

qquad P(66% X > 33%)=\sum_{possible values}P(N_S)=P(10)+P(11)+...+P(19)+P(20)" src="/@api/deki/files/19527/image-813.png">
$$= 0.0280 + 0.0509 + ... + 0.0509 + 0.0280 = 0.9572$$

The probability of heads resulting in 33 – 66% of the trials when a coin is flipped 30 times is 95.72%.

The probability for binomial distribution can also be calculated using Mathematica. This will simplify the calculations, save time, and reduce the errors associated with calculating the probability using other methods such as a calculator. The syntax needed to calculate the probability in Mathematica is shown below in the screen shot. An example is also given for various probabilities based on 10 coin tosses.

```
In[22]:= (*Binomial Distribution*)
         binom[nn_, kk_, pp_] := nn! / (kk! (nn - kk)!) * pp^kk * (1 - pp)^(nn - kk)

         (*Example for Probability of getting 5 heads our of 10 tosses*)
         binom[10, 5, 0.5]
         Sum[binom[10, i, 0.5], {i, 0, 5}]
         d = Table[binom[10, x, 0.50], {x, 0, 10}]

Out[23]= 0.246094   Probability of exactly 5 heads out of 10 tosses

Out[24]= 0.623047   Probability of 0-5 heads out of 10 tosses

         Table of probabilities ranging from 0 to 10 heads out of 10 tosses

Out[25]= {0.000976563, 0.00976563, 0.0439453, 0.117188, 0.205078,
          0.246094, 0.205078, 0.117188, 0.0439453, 0.00976563, 0.000976563}
```

where k is the number of events, n is the number of independent samples, and p is the known probability. The example above also generates a set of values for the probabilities of getting exactly 0 to 10 heads out of 10 total tosses. This will be useful because it simplifies calculating probabilities such as getting 0-6 heads, 0-7 heads, etc. out of 10 tosses because the probabilities just need to be summed according to the exact probabilities generated by the table.

In addition to calling the binomial function generated above, Mathematica also has a built in Binomial Distribution function that is displayed below:

PDF[BinomialDistribution[n,p],k] where n,p, and k still represent the same variables as before

This built in function can be applied to the same coin toss example.

```
In[1]:= PDF[BinomialDistribution[10, 0.5], 5]
Out[1]= 0.246094
In[2]:= Sum[PDF[BinomialDistribution[10, .5], k], {k, 0, 5}]
Out[2]= 0.623047
```

As expected, the probabilities calculated using the built in binomial function matches the probabilities derived from before. Both methods can be used to ensure accuracy of results.

Poisson Distribution

Note: The variables used in this section are defined the same way as seen above in the "Binomial Distribution" section.

A Poisson distribution has several applications, and is essentially a derived limiting case of the binomial distribution. It is most applicably relevant to a situation in which the total number of successes is known, but the number of trials is not. An example of such a situation would be if you know the mean expected number of cancer cells present per sample and wanted to determine the probability of finding 1.5 times that amount of cells in any given sample, you would use a Poisson distribution.

In order to derive the Poisson distribution, you must first start off with the binomial distribution function.

$$P\{X-k\} = \frac{n!}{(n-k)!k!}p^k(1-p)^n \tag{13.9.5}$$

for $k = 1, 2, 3 \ldots$.

Due to the fact that the Poisson distribution does not require an explicit statement of the total number of trials, you must eliminate *n* out of the binomial distribution function. This is done by first introducing a new variable (μ), which is defined as the expected number of successes during the given interval and can be mathematically described as:

$$\mu = np \tag{13.9.6}$$

We can then solve this equation for p, substitute into Equation 13.9.6, and obtain Equation 13.9.7, which is a modified version of the binomial distribution function.

$$P\{X=k\} = \frac{n!}{(n-k)!k!}\left(\frac{\mu}{n}\right)^k\left(1-\frac{\mu}{n}\right)^{n-k} \tag{13.9.7}$$

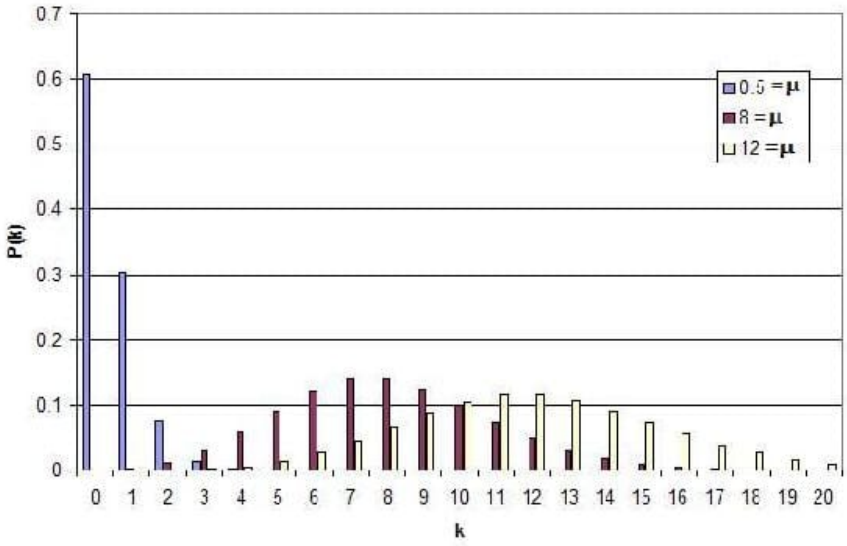
If we keep μ finite and allow the sample size to approach infinity we obtain Equation 13.9.8, which is equal to the Poisson distribution function. (Notice to solve the Poisson distribution, you do not need to know the total number of trials)

$$P\{X=k\} = \frac{\mu^k e^{-\mu}}{k!} \tag{13.9.8}$$

In the above equation:

- *e* is the base of the natural logarithm (*e* = 2.71828...)
- *k* is the number of occurrences of an event
- *k*! is the factorial of *k*
- *μ* is a positive real number, equal to the expected number of occurrences that occur during the given interval.

Using the Poisson distribution a probability versus number of successes plot can be made for several different numbers of successes.



This graph represents several key characteristics of a Poisson process.

1. The number of successes is independent of the time since the last event.
2. Probabilities, or outcomes, are discrete.
3. The probability of two or more successes in a short interval is extremely small (approximately zero).

These rules hint that the Poisson processes only model rare events, yet this is not true. What these characteristics say is that if you are limited to a small interval window, you will observe less than two events. It is important to know that the Poisson distribution is actually an approximation of the binomial distribution. As *n* increases and *p* decreases, the Poisson distribution becomes equal to the binomial distribution.

Hypergeometric Distribution

Note: The definitions of the variables in this section are different than the previous sections.

A hypergeometric distribution function is used only if the following three conditions can be met:

1. Only two outcomes are possible
2. The sample must be random
3. Selections are not replaced

Hypergeometric distributions are used to describe samples where the selections from a binary set of items are not replaced. This distribution applies in situations with a discrete number of elements in a group of **N** items where there are **K** items that are different. As a simple example imagine we are removing 10 balls from a jar of mixed red and green balls without replacing it. What is the chance of selecting 2 green balls? 5 red balls? In general terms, if there are *K* green balls and *N* total balls , the chance of selecting *k* green balls by selecting *n* balls is given by

$$P(x=k)=\frac{(\#\text{ of ways to select } k \text{ green balls })(\#\text{ of ways to select } n-k \text{ red balls })}{(\text{ total number of ways to select})} \tag{13.9.9}$$

As another example, say there are two reactants sitting in a CSTR totaling to **N** molecules. The third reactant can either react with A or B to make one of two products. Say now that there are **K** molecules of reactant A, and **N-K** molecules of reactant B. If we let **n** denote the number of molecules consumed, then the probability that **k** molecules of reactant of A were consumed can be described by the hypergeometric distribution. Note: This assumes that there is no reverse or side reactions with the products.

In mathematical terms this becomes

$$P\{X=k\}=\frac{(N-K)!K!n!(N-n)!}{k!(K-k)!(N-K+k-n)!(n-k)!N!}label10$$

where,

- N = total number of items
- K = total number of items with desired trait
- n = number of items in the sample
- k = number of items with desired trait in the sample

This can be written in shorthand as

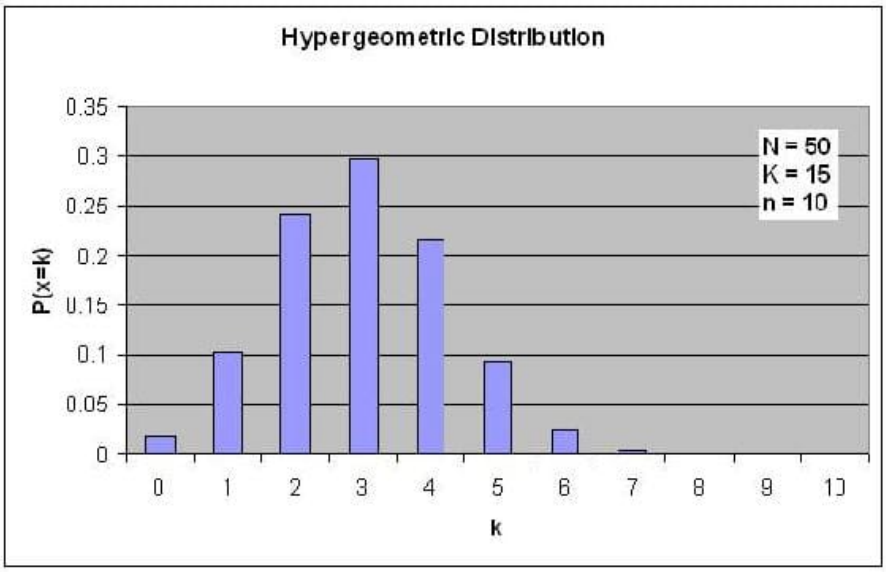
$$P\{X=k\}=\frac{\left(\begin{array}{c}K\\k\end{array}\right)\left(\begin{array}{c}N-K\\n-k\end{array}\right)}{\left(\begin{array}{c}N\\n\end{array}\right)} \tag{13.9.10}$$

where

$$\left(\begin{array}{c}A\\B\end{array}\right)=\frac{A!}{(A-B)!B!} \tag{13.9.11}$$

The formula can be simplified as follows: There are $\binom{N}{n}$ possible samples (without replacement). There are $\binom{K}{k}$ ways to obtain k green balls and there are $\binom{N-k}{n-k}$ ways to fill out the rest of the sample with red balls.

If the probabilities P are plotted versus k, then a distribution plot similar to the other types of distributions is seen.



Example 13.9.1

Suppose that you have a bag filled with 50 marbles, 15 of which are green. What is the probability of choosing exactly 3 green marbles if a total of 10 marbles are selected?

Solution

\[P\{X=3\}=\frac{\left(

$$\begin{pmatrix} 15 \\ 3 \end{pmatrix} \tag{13.9.12}$$

\right)\left(

$$\begin{pmatrix} 50-15 \\ 10-3 \end{pmatrix} \tag{13.9.13}$$

\right)}{\left(

$$\begin{pmatrix} 50 \\ 10 \end{pmatrix} \tag{13.9.14}$$

\right)}]

$$P\{X=3\} = \frac{(50-15)!15!10!(50-10)!}{3!(15-3)!(50-15+3-10)!(10-3)!50!}$$

$$P\{X=3\} = 0.2979$$

Fisher's exact

A special case of the hypergeometric distribution is the Fisher's exact method. Fisher's exact is the probability of sampling a specific configuration of a 2 by 2 table with constrained marginals. Marginals in this case refer to the sums for each row and each column. Therefore, every Fisher's exact will have 4 marginals, those 4 being the sum of the first column, the sum of the second column, the sum of the first row, and the sum of the second row. Since these values are constant, that also means that the sum of **all** the elements in the table will always equal the same thing. This is clarified with the image below:

	Column 1	Column 2	
Row 1	A	B	E = A + B
Row 2	C	D	F = C + D
	G = A + C	H = B+ D	I = A + B + C + D

In the image above, constant marginals would mean that E, F, G, and H would be held constant. Since those values would be constant, I would also be constant as I can be thought of as the sum of E and F or G and H are of which are constants.

In theory this test can be used for any 2 by 2 table, but most commonly, it is used when necessary sample size conditions for using the z-test or chi-square test are violated. The table below shows such a situation:

From this p_{fisher} can be calculated:

$$p_{\text{fisher}} = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{(a+b+c+d)!a!b!c!d!}$$

where the numerator is the number of ways the marginals can be arranged, the first term in the denominator is the number of ways the total can be arranged, and the remaining terms in the denominator are the number of ways each observation can be arranged.

As stated before, this calculated value is only the probability of creating the specific 2x2 from which the p_{fisher} value was calculated.

Another way to calculate p_{fisher} is to use Mathematica. The syntax as well as an example with numbers can be found in the screen shot below of the Mathematica notebook. For further clarification, the screen shot below also shows the calculated value of p_{fisher} with numbers. This is useful to know because it reduces the chances of making a simple algebra error.

```
In[8]:= (*Fisher's Exact*)
        fishd[a_, b_, c_, d_] :=
         (a+b)! (c+d)! (a+c)! (b+d)! / ((a+b+c+d)! a! b! c! d!)

        (*Fisher's Exact Example with Numbers where a=5,b=1,c=1,and d=18*)
        Pfisher = N[fishd[5, 1, 1, 18]]
Out[9]= 0.000643704
```

Another useful tool to calculate p_{fisher} is using available online tools. The following link provides a helpful online calculator to quickly calculate p_{fisher}. [1]

Using Mathematica or a similar calculating tool will greatly simplify the process of solving for p_{fisher}.

After the value of p_{fisher} is found, the p-value is the summation of the Fisher exact value(s) for the more extreme case(s), if applicable. The p-value determines whether the null hypothesis is true or false. An example of this can be found in the worked out hypergeometric distribution example below.

Finding the p-value

As elaborated further here: [2], the p-value allows one to either reject the null hypothesis or not reject the null hypothesis. Just because the null hypothesis isn't rejected doesn't mean that it is advocated, it just means that there isn't currently enough evidence to reject the null hypothesis.

In order to fine the p-value, it is necessary to find the probabilities of not only a specific table of results, but of results considered even more "extreme" and then sum those calculated probabilities. An example of this is shown below in example 2.

What is considered "more extreme" depends on the situation. In the most general sense, more extreme means further from the expected or random result.

Once the more extreme tables have been created and the probability for each value obtained, they can be added together to find the p-value corresponding to the least extreme table.

It is important to note that if the probabilities for every possible table were summed, the result would invariably be one. This should be expected as the probability of a table being included in the entire population of tables is 1.

Example 13.9.1

The example table below correlates the the amount of time that one spent studying for an exam with how well one did on an exam.

	Does not study	Does Study	Totals
Score Poorly on Exam	1	3	4
Score Well on Exam	4	3	7
Totals	5	6	11

There are 5 possible configurations for the table above, which is listed in the Mathematica code below. All the p_{fisher} for each configuration is shown in the Mathematica code below.

```
(*Problem 2c*)
(*Fishsers Exact*)
fishd[a_, b_, c_, d_] := (a + b)! (c + d)! (a + c)! (b + d)! / ((a + b + c + d)! a! b! c! d!)
x1 = N[fishd[1, 3, 4, 3]]
x2 = N[fishd[0, 4, 5, 2]]
x3 = N[fishd[2, 2, 3, 4]]
x4 = N[fishd[3, 1, 2, 5]]
x5 = N[fishd[4, 0, 1, 6]]
x1 + x2 + x3 + x4 + x5 (*all possible cases, should sum to 1*)
```

```
0.30303

0.0454545

0.454545

0.181818

0.0151515

1.
```

Example 13.9.1

What are the odds of choosing the samples described in the table below with these marginals in this configuration or a more extreme one?

	catalyst A	catalyst B	Totals
Bad product	5	1	6
Good product	1	18	19
Totals	6	19	25

Solution

First calculate the probability for the initial configuration.

$$p_{\text{fisher}} = \frac{6!19!6!19!}{25!5!1!1!18!} = 0.000643704$$

Then create a new table that shows the most extreme case that also maintains the same marginals that were in the original table.

	catalyst A	catalyst B	Totals
Bad product	6	0	6
Good product	0	19	19
Totals	6	19	25

Then calculate the probability for the most extreme case.

$$p_{\text{fisher}} = \frac{6!19!6!19!}{25!6!0!0!19!} = 0.00000564653$$

Finally add the two probabilities together.

$$p_{fisher} = 0.000643704 + 0.00000564653 = 0.00064$$

Maximum Entropy Function

The maximum entropy principle uses all the testable information about multiple given probability distributions to describe the true probability distribution by maximizing the information entropy. To derive a distribution with the maximum entropy principle, you obtain some testable information *I* about some quantity *x*. This information is expressed as *m* constraints on the expectation values of the functions f_k. This makes the Bayesian probability distribution to satisfy:

$$F_k = \sum_{i=1}^{n} \Pr\left(x_i \mid I\right) f_k\left(x_i\right)$$

with $k = 1, \cdots, m$.

We also know that all of these probabilities must sum to 1, so the following constraint is introduced:

$$\sum_{i=1}^{n} \Pr\left(x_i \mid I\right) = 1$$

Then the probability distribution with maximum information entropy that satisfies all these constraints is:

$$\Pr\left(x_i \mid I\right) = \frac{1}{Z\left(\lambda_1, \cdots, \lambda_m\right)} \exp\left[\lambda_1 f_1\left(x_i\right) + \cdots + \lambda_m f_m\left(x_i\right)\right]$$

The normalization constant is determined by the classical partition function:

$$Z\left(\lambda_1, \cdots, \lambda_m\right) = \sum_{i=1}^{n} \exp\left[\lambda_1 f_1\left(x_i\right) + \cdots + \lambda_m f_m\left(x_i\right)\right]$$

The λ_k parameters are Lagrange multipliers with values determined by:

$$F_k = \frac{\partial}{\partial \lambda_k} \log Z\left(\lambda_1, \cdots, \lambda_m\right)$$

All of the well-known distributions in statistics are maximum entropy distributions given appropriate moment constraints. For example, if we assume that the above was constrained by the second moment, and then one would derive the Gaussian distribution with a mean of 0 and a variance of the second moment.

The use of the maximum entropy distribution function will soon become more predominant. In 2007 it had been shown that Bayes' Rule and the Principle of Maximum Entropy were compatible. It was also shown that maximum entropy reproduces every aspect of orthodox Bayesian inference methods. This enables engineers to tackle problems that could not be simply addressed by the principle of maximum entropy or Bayesian methods, individually, in the past.

Summary

The three discrete distributions that are discussed in this article include the Binomial, Hypergeometric, and Poisson distributions. These distributions are useful in finding the chances that a certain random variable will produce a desired outcome.

Binomial Distribution Function

The Binomial distribution function is used when there are only two possible outcomes, a *success* or a *faliure*. A success occurs with the probability **p** and a failure with the probability **1-p**. Suppose now that in **n** independent trials the *binomial* random variable **X** represents the number of successes. The following equation applies to binomial random variables:

$$P\left\{X - k\right\} - \frac{n!}{(n-k)!k!} p^k \left(1-p\right)^{n-k} \qquad for \quad k - 0, 1, 2, ...n$$

$$P\{X \leq i\} = \sum_{k=0}^{i} \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k}$$

Poisson Distribution Function

The Poisson distribution can be used in a large variety of situations. It can also be used to approximate the Binomial Distribution when **n** is large and **p** is small yielding a moderate **np**.

The following example is a situation in which the Poisson Distribution applies: Suppose a CSTR is full of molecules and the probability of any one molecule reacting to form product is small (say due to a low temperature) while the number of molecules is large, then the probability distribution function would fit the Poisson Distribution well (Note: all the molecules are the same here, unlike in the hypergeometric example to follow). Other examples include, the number of wrong telephone numbers that are dialed each day or the number of people in a community that live to 100 years of age. The following equation is the Poisson Distribution function:

$$P\{X = k\} = \frac{\mu^k e^{-\mu}}{k!} \qquad for \quad k = 0, 1, 2, \ldots n$$

$$P\{X \leq i\} = \sum_{k=0}^{i} \frac{\mu^k e^{-\mu}}{k!}$$

Summary of Key Distributions

Distribution	Requirements	Distribution Function
Binomial	• Only two outcomes are possible • Sample is random • Success is independent of the interval • The probability of two consecutive successes is small	$P\{X=k\} = \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k}$
Poisson		$P\{X=k\} = \frac{\mu^k e^{-\mu}}{k!}$
Hypergeometric	• Selections are not replaced	$P\{X=k\} = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$

Summary of Distribution Approximations

There are many useful approximations for discrete random variables. An important approximation is that which yields a normal distribution because it allows for confidence intervals and probabilities to be continuous. Additionally, since the normal distribution is so well understood it is used for convenience when speaking to those who have yet to study probability theory. Below is a table of useful approximations. The majority of the theory behind these approximations is in the application of the Central Limit Theorem and Large Number Law but these are outside the scope of this text.

The Following Distribution	Is Approximatly A...	When the Following Are Met...
Hypergeometric(n, K, N)	A continuous Normal random variable where $\mu = nK/N$ and $\sigma^2 = n(K/N)(1-K/N)$. **Normal**$(nK/N, n(K/N)(1-K/N))$	1) Satisfies requirements for approximating a Binomial below 2) There are a large number of samples, n. At least 100.
	A Binomial random variable where $n = n$ and $p = K/N$ **Binomial**$(n, K/N)$	1) Both K & N are large in comparison to n (an order of magnitude or larger)
Binomial(n,p)	A continuous Normal random variable where $\mu = np$ and $\sigma^2 = np(1-p)$. **Normal**$(np, np(1-p))$	1) Large number of samples n. At least 100.
	A Poisson random variable where $\lambda = np$ **Poisson**(np)	1) Large number of samples n. 2) Small Probability, p. 3) Moderate size $\lambda = np$
Poisson(λ)	A continuous Normal random variable where $\mu = \lambda$ and $\sigma^2 = \lambda$. **Normal**(λ, λ)	1) Large number of samples n. At least 100.

The above chart can be read as "A random variable under the following distribution can is approximately a random variable when the following requirements are met." For example, the final line of the table can be read: "A random variable following a Poisson distribution with parameter λ is approximately a continuous Normal random variable where the mean is λ and the variance is λ assuming that there are at least 100 samples that have been taken."

Example 13.9.1

In order for a vaccination of Polio to be efficient, the shot must contain at least 67% of the appropriate molecule, VPOLIO. To ensure efficacy, a large pharmaceutical company manufactured a batch of vaccines with each syringe containing 75% VPOLIO. Your doctor draws a syringe from this batch which should contain 75% VPOLIO. What is the probability that your shot, will successfully prevent you from acquiring Polio? Assume the syringe contains 100 molecules and that all molecules are able to be delivered from the syringe to your blood stream.

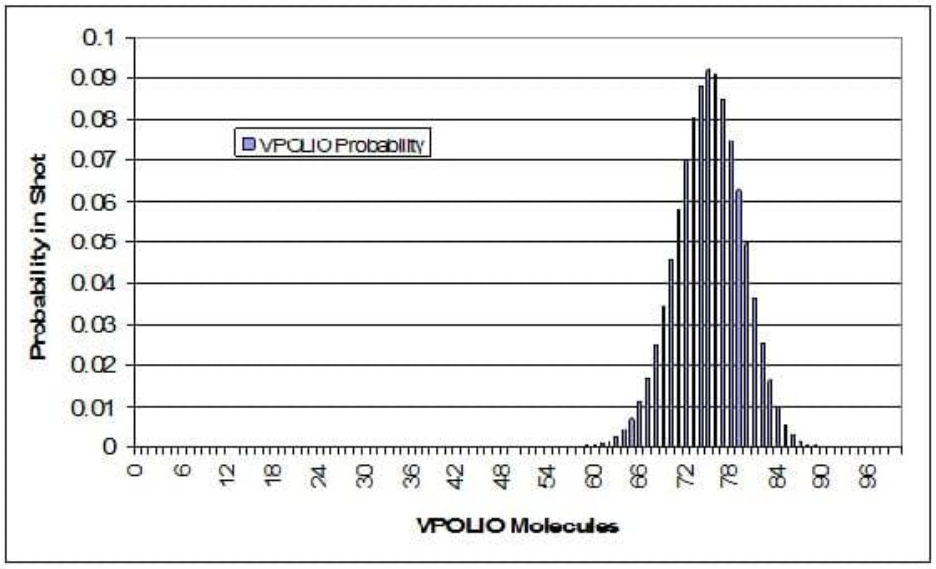
Solution

This can be done by first setting up the binomial distribution function. In order to do this, it is best to set up an Excel spreadsheet with values of k from 0 to 100, including each integer. The frequency of pulling a VPOLIO molecule is 0.75. Randomly drawing any molecule, the probability that this molecule will never be VPOLIO, or in other words, the probability of your shot containing 0 VPOLIO molecules is

$$P = \frac{100!}{(100-0!)!0!}(0.75)^0(1-0.75)^{100-0}$$

Note that 0!=1

Your spreadsheet should contain probabilities for all off the possible values of VPOLIO in the shot. A link to our spreadsheet can be seen at the end of this article. A graph of our distribution is shown below.



The next step is to sum the probabilities from 67 to 100 of the molecules being VPOLIO. This is calculation is shown in the sample spreadsheet. The total probability of at least 67 of the molecules being VPOLIO is 0.9724. Thus, there is a 97.24% chance that you will be protected from Polio.

Example 13.9.1: Gaussian Approximation Of A Binomial Distribution

Calculation of the binomial function with n greater than 20 can be tedious, whereas calculation of the Gauss function is always simple. To illustrate this, consider the following example.

Suppose we want to know the probability of getting 23 heads in 36 tosses of a coin. This probability is given by the following binomial distribution:

$$P = \frac{36!}{23!13!}(0.5)^{36} = 3.36\%$$

To use a Gaussian Approximation, we must first calculate the mean and standard deviation.

$$\mu = np = 18$$

$$\sigma = \sqrt{np(1-p)} = 3$$

Now we can approximate the probability by the Gauss function below.

$$P = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-X)^2}{2\sigma^2}} = 3.32\%$$

This approximation is very close and requires much less calculation due to the lack of factorials. The usefulness of the Gaussian approximation is even more apparent when n is very large and the factorials involve very intensive calculation.

Example 13.9.1: Worked out Hypergeometric Distribution

A teacher has 12 students in her class. On her most recent exam, 7 students failed while the other 5 students passed. Curious as to why so many students failed the exam, she took a survey, asking the students whether or not they studied the night before. Of the students who failed, 4 students did study and 3 did not. Of the students who passed, 1 student did study and 4 did not. After seeing the results of this survey, the teacher concludes that those who study will almost always fail, and she proceeds to enforce a "no-studying" policy. Was this the correct conclusion and action?

Solution

This is a perfect situation to apply the Fisher's exact test to determine if there is any association between studying and performance on the exam. First, create a 2 by 2 table describing this particular outcome for the class, and then calculate the probability of seeing this exact configuration. This is shown below.

	Did Study	Did Not Study	Totals
Failed Exam	4	3	7
Passed Exam	1	4	5
Totals	5	7	12

$$p_{\text{fisher },1} = \frac{(4+3)!(1+4)!(4+1)!(3+4)!}{(4+3+1+4)!4!3!1!4!} = 0.221$$

Next, create 2 by 2 tables describing any configurations with the exact same marginals that are more extreme than the one shown above, and then calculate the probability of seeing each configuration. Fortunately, for this example, there is only one configuration that is more extreme, which is shown below.

	Did Study	Did Not Study	Totals
Failed Exam	5	2	7
Passed Exam	0	5	5
Totals	5	7	12

$$p_{\text{fisher },2} = \frac{(5+2)!(0+5)!(5+0)!(2+5)!}{(5+2+0+5)!5!2!0!5!} = 0.0265$$

Finally, test the significance by calculating the p-value for the problem statement. This is done by adding up all the previously calculated probabilities.

$$p = p_{\text{fisher },1} + p_{\text{fisher },2} = 0.221 + 0.0265 = 0.248$$

Thus, the p-value is greater than 0.05, which is the standard accepted level of confidence. Therefore, the null hypothesis cannot be rejected, which means there is no significant association between studying and performance on the exam. Unfortunately, the teacher was wrong to enforce a "no-studying" policy.

Example 13.9.1: Poisson

The hormone PREGO is only found in the female human body during the onset of pregnancy. There is only 1 hormone molecule per 10,000 found in the urine of a pregnant woman. If we are given a sample of 5,000 hormone molecules what is the probability of finding exactly 1 PREGO? If we need at least 10 PREGO molecules to be 95% positive that a woman is pregnant, how many total hormone molecules must we collect? If the concentration of hormone molecules in the urine is 100,000 molecules/mL of urine,what is the minimum amount of urine (in mL) necessary to insure an accurate test(95% positive for pregnancy)?

Solution

This satisfies the characteristics of a Poisson process because

1. PREGO hormone molecules are randomly distributed in the urine
2. Hormone molecule numbers are discrete
3. If the interval size is made smaller (i.e. our sample size is reduced), the probability of finding a PREGO hormone molecule goes to zero

Therefore we will assume that the number of PREGO hormone molecules is distributed according to a Poisson distribution.

To answer the first question, we begin by finding the expected number of PREGO hormone molecules:

$$\mu = np$$

$$\mu = (5,000)\left(\frac{1}{10,000}\right) = 0.5$$

Next we use the Poisson distribution to calculate the probability of finding exactly one PREGO hormone molecule:

$$P\{X = k\} = \frac{\mu^k e^{-\mu}}{k!} = 0.303$$

where $k = 1$ and $\mu = 0.5$.

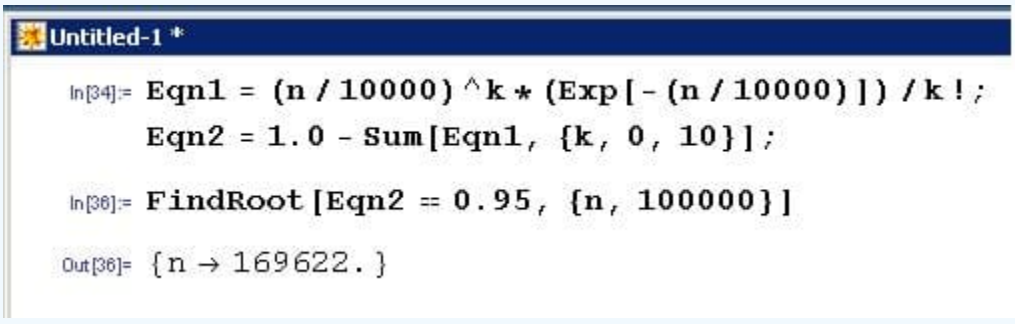
The second problem is more difficult because we are calculating **n**. We begin as before by calculating the expected number of PREGO hormone molecules:

$$\mu = np = \frac{n}{10,000}$$

Now we apply the Poisson Distribution Function equation:

$$F\{X \geq 10\} - 0.95 - \sum_{k=1}^{\infty} \frac{\mu^k e^{\mu}}{k!} - 1.0 - \sum_{k=0}^{-5} \frac{\mu^k e^{\mu}}{k!} - 1.0 - \sum_{k=1}^{10} \left(\left(\frac{n'}{10.000}\right)^k \frac{e^{-\left(\frac{n}{k.00!}\right)}}{k!} \right)$$

It is easiest to solve this in Mathematica:



The function **FindRoot[]** was used in Mathematica because the **Solve[]** function has problems solving polynomials of degree order greater than 4 as well as exponential functions. However, **FindRoot[]** requires an initial guess for the variable trying to be solved, in this case **n** was estimated to be around 100,000. As you can see from the Mathematica screen shot the total number of hormone molecules necessary to be 95% sure of pregnancy (or 95% chance of having atleast 10 PREGO molecules) was 169,622 molecules.

For the last step we use the concentration of total hormone molecules found in the urine and calculate the volume of urine required to contain 169,622 total hormone molecules as this will yield a 95% chance of an accurate test:

$$Concentration = \frac{100,000 \quad molecules}{mL \quad urine}$$

$$Volume \quad of \quad urine = \frac{169,622}{100,000} = 1.69622 \quad mL$$

Example 13.9.1: Gaussian Approximation to a Poisson Distribution

To illustrate the Gaussian approximation to the Poisson distribution, consider a distribution where the mean (μ) is 64 and the number of observations in a definite interval (N) is 72. Determine the probability of these 72 observations?

Solution

Using the Poisson Distribution

$$P(N) = e^{-\mu} \cdot \frac{\mu^N}{N!}$$

$$P(72) = e^{-64} \cdot \frac{64^7 2}{72!} = 2.9\%$$

This can be difficult to solve when the parameters N and μ are large. An easier approximation can be done with the Gaussian function:

$$P(72) = G_{64,8} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-X)^2}{2\sigma^2}} = 3.0\%$$

where, $X = \mu$ and $\sigma = \sqrt{\mu}$.

Exercise 13.9.1

All of the following are characteristics of the Poisson distribution **EXCEPT**:

a. The number of successes is independent of the interval
b. There are imaginary numbers
c. Probabilities, or outcomes, are discrete
d. Two or more success in a short interval is extremely small

Answer available at the end of this section.

Exercise 13.9.1

If there are **K** blue balls and **N** total balls , the chance of selecting **k** blue balls by selecting **n** balls in shorthand notation is given by:

a) $$P\{X=k\}=\frac{\binom{N}{n}\binom{K-k}{N-n}}{\binom{K}{n}}$$

b) $$P\{X=k\}=\frac{\binom{n}{k}\binom{n-K}{N-k}}{\binom{K}{n}}$$

c) $$P\{X=k\}=\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

d) $$P\{X=k\}=\frac{\binom{K}{n}\binom{N-n}{K-N}}{\binom{N}{n}}$$

Answer available at the end of this section.

Discrete Distribution Presentation: Clown Time

www.youtube.com/v/6WOieU664Ew

References

- Ross, Sheldon: A First Course in Probability. Upper Saddle River: Prentice Hall, Chapter 4.
- Uts, J. and R. Hekerd. *Mind on Statistics.* Chapter 15 - More About Inference for Categorical Variables. Belmont, CA: Brooks/Cole - Thomson Learning, Inc. 2004.
- Weisstein, Eric W.: MathWorld - Discrete Distributions. Date Accessed: 20 November 2006. MathWorld
- Woolf, Peter and Amy Keating, Christopher Burge, Michael Yaffe: Statistics and Probability Primer for Computational Biologists. Boston: Massachusetts Institute of Technology, pp 3.1 - 3.21.
- Wikipedia-Principle of maximum entropy. Date Accessed: 10 December 2008. [3]
- Giffin, A. and Caticha, A., 2007,[4].

Answers

Exercise 13.9.1

TBA

Exercise 13.9.1

TBA

This page titled 13.9: Discrete Distributions - Hypergeometric, Binomial, and Poisson is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Marc Palmer, Matt Pickvet, Rob Carr, & Tommy DiRaimondo via source content that was edited to the style and standards of the LibreTexts platform.

13.10: Multinomial Distributions

Introduction

Statistical events generate two types of outcomes: continuous or discrete. Continuous outcomes can take an infinite number of values; for instance, a person's height can take any real number, so there are infinite possibilities. Typical events generating continuous outcomes may follow a normal, exponential, or geometric distribution. Discrete outcomes can only on take prescribed values; for instance, a dice roll can only generate an integer between 1 to 6. Discrete outcomes are typically distributed either binomially or multinomially. It is with multinomial distribution that this article is concerned.

Multinomial Distributions: Mathematical Representation

Multinomial distributions specifically deal with events that have **multiple discrete outcomes**. The Binomial distribution is a specific subset of multinomial distributions in which there are only two possible outcomes to an event.

Multinomial distributions are not limited to events only having discrete outcomes. It is possible to categorize outcomes with continuous distributions to different degrees (high, medium, low). For instance, the water level - a continuous entity - in a storage tank can be made discrete by categorizing them into either "desirable" or "not desirable." Multinomial distributions, therefore, have expansive applications in process control.

Probability Density Function

One way of describing the probability of an outcome occurring in a trial is the **probability density function**. The probability density function (PDF) mathematically represents the probability of having a specified outcome. The probability density function is a useful way to find the probability of simultaneous occurrence of specific results (i.e. having $n_1 = 1$, $n_2 = 1$, and $n_3 = 1$ with 3 trials as opposed to other outcomes, such as $n_1 = 3$, $n_2 = 0$, $n_3 = 0$ or $n_1 = 0$, $n_2 = 1$, $n_3 = 2$).

$$P\left(n_1, n_2, \ldots, n_k\right) = \left(n_1! n_2! \ldots n_k!\right) \prod_{i=1}^{N!} p_i^{n_i} = \left(n_1! n_2! \ldots n_{l\ell}!\right) * \left(p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k}\right)$$

where

- N is the number of trials
- k is the number of possible outcomes
- n_i is the number of occurrences of outcome i
- p_i is the probability of observing outcome i

We know that the the sum of the probabilities of all possible outcomes that can occur in a trial must be unity (since one outcome must occur). The probability density function yields this result for both continuous and discrete outcomes. However, it is important to note that to get this result for continuous outcomes, one must take the integral of the probability density function over all possible outcomes. To get this unity result for discrete outcomes, one must sum the probabilities of each outcome (similar to taking Riemann sums).

Cumulative Distribution Function

While the probability density function calculates the probability of a single outcome, the cumulative distribution function (CDF) is a useful way to find the probability that an outcome lies within a given range of values.

$$P\left(n_1 \leq c_1, n_2 \leq c_2, \ldots, N - n_1 - n_2 - \ldots\right) = \sum_{n_1=0}^{c_1} \sum_{n_2=0}^{c_2} \ldots \sum_{n_{k-1}=0}^{c_k} \frac{N!}{n_1! n_2! \ldots \left(N - n_1 - n_2 - \ldots\right)!} \prod_{i=1}^{k} p_i^{n_i}$$

where

- N is the number of trials
- k is the number of possible outcomes
- n_i is the number of occurrences of outcome i
- p_i is the probability of seeing outcome i
- c_i is the maximum number of occurrences of outcome i

Visualizing Probability Density Function with Mathematica

Before using the functions for multinomial probability distributions, a special package must be loaded using the following command (depending on the version of Mathematica):

<< *Needs["MultivariateStatistics`"]* (Mathematica 6.0)

<< *Statistics`MultiDiscreteDistributions`* (Mathematica 5.2)

Note that the ` (tilde) is the key next to 1 and not the single quote character '.

Table of pertinent Mathematica commands:

Command	Description
PDF[multinomial,x]	probability density function at x, where x is scalar-, vector-, or matrix-valued depending on distribution
CDF[multinomial,x]	cumulative distribution function at x
Mean[multinomial]	mean
Variance[multinomial]	variance
StandardDeviation[multinomial]	standard deviation

To plot the multinomial distribution probability density function (PDF) in Mathematica, follow three simple steps:

Defining the Multinomial Distribution

multinomial = MultinomialDistribution[n,{p1,p2,...pk}] where k is the number of possible outcomes, n is the number of outcomes, and p1 to pk are the probabilities of that outcome occurring. n and p1 to pk are usually given as numbers but can be given as symbols as long as they are defined before the command.

Defining the PDF of the Multinomial Distribution

pdf=PDF[multinomial, {x1,x2,...,xk}]; The x here simply refers to the variable so this command can be typed as is, and leave the x as a variable not a number.

Plotting the PDF

Plot3D[pdf, {x1, 0, 6}, {x2, 0, 5}, AxesLabel -> {x1, x2, probability}] the 0,6 and 0,5 are the ranges of x1 and x2 on the plot respectively, and the AxesLabel argument is to help us see which is which on the plot created. This command can also be typed as is, by leaving all the x's as variables.

- Note that this will only work if you have 2 variables. If there are more variables, constraints can be set so that it can be plotted. For 3 variables, set the third variable x3 as n-x1-x2. See the attached Mathematica notebook for more information.

Media:Multinomial.nb

Other Characteristics

Statistics have historically been useful in descriptive and inferential analysis of data. Likewise, multinomial distribution is also applicable to the aforementioned areas: descriptive statistics, inferential statistics, and six-sigma. Several key variables are used

in these applications:

The expected value below describes the mean of the data. Discrete random variables can take on a range of values; the mean of the data describes the location of the data within this range.

$$\mathrm{E}\left(X_i\right) = np_i$$

The data's standard deviation describes the variance of the data. The standard deviation of the data describes the spread of the data with respect to the center value (the mean of the data).

$$\operatorname{var}\left(X_i\right) = np_i\left(1 - p_i\right)$$

Using the above parameters, it is possible to find the probability of data lying within n standard deviations of the mean. By setting n equal to 6, six sigma quality control can be implemented on the event and outcomes in question.

Derivation of Binomial Distribution

As was stated above, the binomial distribution is simply a special case of the multinomial distribution. Using the multinomial distribution, the probability of obtaining two events n1 and n2 with respective probabilities p_1 and p_2 from N total is given by:

$$P\left(n_1, n_2\right) = \frac{N!}{n_1! n_2!}\left(p_1^{n_1} p_2^{n_2}\right)$$

If we label the event of interest, say n1 in this case, as "k," then, since only two outcomes are possible, n2 must equal N-k. Finally, If we label the probability of obtaining k as simply "p," then the probability of obtaining event n2 (p2) must be 1-p, because again only two outcomes are possible. With these subsitutions, the above equation simplifies to

$$P(k, N, p) = \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k}$$

Which is the familiar binomial distribution, where k is the number of events of interest, N is the total number of events, and p is the probability of attaining the event of interest.

Applications of Multinomial Distributions

As mentioned before, multinomial distributions are a generalized version of binomial distributions. In chemical engineering applications, multinomial distributions are relevant to situations where there are more than two possible outcomes (temperature = {high, med, low}). Multinomial systems are a useful analysis tool when a “success-failure” description is insufficient to understand the system. A continuous form of the multinomial distribution is the Dirichlet distribution.

Using Bayes' Rule is one of the major applications of multinomial distributions. For example, Bayes' Rule can be used to predict the pressure of a system given the temperature and statistical data for the system. Bayes' Rule can be used to determine the probability of an event or outcome as mentioned above. Additional details on Bayes' Rule can be found at Bayes' Rule, conditional probability, independence.

Bayes' Rule Example

In case you’ve forgotten how to use Bayes’ Rule, here is an example that shows how to use it to solve a problem with a multinomial outcome which results from combining dual possibilities (such as desired flow vs. undesired flow and configuration A vs. configuration B).

You are given a flow apparatus with two possible specific valve configurations, A and B. For specific valve configuration A, desired flow rates are achieved 98.5% of the time. And for specific valve configuration B, the desired flow rates are achieved 89.3% of the time. An operator might use the configuration from A 79% of the time when desired flow rates are achieved. Can you determine the probability that a randomly chosen configuration is the same as the one in specific valve configuration B and produces undesirable flow rates?

Worked Out Solution

We'll be using the following symbols:

F_U: {undesirable flow rates}
F_D: {desirable flow rates}
C_B: {configuration from Apparatus B}

Thus we want to find : $P(F_U \cap C_B)$, which is the probability of undesirable flow rates and valve configuration B.

If we use Apparatus A 79% of the time then the probability of using Apparatus B would be 21% or : $P(F_D \cap C_B) = 0.21$, which is the probability of using configuration B and getting desired flow rates.

Additionally, we are given : $P(F_D \mid C_B) = 0.893$, which is the probability of a desired flow rate, given that configuration B is used.

Then $P(F_U \mid C_B) = 1 - P(F_D \mid C_B) = 1 - 0.893 = 0.107$, which is the probability that an undesirable flow rate is obtained, given that configuration B is used.

The information from above is then used below: $P(F_U \cap C_B) = P(F_U \mid C_B) * P(F_D \cap C_B) = (0.107)(0.21) = 0.0225 = 2.25\%$, which is the probability of getting an undesirable flow rate and configuration B which is found by taking the probability of getting an undesirable flow rate given that configuration B is used and multiplying that by the probability of a desirable flow rate and configuration B.

Example 13.10.1

Jan 30, 2021, 10:11 AM

from Perry's, page 3-72

Consider the scenario in which you toss a fair die 12 times. What is the probability that each face value (1-6) will occur exactly twice?

Solution

The probability can be determined using a multinomial distribution in which 6 outcomes are possible. The individual probabilities are all equal given that it is a fair die, p = 1/6. The total number of trials N is 12, and the individual number of occurrences in each category n is 2.

$$P(2,2,2,2,2,2) - \frac{12!}{2!2!2!2!2!2!} * \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 - 0.003488$$

Therefore, the probability of rolling exactly 2 of each face value on a fair die is about 0.35%.

Example 13.10.2

A bowl has 2 maize marbles, 3 blue marbles and 5 white marbles. A marble is randomly selected and then placed back in the bowl. You do this 5 times. What is the probability of choosing 1 maize marble, 1 blue marble and 3 white marbles?

Solution

- *N* is the number of trials = 5
- *k* is the number of possible outcomes = 3
- n_i is the number of occurrences of outcome i
- p_i is the probability of seeing outcome i

The three possible outcomes are choosing a maize marble, a blue marble or a white marble.

We must determine n_i and p_i to solve the multinomial distribution.

The number of occurrences of the outcome are the number of times we wish to see each outcome. These are given in the problem statement.

$n_{maize} = 1$

$n_{blue} = 1$

$n_{white} = 3$

The probability of seeing each outcome is easy to find. For example, there are two maize marbles in the bowl of 10, so the probability of choosing a maize marble is $\frac{2}{10}$.

$p_{maize} = \frac{2}{10}$

$p_{blue} = \frac{3}{10}$

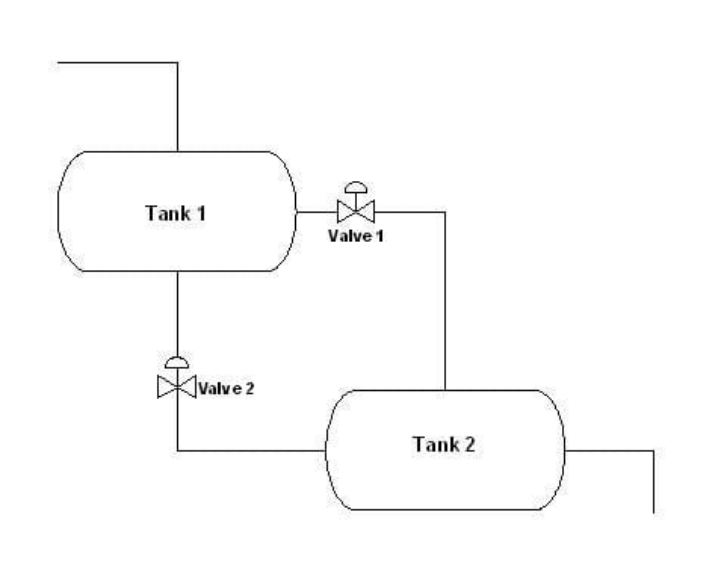
$p_{white} = \frac{5}{10}$

We can now solve the multinomial distribution as shown below. The probability of choosing 1 maize marble, 1 blue marble and 3 white marbles is 0.15.

$$P(1,1,3) = \frac{5!}{1!1!3!} * \left(\frac{2}{10}\right)^1 \left(\frac{3}{10}\right)^1 \left(\frac{5}{10}\right)^3 = 0.15$$

Example 13.10.3

Two valves are used to control the flow of liquid out of a storage tank (Tank 1) in to another storage tank (Tank 2) and controlled by Valves 1 and 2 as seen below. The valves exist in two states: open and closed.



The following table describes the four different valve configurations and the frequency of the desired flow for each valve configuration based on experimental data.

Configuration	Valve 1	Valve 2	Probability of that configuration having the best flow
1	Open	Closed	0.305
2	Open	Open	0.476
3	Closed	Open	0.218
4	Closed	Closed	0.001

To reduce costs, two different apparatuses with a similar configuration, but with different pipe dimensions are being considered to replace the original apparatus. Below are data tables for each potential replacement apparatus with different valve configurations and the number of times out of 100 measurements that the valve configuration gave the desired flow.

Apparatus 1

Configuration	Valve 1	Valve 2	# of Best Flow
1	Open	Closed	42
2	Open	Open	37
3	Closed	Open	21
4	Closed	Closed	0

Apparatus 2

Configuration	Valve 1	Valve 2	# of Best Flow
1	Open	Closed	20
2	Open	Open	22
3	Closed	Open	56
4	Closed	Closed	2

Based on the information above, which apparatus is more like the desired model? By how much?

Solution

With the 4 different valve configurations, multinomial distribution can be utilized to calculate the probability of a measurement. The goal here is to find the model with the highest probability density function, because that one is more like the desired model.

The probability of a measurement can be calculated as follows given that there are 4 different possible valve configurations (see section on pdf for more information on where this equation came from):

$$P\left(n_1, n_2, n_3, n_4\right) = \frac{N!}{n_1! n_2! n_3! n_4!} p_1^{n_1} p_2^{n_2} p_3^{n_3} p_4^{n_4}$$

where

- n_i = total number of measurements with the best flow rate from each valve configuration
- i = configuration
- p_i = probability of that configuration having the best flow
- N = Total # of observations

The calculation for the probability of each apparatus can be done similarly:

Apparatus 1:

$$P(42,37,21,0) = \frac{100!}{42!\,37!\,21!\,0!}0.305^{42}0.476^{37}0.218^{21}0.001^{0}$$
$$P(42,37,21,0) = 3.125 \times 10^{-4}$$

Apparatus 2:

$$P(20,22,56,2) = \frac{100!}{20!22!56!2!}0.305^{20}0.476^{22}0.218^{56}0.001^{2}$$
$$P(20,22,56,2) = 8.455\times10^{-18}$$

Based on the calculations above for both apparatuses, Apparatus 1 is more like the desired model. Apparatus 1 has a higher probability density function, based on the relative likelihood of each configuration flow.

MultinomialDistProb.nb

```
In[1]:= Prob[n1_, n2_, n3_, n4_, n5_, p1_, p2_, p3_, p4_, p5_] :=
          N[((n1 + n2 + n3 + n4 + n5)! / (n1! * n2! * n3! * n4! * n5!) * (p1^n1) *
              (p2^n2) * (p3^n3) * (p4^n4) * (p5^n5))]
        Prob[42, 37, 21, 0, 0, 0.305, 0.476, 0.218, 0.001, 1]
        Prob[20, 22, 56, 2, 0, 0.305, 0.476, 0.218, 0.001, 1]

Out[2]= 0.000312523

Out[3]= 8.45526×10^-18
```

This figure shows Mathematica code that can be used in order to solve the probability of a multinomial distribution. The n values are the number of occurrences of each outcome and the p values are the probabilities of eachout come. The function is set for a multinomial distribution with five different outcomes. However, it can be used for multinomial distributions with fewer outcomes by setting the unused n values to 0 and the unused p values to any number other than 0. This figure also shows the probabilities calculated from Apparatus 1 and Apparatus 2.

Example 13.10.4

A runaway reaction occurs when the heat generation from an exothermic reaction exceeds the heat loss. Elevated temperature increases reaction rate, further increasing heat generation and pressure buildup inside the reactor. Together, the uncontrolled escalation of temperature and pressure inside a reactor may cause an explosion.

The precursors to a runaway reaction - high temperature and pressure - can be detected by the installation of reliable temperature and pressure sensors inside the reactor. Runaway reactions can be prevented by lowering the temperature and/or pressure inside the reactor before they reach dangerous levels. This task can be accomplished by sending a cold inert stream into the reactor or venting the reactor.

Les Lloyd is a process engineer at the Miles Reactor Company that has been assigned to work on a new reaction process. Using historical data from all the similar reactions that have been run before, Les has estimated the probabilities of each outcome occurring during the new process. The potential outcomes of the process include all permutations of the possible reaction temperatures (low and high) and pressures (low and high). He has combined this information into the table below:

Outcome	Temperature	Pressure	Probability
1	high	high	0.013
2	high	low	0.267
3	low	high	0.031
4	low	low	0.689

Worried about risk of runaway reactions, the Miles Reactor Company is implementing a new program to assess the safety of their reaction processes. The program consists of running each reaction process 100 times over the next year and recording the reactor conditions during the process every time. In order for the process to be considered safe, the process outcomes must be within the following limits:

Outcome	Temperature	Pressure	Frequency
1	high	high	$n_1 = 0$
2	high	low	$n_2 \leq 20$
3	low	high	$n_3 \leq 2$
4	low	low	$n_4 = 100 - n_1 - n_2 - n_3$

Help Les predict whether or not the new process is safe by answering the following question: What is the probability that the new process will meet the specifications of the new safety program?

Solution

The probability of the safety guidelines being met is given by the following CDF expression:

$$P(n_1 = 0, n_2 \leq 20, n_3 \leq 2, n_4 = 100 - n_1 - n_2 - n_3) = \sum_{n_1=0}^{0} \sum_{n_2=0}^{20} \sum_{n_3=0}^{2} \frac{100!}{n_1! n_2! n_3! (100 - n_1 - n_2 - n_3)!} p_1^{n_1} p_2^{n_2} p_3^{n_3} p_4^{100 - n_1 - n_2 - n_3}$$

where

- N is the number of trials
- k is the number of possible outcomes
- n_i is the number of occurrence of outcome i
- p_i is the probability of seeing outcome i
- c_i is the maximum number of occurrence of outcome i

This CDF expression can be evaluated using the following commands in Mathematica:

```
<< Needs["MultivariateStatistics`"]
<< multinomial = MultinomialDistribution[100, {0.013, 0.267, 0.031, 0.689}]
<< CDF[multinomial, {0, 20, 2, 78}]
```

The result is:

$$P\left(n_{-}=0, n_2 \leq 20, n_3 \leq 2, n_4=1000-n_1-n_2-n_3\right)=0.00132705$$

Based on this probability calculation, it appears unlikely that this new process will pass the new safety guidelines.

Sage's Corner

www.youtube.com/v/aAlQpREhy5c
For slides of this presentation by Group Si:Media:Multinomial Distribution.ppt

References

- Woolf, Peter, Amy Keating, Christopher Burge, and Michael Yaffe. (2004). *Statistics and Probability Primer for Computational Biologists*. Massachusetts Institute of Technology.
- "Multinomial Distribution." Wikipedia. en.Wikipedia.org/wiki/Multinomial_distribution.
Perry,R.H.(1997). Perry's Chemical Engineer's Handbook, 7th Edition, McGraw-Hill. ISBN 978-0-07-049841-9
- *Multinomial Calculator: Online Statistical Calculator*. Accessed 27 Nov. 2007. stattrek.com/Tables/multinomial.aspx#calculator

Contributors

- **Authors:** Hillary Kast, Andrew Kim, Adhi Paisoseputra, Stephanie Van Kirk
- **Stewards**: Gillian Berberich, Katie Feldt, Christopher Mark, Jason Wong

This page titled 13.10: Multinomial Distributions is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

13.11: Comparisons of two means

Engineers often must compare data sets to determine whether the results of their analyses are statistically equivalent. A sensor outputs a series of values and the engineer must both determine whether the sensor is precise and whether the values are accurate according to a standard. To make this evaluation, statistical methods are used. One method compares the probability distributions and another uses the Students t-test on two data sets. Microsoft Excel also has functions that perform the t-test that output a fractional probability to evaluate the null hypothesis (Basic Statistics).

Distributions

General Distributions

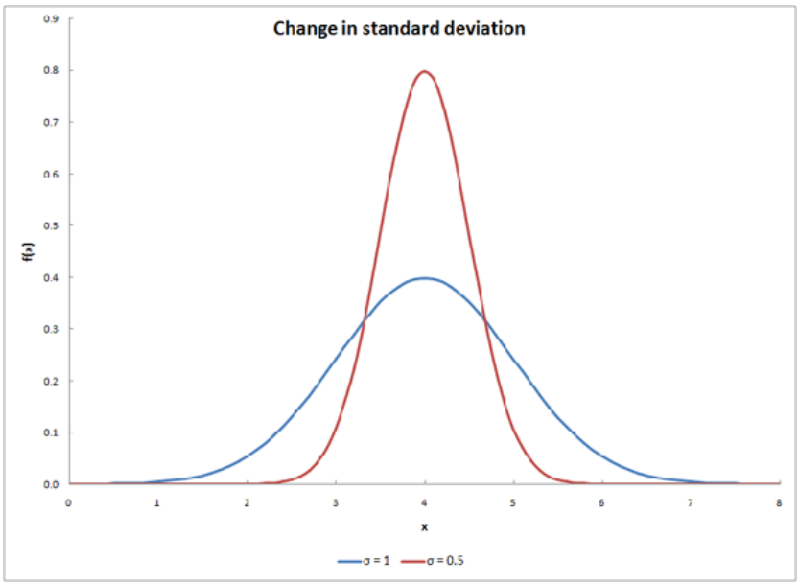
Distributions are governed by the probability density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

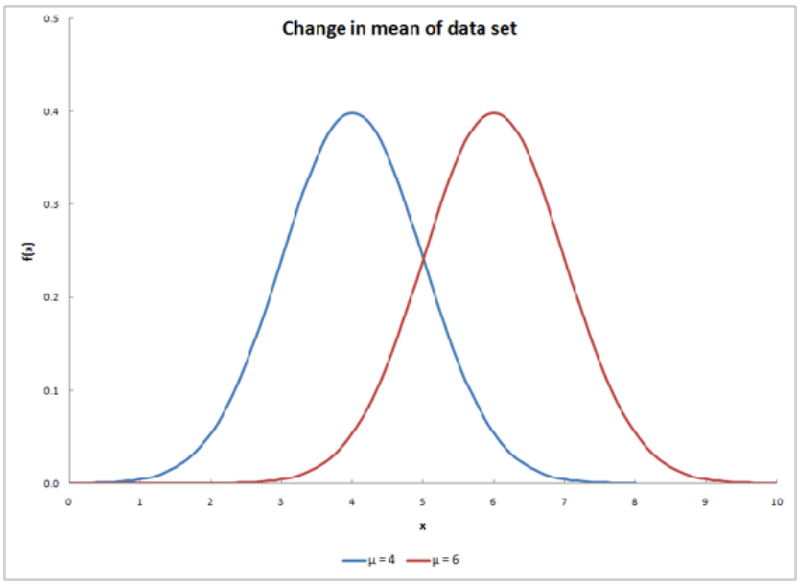
where

- σ is the standard deviation of a data set
- σ is the mean of the data set
- x is the input value

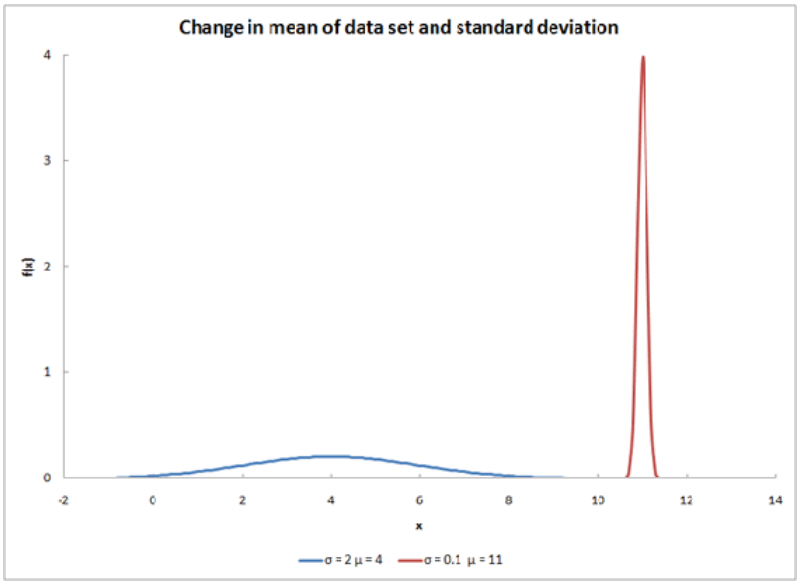
This equation gives a typical bell curve. Changing σ will alter the shape of the curve, as shown in the graph below.



Changing μ will shift the curve along the x-axis, as shown below.



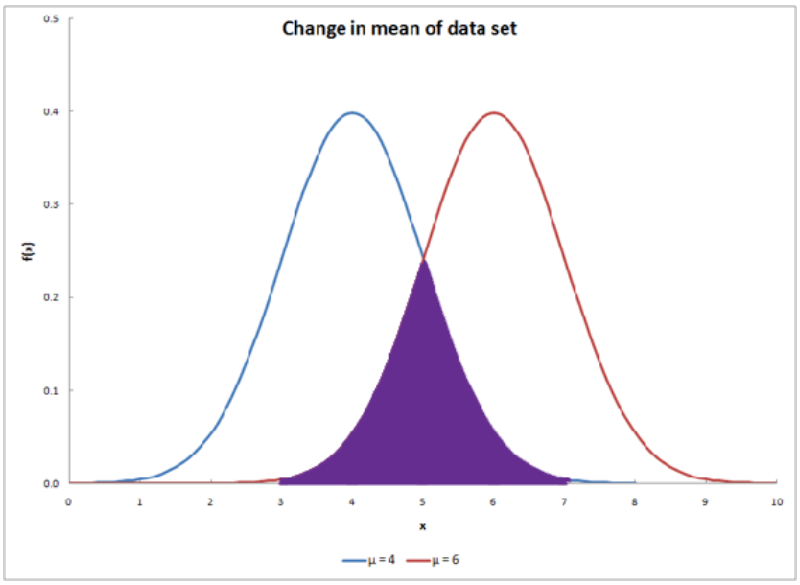
Changing both variables will have a result similar to the graph below.



For a more in-depth look at distributions, go to (Distributions).

Overlapping Distributions

The overlap between two distribution curves helps determine the probability of the two data sets being from the same distribution. The probability value increases as overlapping increases. The overlap is shown as the purple region in the graph below.



Comparison of Two Means

Probability

The similarity of two data sets can be determined by finding the probability of overlap. This is illustrated by the following equation:

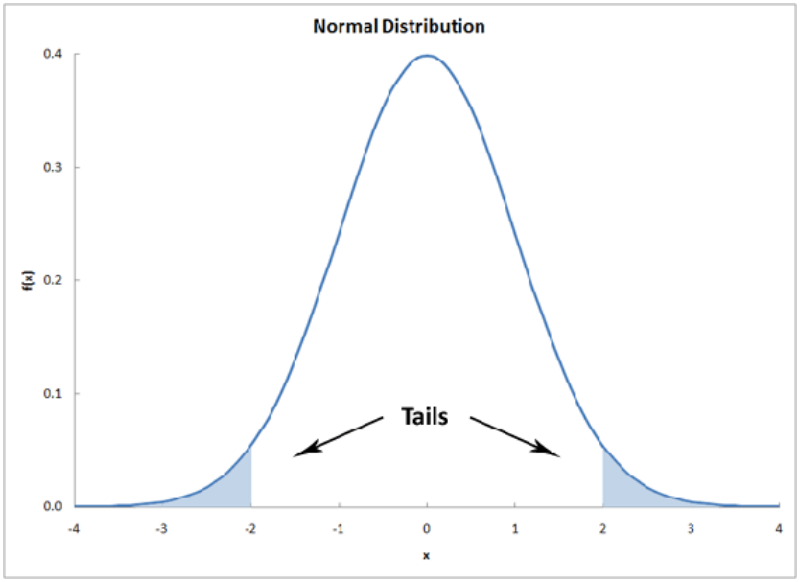
$$P(\text{ overlap }) = \int_{-\infty}^{\infty} \min \begin{Bmatrix} p_s\,(k \mid \theta_s) \\ p_o\,(k \mid \theta_0) \end{Bmatrix}$$

The functions contained within the integral are the probability distributions of each respective data set. The equation sums the lesser probability distribution from each data set. After using this equation, the solution will be a value with a range between 0 and 1 indicating the magnitude of overlapping probability. A value of 0 demonstrates that the two data sets do not overlap at all. A value of 1 demonstrates that the two data sets completely overlap. Values between 0 and 1 give the extent of overlap between the two data sets. This probability is not the same as the confidence interval that can be computed with t-tests.

Student's T-Test

The Student's t-test is extremely useful for comparing two means. There are various versions of the student t-test, depending on the context of the problem. Generally, the test quantifies the signal to noise ratio - where the signal is the difference of means and the noise is a function of the error around the means. If the signal is large and the noise is small, one can be more confident that the difference between the means is "real" or significant. To prove a significant difference, we need to disprove the null hypothesis. The "null hypothesis" (Ho) is that there is no difference between the two means. If we are able to disprove the "null hypothesis" then we will be able to say that the two groups are statistically different within a known level of confidence.

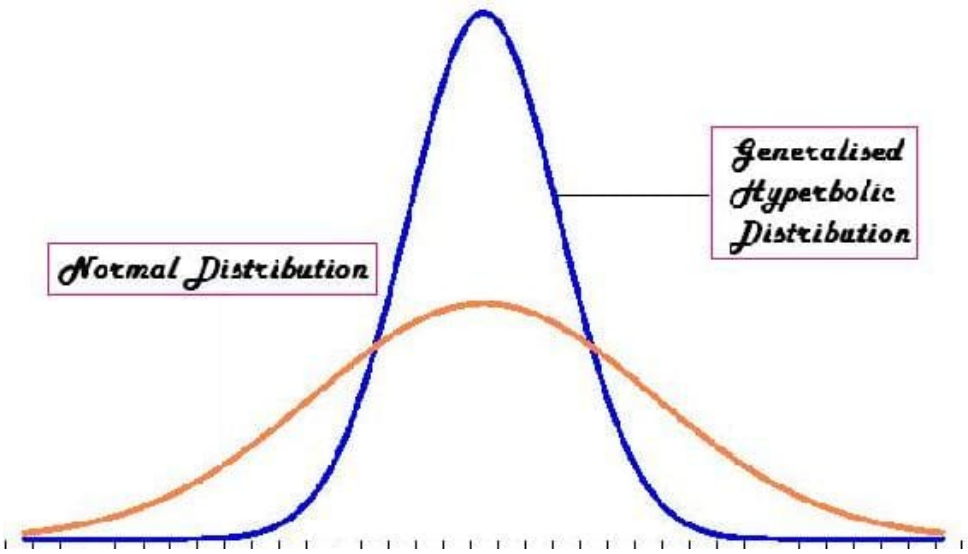
The extremities at either ends of a probability distribution are referred to as the "tails" of the distribution. The assessment of the significance of a calculated t-value will depend upon whether or not both tails of the distribution need to be considered. This will depend on the form of the null hypothesis. If your null hypothesis is an equality, then the case where one mean is larger *and* smaller must be considered; i.e. only one tail of the distribution should be accounted for. Conversely, if the null hypothesis is an inequality, then you are only concerned with the domain of values for a mean either less than *or* greater than the other mean; i.e. both tails of the distribution should be accounted for.



More Info: the student t-distribution

The t-distribution is the resulting probability distribution with a small sample population. This distribution is the basis for the t-test, in order to find the statistical significance between two means of data. This distribution is in the form of a generalised hyperbolic function (Which goes into details that would only clutter here. For more information, the Wikipedia site holds a lot of information on the subject: en.Wikipedia.org/wiki/Generalised_hyperbolic_distribution).

The t-distribution is commonly used when the standard deviation is unknown, or cannot be known (ie: a very small population set). When the data sets are large, or a standard deviation is assumed, the t-distribution is not very useful for a statistical analysis, and other methods of analysis should be used. An example of the t-distribution can be seen below:



Comparing Two Unknown True Means when Sample Standard Deviations are approximately Equal

The first accepted assumption is that when two sample means are being compared, the standard deviations are approximately equal. This method requires the average, standard deviation, and the number of measurements taken for each data set. The deviations are then pooled into one standard deviation. The equation for this t-test is as follows:

$$t = \frac{\text{Signal}}{\text{Noise}} = \frac{\bar{x}_1 - \bar{x}_2}{S_{\text{pooled}}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

where:

$$S_{\text{pooled}} = \sqrt{\frac{s_1^2\,(n_1-1)+s_2^2\,(n_2-1)}{n_1+n_2-2}}$$

where:

- $\overline{x}_1$ is the average of the first data set
- $\overline{x}_2$ is the average of the second data set
- n_1 is the number of measurements in the first data set
- n_2 is the number of measurements in the second data set
- s_1 is the standard deviation of the first data set
- \(s_2) is the standard deviation of the second data set
- t is a result of the t-test; it relates to values from the Student's t-distribution

Also note that the variance is defined as the square of the standard deviation.

Using t-distribution tables (sample shown below), the confidence level for the two means can then be determined. This confidence level determines whether the two means are significantly different. The confidence level can be found with the degrees of freedom for the measurements and the t-value computed above. The degree of freedom is equal to two less than the total number of measurements from the two data sets, as shown below:

$$DOF = n_1 + n_2 - 2$$

The following table is an image of a t value table, which can also be found (here):

	Confidence Level %			
Degrees of Freedom	50	90	95	98
1	1.000	6.314	12.706	31.821
2	0.816	2.920	4.303	6.965
3	0.756	2.353	3.182	4.541
4	0.741	2.132	2.776	3.747
5	0.727	2.015	2.571	3.365
6	0.718	1.943	2.447	3.143
7	0.711	1.895	2.365	2.998
8	0.706	1.860	2.306	2.896
9	0.703	1.833	2.262	2.821
10	0.700	1.812	2.228	2.764
15	0.691	1.753	2.131	2.602
20	0.687	1.725	2.086	2.528
25	0.684	1.708	2.060	2.485
30	0.683	1.697	2.042	2.457
40	0.681	1.684	2.021	2.423
60	0.679	1.671	2.000	2.390
120	0.677	1.658	1.980	2.358
∞	0.674	1.645	1.960	2.326

For example, if you had two data sets totaling 10 measurements and you calculated a t-value of 2.305 the confidence level would be 95%. This means that there is a 95% chance that the two data sets are statistically different and only a 5% chance that the two data sets are statistically similar. Also, degrees of freedom between the values listed on the table can be found by interpolating between those two values.

Note that there are some drawbacks when evaluating two means to see if they are significant or not. This problem mainly stems from the standard deviation. If say a set of values has a certain mean x, but the standard deviation was high due to the fact that some numbers in may have been greatly out of the range of the mean. This standard deviation may imply, from the student's t-test, that the mean x is significantly different from the mean of another set of data, when in actuality it may not seem that different. Hence, this must be taken into account when comparing two means using the student's t-test.

Comparing Two Unknown True Means (**μ_1 = ?** and **μ_2 = ?**) with Known True Unequal Standard Deviations (

$$sigma_1 \neq \sigma_2$$

)

The z-test is used when the difference between a sample mean and the population mean is large enough to be statistically significant. The t-test and z-test are essentially the same but in the z-test the actual population means (μ) and standard deviations(σ) are known. Since the estimate for difference in standard deviation used here is biased, two sample z-tests are rarely used.

The two sample z-statistic is described by:

$$z = \frac{\text{Signal}}{\text{Noise}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

where:

- $\bar{x}_1$ is the average of the first data set
- $\bar{x}_2$ is the average of the second data set
- n_1 is the number of measurements in the first data set
- n_2 is the number of measurements in the second data set
- σ_1 is the known standard deviation of the first population
- σ_2 is the known standard deviation of the second population

A different table is used to look up the probability of significance, please refer to Z-score table. If p < 0.05 (using a 95% confidence interval), we can declare a significant difference exists. The p-value is the probability that the observed difference between the means is caused by sampling variation, or the probability that these two samples came from the same population.

Comparing Two Unknown True Means (**μ_1 = ?** and **μ_2 = ?**) with Unknown True Standard Deviations (**σ_1 = ?** and **σ_2 = ?**)

This is known as the two sample t-statistic, which is used in statistical inference for comparing the means of two independent, normally distributed populations with unknown true standard deviations. The two sample t-statistic is described by:

$$t = \frac{\text{Signal}}{\text{Noise}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where:

- $overlinex_1$ is the average of the first data set
- $overlinex_2$ is the average of the second data set
- n_1 is the number of measurements in the first data set
- n_2 is the number of measurements in the second data set
- s_1 is the standard deviation of the first data set
- s_2 is the standard deviation of the second data set

Comparing the Mean of Differences for Paired Data

This is used in statistical analysis for the case where a single mean of a population occurs when two quantitative variables are collected in pairs, and the information we desire from these pairs is the difference between the two variables.

Two examples of paired data:

1. Each unit is measured twice. The two measurements of the same data are taken under varied conditions.
2. Similar units are paired prior to the experiment or run. During the experiment or run each unit is placed into different conditions.

The mean of differences t-statistic is described by:

$$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$$

where:

- $\bar{d}$ is the the mean of the differences for a sample of the two measurements
- s_d is the standard deviation of the sampled differences
- n is the number of measurements in the sample

Summary of Two Sample Mean Tests

Tests of Two Population Means			
Case	USE	Test Statistic	Formula
I	known σ's , and normal populations	z_0	$z = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$
II	unknown σ's , and normal populations (unequal variances)	t_c	$t = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$
III	unknown σ's , and normal populations (equal variances)	t_o	$t = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{s_p^2\left(\frac{1}{m} + \frac{1}{n}\right)}}$
IV	Paired t-test	t_0	$t = \frac{\bar{d} \quad \Delta_0}{s_d / \sqrt{n}}$

Excel Method

Instead of using the t-distribution tables to interpolate values, it is often easier to use built-in tools in Excel. The following three functions can be used for most of the common situations encountered when comparing two means:

- The TDIST function is useful when you have calculated a t-value and you want to know the probability that the t-value is significant.
- The TINV function is useful when you know the probability of significance you are interested in, and you desire the t-value (essentially the reverse of the TDIST function). This is helpful if you are designing an experiment and would like to determine the number of experimental runs needed to test for the difference of two means.
- The TTEST function is useful if you have two sets of data and you would like to know the probability that the mean of the two data sets are significantly different.

TDIST Function

The TDIST function has the syntax "=TDIST(x,deg_freedom,tails)"

Where:

- **x** is the t-value of the statistic
- **deg_freedom** is the number of degrees of freedom of the t-statistic. For comparing sample means $\overline{x}_o$ and $\overline{x}_s$, with sample

sizes n_o and n_s respectively, has $n_o + n_s - 2$ degrees of freedom.
- **tails** is the number of tails to be summed for probability. If null hypothesis is an equality, 2 tails will be used. If the null hypothesis is an inequality, 1 tail will be used.

The output of the function is the fractional probability of the Students t-distribution. For example, if the function returned a value of 0.05, this would correspond to a 95% or equivalently (1 - 0.05) confidence level for rejecting the null hypothesis.

TINV Function

The TINV function has the syntax "=TINV(probalility, deg_freedom)"

where:

- **probability** is the fractional probability of the Students t-distribution. This is identical to the output of the "TDIST" function.
- **deg_freedom** is the number of degrees of freedom of the t-statistic. For comparing sample means $\overline{x}_o$ and $\overline{x}_s$, with sample sizes n_o and n_s respectively, has $n_o + n_s - 2$ degrees of freedom.

The output of the function is the t-value of the Student's t-distribution.

TTEST Function

The TTEST function has the syntax "=TTEST(array 1, array 2, tails, type)"

where:

- **array 1** is the first data set
- **array 2** is the second data set
- **tails** is the number of tails to be summed for probability (1 or 2). If null hypothesis is an equality 2 tails will be used, if the null hypothesis is an inequality 1 tail will be used.
- **type** is the type of t-test to be performed the values that correspond to each type of test are listed below.

If type equals | This test is performed 1 | Paired 2 | Two-sample equal variance (homoscedastic) 3 | Two-sample unequal variance (heteroscedastic)

For our purposes we will only be concerned with type = 3. This corresponds to unequal variance (independent data sets). The other two types are useful and may prove interesting for the curious, but are beyond our scope.

Alternatively if you are not fond of Excel, a website located here will do the TTEST calculation for you.

The output of the function is the fractional probability of the Student's t-distribution. For example, if the function returned a value of 0.05, this would correspond to a 95% or equivalently (1 - 0.05) confidence level for rejecting the null hypothesis.

This function is very useful when the data-sets of the two means to be compared are known.

Worked out Example 1

You are a product quality engineer at a company that manufactures powdered laundry detergent in 100 ounce boxes. You are in charge of determining whether the product meets the specifications the company promises the customer. In the past, 100 samples (n_o) of the process established normal process conditions. The data is given below. Note that weighing samples using an imprecise scale does not affect your results. Also assume that the standard deviations are close enough to pool them.

Sample	Mass (oz)	Sample	Mass (oz)	Sample	Mass (oz)	Sample	Mass (oz)
1	101.5323	26	103.6760	51	102.8321	76	102.2068
2	102.6114	27	99.4326	52	103.1561	77	100.0458
3	98.5144	28	101.0811	53	104.9740	78	99.1730
4	102.5376	29	103.0467	54	102.8182	79	101.1648
5	99.6549	30	104.5012	55	104.6667	80	102.7247
6	103.2026	31	103.3905	56	102.7701	81	104.5956
7	102.9082	32	103.3171	57	100.3976	82	100.5764
8	103.5021	33	102.4609	58	104.4545	83	101.6134
9	102.3490	34	100.2123	59	101.3030	84	104.6165
10	104.0031	35	102.0465	60	103.6512	85	101.5386
11	98.9170	36	104.3919	61	101.2811	86	102.2764
12	103.1442	37	104.3028	62	104.8338	87	99.8757
13	103.9829	38	104.7788	63	100.0149	88	102.9348
14	102.0530	39	99.9478	64	101.5113	89	104.2985
15	100.8094	40	102.4054	65	103.6051	90	99.1000
16	102.4637	41	100.5217	66	104.5143	91	104.3817
17	100.1330	42	99.2822	67	100.6274	92	100.0774
18	101.1803	43	100.0083	68	104.7486	93	103.3923
19	102.4206	44	103.0967	69	103.0690	94	103.1890
20	99.2446	45	104.3340	70	101.9572	95	104.2288
21	101.1265	46	99.4872	71	102.6948	96	104.7568
22	104.5352	47	102.7668	72	104.4613	97	99.7750
23	103.0829	48	104.0406	73	102.1133	98	103.7144
24	101.3206	49	103.9086	74	99.6126	99	104.7967
25	104.7705	50	103.6373	75	104.3923	100	101.8893

You randomly sample 25 (n_s) of the products and you get the following data:

Sample	Mass (oz)
1	103.0484
2	103.2938
3	102.0150
4	100.2619
5	97.5217
6	97.8343
7	100.5561
8	101.7700
9	97.5080
10	99.6632
11	98.4071
12	98.6208
13	99.9031
14	97.6498
15	98.9091
16	98.0822
17	99.1118
18	102.5405
19	103.7024
20	102.6328
21	97.4841
22	99.9671
23	100.7755
24	99.9557
25	98.7856

The data can also be found in the first tab of this Excel file: Data File

Is the process running significantly different from normal with 95% confidence? In this example, do not use built-in Excel functions.

Solution

Using the data, the normal process produces an average $(\overline{x}_o)$ product mass of 102.37 and a variance $\left(s_o^2\right)$ of 3.17.

The sample has an average mass $(\overline{x}_s)$ of 100.00 with a variance $\left(s_s^2\right)$ of 4.00.

These were calculated using the AVERAGE function and VAR function in Excel. See the second tab of this Excel file for further explanation: Example 1

The null hypothesis is that the means are identical and the difference between the two means is due purely to chance.

Using the t-test:

$$t = \frac{\bar{x}_o - \bar{x}_s}{\sqrt{\frac{s_o^2(n_o-1)+s_s^2(n_s-1)}{n_o+n_s-2}}}\sqrt{\frac{n_o n_s}{n_o+n_s}}$$

$$t = \frac{102-100}{\sqrt{\frac{3(100-1)+4(25-1)}{100+25-2}}}\sqrt{\frac{25*100}{25+100}} = 5.00$$

Looking at the table that is reprinted below for reference:

	Confidence Level %			
Degrees of Freedom	50	90	95	98
120	0.677	1.658	1.980	2.358
∞	0.674	1.645	1.960	2.326

The number of degrees of freedom is:

$$DOF = n_o + n_s - 2$$

$$DOF = 25 + 100 - 2 = 123$$

The t-value corresponding to 95% confidence level at 123 degrees of freedom is between 1.960 and 1.980. Since the calculated t-value, 5.00, is much greater than 1.980, the null hypothesis is rejected at the 95% confidence level.

We conclude that the two means are significantly different. Thus, the process is not running normally and it is time to troubleshoot to find the problems occurring in the system.

Worked out Example 2

Same problem as "Worked Out Example 1." Instead, use the TTEST function in Excel.

Solution

The solution can also be seen in the third tab of this Excel file: Example 2

Using the TTEST function, with tails = 2 and test type = 3, the function gives a T-test value of $4.78x10^{-6}$.

Since $4.78x10^{-6}$ is less than 0.05 (from our 95% confidence level) we can once again conclude that we can reject the null hypothesis.

Exercise 13.11.1

If there are two sets of data, one with 15 measurements and another with 47 measurements, how many degrees of freedom would you enter in the Excel functions?

a. 20
b. 47
c. 15

d. 60

Answer available at the end of this section.

Exercise 13.11.2

What does the statement "Accurate within 95% confidence" mean?

a. The average of the data set is statistically equal to the true value within 95%
b. The value is the average plus or minus 95%
c. The average is statistically equal to the true value within 5%

Answer available at the end of this section.

References

- "Comparison of Two Means." State.Yale.Edu. Yale. 19 Nov. 2006 <http://www.stat.yale.edu/Courses/1997-98/101/meancomp.htm>.
- Excel Help File. Microsoft 2006.
- Harris, Daniel C. Exploring Chemical Analysis. 3rd ed. New York: W. H. Freeman and Company, 2005. 77-151.
- Woolf, Peter, et al. Statistics and Probability Primer for Computational Biologists. Massachusetts Institute of Technology. BE 490/Bio 7.91. Spring 2004. 52-68.
- "Z-test." Wikipedia. en.Wikipedia.org/wiki/Z-test

Answers

Exercise 13.11.1

The number of degrees of freedom is calculated as follows:

$$DOF = n_1 + n_2 - 2$$

where n_1 is the number of measurements in the first data set and n_2 is the number of measurements in the second data set.

Therefore,

$$DOF = 15 + 47 - 2 = 60$$

the answer is **d**.

Exercise 13.11.2

c

This page titled 13.11: Comparisons of two means is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Daniel Katzman, Jason Noelanders, Jessica Moreno, & Mark Winston-Galant via source content that was edited to the style and standards of the LibreTexts platform.

13.12: Factor analysis and ANOVA

First invented in the early 1900s by psychologist Charles Spearman, **factor analysis** is the process by which a complicated system of many variables is simplified by completely defining it with a smaller number of "factors." If these factors can be studied and determined, they can be used to predict the value of the variables in a system. A simple example would be using a person's intelligence (a factor) to predict their verbal, quantitative, writing, and analytical scores on the GRE (variables).

Analysis of variance (ANOVA) is the method used to compare continuous measurements to determine if the measurements are sampled from the same or different distributions. It is an analytical tool used to determine the significance of factors on measurements by looking at the relationship between a quantitative "response variable" and a proposed explanatory "factor." This method is similar to the process of comparing the statistical difference between two samples, in that it invokes the concept of hypothesis testing. Instead of comparing two samples, however, a variable is correlated with one or more explanatory factors, typically using the F-statistic. From this F-statistic, the P-value can be calculated to see if the difference is significant. For example, if the P-value is low (P-value<0.05 or P-value<0.01 - this depends on desired level of significance), then there is a low probability that the two groups are the same. The method is highly versatile in that it can be used to analyze complicated systems, with numerous variables and factors. In this article, we will discuss the computation involved in *Single-Factor*, *Two-Factor: Without Replicates*, and *Two-Factor: With Replicates* ANOVA. Below, is a brief overview of the different types of ANOVA and some examples of when they can be applied.

Overview and Examples of ANOVA Types

ANOVA Types

Single-Factor ANOVA (One-Way):

One-way ANOVA is used to test for variance among two or more independent groups of data, in the instance that the variance depends on a single factor. It is most often employed when there are at least three groups of data, otherwise a t-test would be a sufficient statistical analysis.

Two-Factor ANOVA (Two-Way):

Two-way ANOVA is used in the instance that the variance depends on two factors. There are two cases in which two-way ANOVA can be employed:

- Data *without replicates*: used when collecting a single data point for a specified condition
- Data *with replicates*: used when collecting multiple data points for a specified condition (the number of replicates must be specified and must be the same among data groups)

When to Use Each ANOVA Type

- Example: There are three identical reactors (R1, R2, R3) that generate the same product.

- One-way ANOVA: You want to analyze the variance of the product yield as a function of the reactor number.
- Two-way ANOVA without replicates: You want to analyze the variance of the product yield as a function of the reactor number and the catalyst concentration.
- Two-way ANOVA with replicates: For each catalyst concentration, triplicate data were taken. You want to analyze the variance of the product yield as a function of the reactor number and the catalyst concentration.

ANOVA is a Linear Model

Though ANOVA will tell you if factors are significantly different, it will do so according to a linear model. ANOVA will always assumes a linear model, it is important to consider strong nonlinear interactions that ANOVA may not incorporate when determining significance. ANOVA works by assuming each observation as overall mean + mean effect + noise. If there are non-linear relationship between these (for example, if the difference between column 1 and column 2 on the same row is that column2 = column1^2), then there is the chance that ANOVA will not catch it.

Key Terms

Before further explanation, please review the terms below, which are used throughout this Wiki.

Term or Variable	Description
Factor	Characteristics differentiating populations from one another
Group	Sample set that is influenced by the same factor
Levels	Populations (or treatments)
n_j	Number of elements in group j
x_{ji}	Element within group j
s_j^2	Variance within group j
N	Total number of samples
k	Total number of groups
$\bar{x}_j$	Mean of group j
$\bar{x}$	Mean of all samples within all groups
SSGroups	Sum of squares between groups
SSE	Sum of squared errors within groups
MSGroups	Mean square between groups
MSE	Mean square error within groups
SSTotal	Total sum of squares
df	Degrees of freedom

Comparison of Sample Means Using the F-Test

The F-Test is the ratio of the sample variances. The F-statistic and the corresponding F-Test are used in single-factor ANOVA for purposes of hypothesis testing.

Null hypothesis (H_o): all sample means arising from different factors are equal

Alternative hypothesis (H_a): the sample means are not all equal

Several assumptions are necessary to use the F-test:

- The samples are independent and random
- The distribution of the response variable is a normal curve within each population
- The different populations may have different means
- All populations have the same standard deviation

Introduction to the F-Statistic

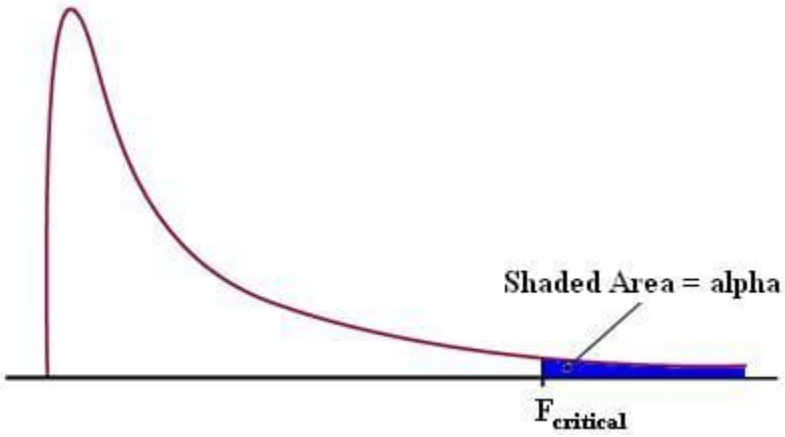
The F-statistic is the ratio of two variance estimates: the variance between groups divided by the variance within groups. The larger the F-statistic, the more likely it is that the difference between samples is due to the factor being tested, and not just the natural variation within a group. A standardized table can be used to find $F_{critical}$ for any system. $F_{critical}$ will depend on alpha, which is a measure of the confidence level. Typically, a value of alpha = 0.05 is used, which corresponds to 95% confidence. If $F_{observed} > F_{critical}$, we conclude with 95% confidence that the null hypothesis is false. For an explanation of how to read an F-Table, see *Interpreting the F-statistic* (below). In a similar manner, F tables can also be used to determine the p-value for a given data set. The p-value for a given data set is the probability that you could obtain this data set if the null hypothesis were true: that is, if the results were strictly due to chance. When H_o is true, the F-statistic has an F distribution.

F-Distributions

The F-distribution is important to ANOVA, because it is used to find the p-value for an ANOVA F-test. The F-distribution arises from the ratio of two Chi squared distributions. Thus, this family has a numerator and denominator degrees of freedom. (For information on the Chi squared test, click here.) Every function of this family has a a skewed distribution and minimum value of

zero.

Figure 1 - F distribution with alpha and $F_{critical}$ indicated



Single-Factor Analysis of Variance

In the case of single-factor analysis, also called single classification or one-way, a factor is varied while observing the result on the set of dependent variables. These dependent variables belong to a specific related set of values and hence, the results are expected to be related.

This section will describe some of the computational details for the F-statistic in one-way ANOVA. Although these equations provide insight into the concept of analysis of variance and how the F-test is constructed, it is not necessary to learn formulas or to do this analysis by hand. In practice, computers are always used to do one-way ANOVA

Setting up an Analysis of Variance Table

The fundamental concept in one-way analysis of variance is that the variation among data points in all samples can be divided into two categories: variation between group means and variation between data points in a group. The theory for analysis of variance stems from a simple equation, stating that the total variance is equal to the sum of the variance between groups and the variation within groups –

Total variation = variation between groups + variation within groups

An analysis of variance table is used to organize data points, indicating the value of a response variable, into groups according to the factor used in each case. For example, Table 1 is an ANOVA table for comparing the amount of weight lost over a three month period by dieters on various weight-loss programs.

Table 1 - Amount of weight lost by dieters on various programs over a 3 month period

Program 1	Program 2	Program 3
7	9	15
9	11	12
5	7	18
7		

A reasonable question is, can the type of program (a factor) be used to predict the amount of weight a dieter would lose on that program (a response variable)? Or, in other words, is any one program superior to the others?

Measuring Variation Between Groups

The variation between group means is measured with a weighted sum of squared differences between the sample means and the overall mean of all the data. Each squared difference is multiplied by the appropriate group sample size, ni, in this sum. This quantity is called **sum of squares between groups** or **SS Groups**.

$$\text{SSGroups } = n_1 (x_1 - x)^2 + n_2 (x_2 - x)^2 + ... + n_k (x_k - x)^2 = \sum_{\text{yrcups}} n_j (\bar{x}_{jj} - \bar{x}'^2$$

The numerator of the F-statistic for comparing means is called the **mean square between groups** or **MS Groups**, and it is calculated as -

$$\text{MSGroups } = \frac{\text{SSGroups}}{k-1}$$

Measuring Variation Within Groups

To measure the variation among data points within the groups, find the sum of squared deviations between data values and the sample mean in each group, and then add these quantities. This is called the **sum of squared errors**, **SSE**, or **sum of squares within groups**.

$$SSE = (n_1 - 1) s_1^2 + (n_2 - 1) s_2^2 + ... + (n_k - 1) s_k^2 = \sum_{\text{allgroups}} (n_j - 1) s_j^2$$

where

$$s_j^2 = \sum_{\text{groupj}} \frac{(x_{ij} - \bar{x}_j)^2}{n_j - 1} =$$

and is the variance within each group.

The denominator of the F-statistic is called the **mean square error**, **MSE**, or **mean squares within groups**. It is calculated as

$$MSE = \frac{SSE}{N-k} = \frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2 + ... + (n_k - 1) s_k^2}{n_1 + n_2 + ... + n_k - k}$$

MSE is simply a weighted average of the sample variances for the k groups. Therefore, if all n_i are equal, MSE is simply the average of the k sample variances. The square root of MSE (s_p), called the **pooled standard deviation,** estimates the population standard deviation of the response variable (keep in mind that all of the samples being compared are assumed to have the same standard deviation σ).

Measuring the Total Variation

The total variation in all samples combined is measured by computing the sum of squared deviations between data values and the mean of all data points. This quantity is referred to as the **total sum of squares** or SS Total. The total sum of squares may also be referred to as SSTO. A formula for the sum of squared differences from the overall mean is

$$\text{SSTotal } = \sum_{\text{values}} (x_{ij} - \bar{x})^2$$

where x_{ij} represents the jth observation within the ith group, and $\bar{x}$ is the mean of all observed data values. Finally, the relationship between SS Total, SS Groups, and SS Error is

SS Total = SS Groups + SS Error

Overall, the relationship between the total variation, the variation between groups, and the variation within a group is illustrated by Figure 2.

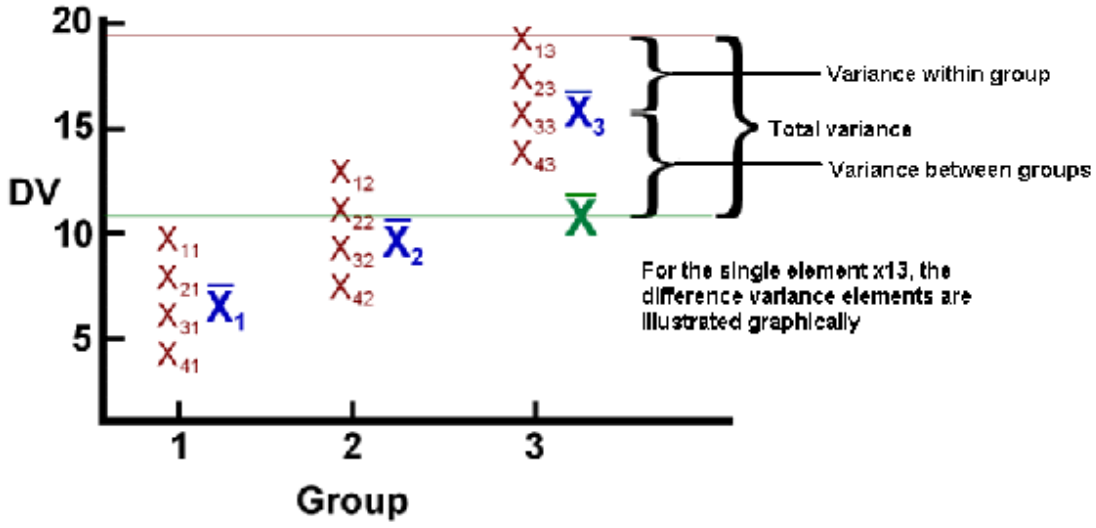


Figure 2: Visual illustration of variance in ANOVA, adapted from www.uwsp.edu/psych/stat/12/anova-1w.htm

A general table for performing the one-way ANOVA calculations required to compute the F-statistic is given below

Table 2 - One-Way ANOVA Table

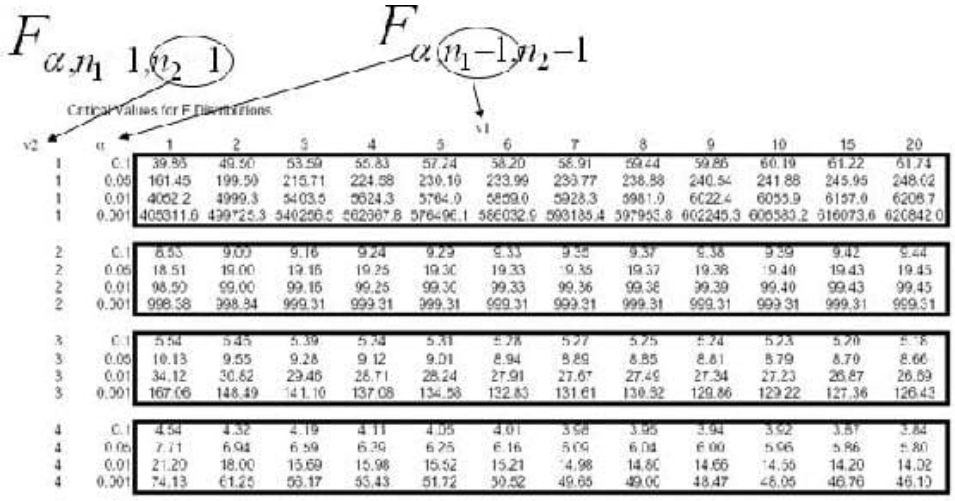
Source	Degrees of Freedom	Sum of Squares	Mean Sum of Squares	F-Statistic
Between groups	k-1	$SSGroups = \sum_{groups} n_i(\bar{x}_i - \bar{x})^2$	$\frac{SSGroups}{k-1}$	$F = \frac{MSGroups}{MSE}$
Within groups (error)	N-k	$SSE = \sum_{groups} (n_i - 1)s_i^2$	$\frac{SSE}{N-k}$	
Total	N-1	$SSTO = \sum_{values} (x_{ij} - \bar{x})^2$		

Interpreting the F-statistic

Once the F-statistic has been found, it can be compared with a critical F value from a table, such as this one: F Table. This F table is calculated for a value of alpha = 0.05, indicating a 95% confidence level. This means that if $F_{observed}$ is larger than $F_{critical}$ from the table, then we can reject the null hypothesis and say with 95% confidence that the variance between groups is not due to random chance, but rather due to the influence of a tested factor. Tables are also available for other values of alpha and can be used to find a more exact probability that the difference between groups is (or is not) caused by random chance.

Finding the Critical F value

In this F Table, the first row of the F table is the number of degrees of between groups (number of groups - 1), and the first column is the number of degrees of freedom within groups (total number of samples - number of groups).



v2	α	1	2	3	4	5	6	7	8	9	10	15	20
1	0.1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	61.22	61.74
1	0.05	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	245.95	248.02
1	0.01	4052.2	4999.3	5403.5	5624.3	5764.0	5859.0	5928.3	5981.0	6022.4	6055.9	6157.0	6208.7
1	0.001	405311.0	499725.3	540266.5	562367.8	576496.1	586032.9	593185.4	597953.8	602245.3	605583.2	616073.6	620842.0
2	0.1	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.42	9.44
2	0.05	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.43	19.45
2	0.01	98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.43	99.45
2	0.001	998.58	998.84	999.31	999.31	999.31	999.31	999.31	999.31	999.31	999.31	999.31	999.31
3	0.1	5.54	5.45	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.20	5.18
3	0.05	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.70	8.66
3	0.01	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	26.87	26.69
3	0.001	167.06	148.49	141.10	137.08	134.58	132.83	131.61	130.62	129.86	129.22	127.36	126.43
4	0.1	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.87	3.84
4	0.05	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.86	5.80
4	0.01	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.20	14.02
4	0.001	74.13	61.25	56.17	53.43	51.72	50.52	49.65	49.00	48.47	48.05	46.76	46.10

For the diet example in Table 1, the degree of freedom between groups is (3-1) = 2 and and the degree of freedom within groups is (13-3) = 10. Thus, the critical F value is 4.10.

Computing the 95% Confidence Interval for the Population Means

It is useful to know the confidence interval at which the means of the different groups are reported. The general formula for calculating a confidence interval is $samplemean \pm Multiplier * Standarderror$. Because it is assumed that all populations have the same standard deviation $s_p = \sqrt{(MSE)}$ can be used to estimate the standard deviation within each group. Although the population standard deviation is assumed to be the same, the standard error and the multiplier may be different for each group, due to differences in group size and degrees of freedom. The standard error of a sample mean is inversely proportional to the square root of the number of data points within the sample. It is calculated as $SE(\bar{x}_i) = \frac{s_p}{\sqrt{(n_i)}}$. The multiplier is determined using a t-distribution where the degrees of freedom are calculated as df = N-k. Therefore,*Insertformulahere* the confidence interval for a population mean is $\bar{x}_i \pm t * \frac{s_p}{\sqrt{(n_i)}}$. More details on confidence intervals can be found in Comparison of two means

An example for using factor analysis is the following:

You have two assembly lines. Suppose you sample 10 parts from the two assembly lines. Ho: $s1^2 = s2x^2$ Ha: variances are not equal Are the two lines producing similar outputs? Assume a=0.05 $F_{.025,9,9} = 4.03$ $F_{1-.025,9,9} = ?$

Observation	Line1	Line2
1	0.61	-1.01
2	0.15	-0.36
3	-0.24	-0.34
4	-0.65	-0.54
5	-0.58	-0.13
6	-0.48	-0.78
7	-0.35	-0.90
8	-0.07	-0.22
9	-0.45	-0.72
10	-0.25	-0.88

n	10	10
Mean	-0.231	-0.588
StDev	0.380	0.312
Variance	0.144	0.097

Are variances different?

Test for Equal Variances

Bonferroni confidence intervals for standard deviations

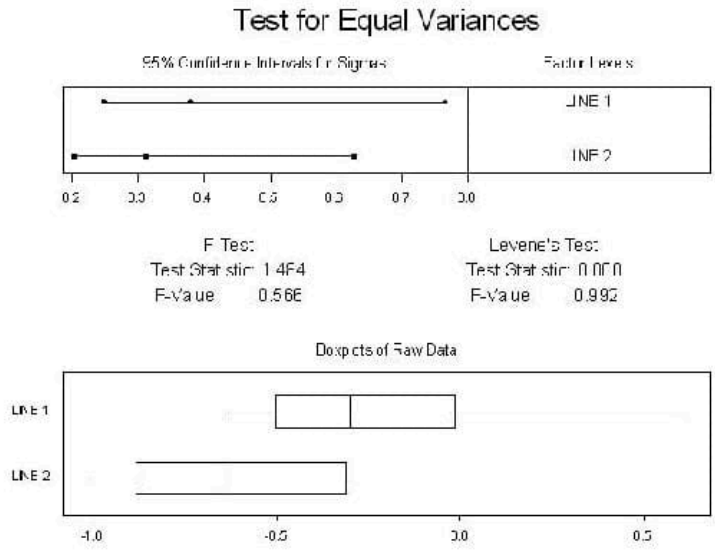
Lower	Sigma	Upper	N	Factor Levels
0.240633	0.380101	0.765391	10	LINE 1
0.204071	0.311976	0.628211	10	LINE 2

F Test (normal distribution)

Test Statistic: 1.484

P-Value : 0.566

How would we test if the means are different?



Two-Factor Analysis of Variance

A two factor or two-way analysis of variance is used to examine how two qualitative categorical variables (male/female) affect the mean of a quantitative response variable. For example, a psychologist might want to study how the type and volume of background music affect worker productivity. Alternatively, an economist maybe be interested in determining the affect of gender and race on mean income. In both of these examples, there is interest in the effect of each separate explanatory factor, as well as the combined effect of both factors.

Assumptions

In order to use the two-way ANOVA, the following assumptions are required:

- Samples must be independent.
- Population variances must be equal.
- Groups must have same sample size. The populations from which the samples were obtained must be normally distributed (or at least approximately so).
- The null hypothesis is assumed to be true.

The null hypothesis is as follows:

- The population means for the first factor have to be equal. This is similar to the one-way ANOVA for the row factor.
- The population means for the second factor must also be equal. This is similar to the one-way ANOVA for the column factor.
- There isn't an interaction between the two factors. This is similar to performing an independence test using contingency tables.

More simply, the null hypothesis implies that the populations are all similar and any differences in the populations are caused by chance, not by the influence of a factor. After carrying out two-way ANOVA it will be possible to analyze the validity of this assumption.

Terms Used in Two-Way ANOVA

The **interaction** between two factors is the most unique part of a two-way analysis of variance problem. When two factors interact, the effect on the response variable depends on the value of the other factor. For example, the statement *being overweight caused greater increases in blood pressure for men than women* describes an interaction. In other words, the effect of weight (factor) on blood pressure (response) depends on gender (factor).

The term **main effect** is used to describe the overall effect of a single explanatory variable. In the music example, the main effect

of the factor "music volume" is the effect on productivity averaged over all types of music. Clearly, the main effect may not always be useful if the interaction is unknown.

In a two-way analysis of variance, three F-statistics are constructed. One is used to test the statistical significance of the interaction, while the other two are used to test the significance of the two separate main effects. The p-value for each F-statistic is also reported--a p-value of <.05 is usually used to indicate significance. When an F-factor is found to have statistical significance, it is considered a main effect. The p-value is also used as an indicator to determine if the two factors have a significant interaction when considered simultaneously. If one factor depends strongly on the other, the F-statistic for the interaction term will have a low p-value. An example output of two-way analysis of variance of restaurant tip data is given in Table 4.

Table 4 - Two-Way Analysis of Variance of Restaurant Tipping Data

Source	DF	Adj SS	Adj MS	F-Statistic	P-Value
Message	1	14.7	14.7	.13	.715
Sex	1	2602.0	2602.0	23.69	0.00
Interaction	1	438.7	438.7	3.99	.049
Error	85	9335.5	109.8		
Total	88	12407.9			

In this case, the factors being studied are sex (male or female) and message on the receipt (:-) or none). The p-values in the last column are the most important information contained in this table. A lower p-value indicates a higher level of significance. Message has a significance value of .715. This is much greater than .05, the 95% confidence interval, indicating that this factor has no significance (no strong correlation between presence of message and amount of tip). The reason this occurs is that there is a relationship between the message and the sex of the waiter. The interaction term, which was significant with a value of p= 0.049, showed that drawing a happy face increased the tip for women but decreased it for men. The main effect of waiter sex (with a p-value of approximately 0) shows that there is a statistical difference in average tips for men and women.

Two-Way ANOVA Calculations

Like in one-way ANOVA analysis the main tool used is the square sums of each group. Two-way ANOVA can be split between two different types: with repetition and without repetition. With repetition means that every case is repeated a set number of times. For the above example that would mean that the :-) was given to females 10 times and males 10 times, and no message was given to females 10 times and males 10 times

Using the SS values as a start the F-statistics for two-way ANOVA with repetition are calculated using the chart below where a is the number of levels of main effect A, b is the number of levels of main effect B, and n is the number of repetitions.

Source	SS	DF	Adj MS	F-Statistic
Main Effect A	From data given	a-1	SS/df	MS(A)/MS(W)
Main Effect B	From data given	b-1	SS/df	MS(B)/MS(W)
Interaction Effect	From data given	(a-1)(b-1)	SS/df	MS(A*B)/MS(W)
Within	From data given	ab(n-1)	SS/df	
Total	sum of others	abn-1		

Without repetition means there is one reading for every case. For example is you were investigating whether or not difference in yield are more significant based on the day the readings were taken or the reactor that the readings were taken from you would have one reading for Reactor 1 on Monday, one reading for Reactor 2 on Monday etc... The results for two-way ANOVA without repetition is slightly different in that there is no interaction effect measured and the within row is replaced with a similar (but not equal) error row. The calculations needed are shown in the table below.

Source	SS	DF	MS	F-Statistic
Main Effect A	From data given	a-1	SS/df	MS(A)/MS(E)
Main Effect B	From data given	b-1	SS/df	MS(B)/MS(E)
Error	From data given	(a-1)(b-1)	SS/df	
Total	sum of others	ab-1		

These calculations are almost never done by hand. In this class you will usually use Excel or Mathematica to create these tables. Sections describing how to use these programs are found later in this chapter.

Other Methods of Comparison

Unfortunately, the conditions for using the ANOVA F-test do not hold in all situations. In this section, several other methods which do not rely on equal population standard deviations or normal distribution. It is important to realize that *no* method of factor analysis is appropriate if the data given is not representative of the group being studied.

Hypotheses About Medians

In general, it is best construct hypotheses about a population median, rather than the mean. Using the median accounts for the sample being skewed based on extreme outliers. Median hypotheses should also be used for dealing with ordinal variables (variables which are described only as being higher or lower than one other and do not have a precise value). When several populations are compared, the hypotheses are stated as -

H_0: Population medians are equal
H_a: Population medians are not all equal

Kruskal-Wallis Test for Comparing Medians

The **Kruskal-Wallis Test** provides a method of comparing medians by comparing the relative rankings of data in the observed samples. This test is therefore referred to as a rank test or non-parametric test because the test does not make any assumptions about the distribution of data.

To conduct this test, the values in the total data set are first ranked from lowest to highest, with 1 being lowest and N being highest. The ranks of the values within each group are averaged, and the test statistic measures the variation among the average ranks for each group. A p-value can be determined by finding the probability that the variation among the set of rank averages for the groups would be as large or larger as it is if the null hypothesis is true. More information on the Kruskal-Wallis test can be found [here].

Mood's Median Test for Comparing Medians

Another nonparametric test used to compare population medians is **Mood's Median Test**. Also called the Sign Scores Test, this test involves multiple steps.

1. Calculate the median (M) using all data points from every group in the study
2. Create a contingency table as follows

	A	B	C	Total
Number of values greater than M				
Number of values less than or equal to M				

3. Calculate the expected value for each data set using the following formula:

$$\text{expected} = \frac{(\text{ rowtotal })(\text{ columntotal })}{\text{grandtotal}}$$

4. Calculate the chi-square value using the following formula

$$\chi = \frac{(\text{ actual - expected })^2}{\text{expected}}$$

A chi-square statistic for two-way tables is used to test the null hypothesis that the population medians are all the same. The test is equivalent to testing whether or not the two variables are related.

ANOVA and Factor Analysis in Process Control

ANOVA and factor analysis are typically used in process control for troubleshooting purposes. When a problem arises in a process control system, these techniques can be used to help solve it. A factor can be defined as a single variable or simple process that has an effect on the system. For example, a factor can be temperature of an inlet stream, flow rate of coolant, or the position of a specific valve. Each factor can be analyzed individually to determine the effect that changing the input has on the process control system as a whole. The input variable can have a large, small, or no effect on what is being analyzed. The amount that the input variable affects the system is called the “factor loading”, and is a numerical measure of how much a specific variable influences the system or the output variable. In general, the larger the factor loading is for a variable the more of an affect that it has on the output variable.

A simple equation for this would be:

$$Output = f_1 * input_1 + f_2 * input_2 + ... + f_n * input_n$$

where f_n is the factor loading for the nth input.

Factor analysis is used in this case study to determine the fouling in an alcohol plant reboiler. This article provides some additional insight as to how factor analysis is used in an industrial situation.

Using Mathematica to Conduct ANOVA

Mathematica can be used for one-way and two-way factor anaylses. Before this can be done, the ANOVA package must be loaded into Mathematica using the following command:

Needs["ANOVA`"**]**

Once this command is executed, the 'ANOVA' command can be utilized.

One-Way Factor Analysis

The basic form of the 'ANOVA' command to perform a one-way factor analysis is as follows:

ANOVA[*data*]

An example set of data with five elements would look like:

ANOVA[

ParseError: EOF expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

]

An output table that includes the degrees of freedom, sum of the squares, mean sum of the squares, F-statistic, and the P-value for the model, error, and total will be displayed when this line is executed. A list of cell means for each model will be displayed beneath the table.

Two-Way Factor Analysis

The basic form of the 'ANOVA' command to perform a two-way factor analysis is as follows:

ANOVA[*data, model, vars*]

An example set of data with seven elements would look like:

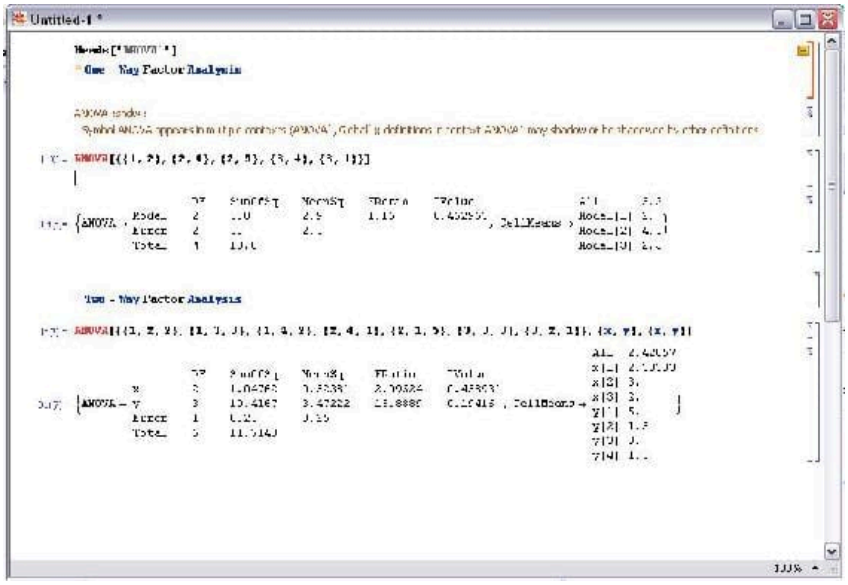
ANOVA[

ParseError: EOF expected (click for details)

```
Callstack:
    at (Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_
```

,{x,y},{x,y}]

An output table will appear similar to the one that is displayed in the one-way analysis except that there will be a row of statistics for each variable (i.e. x,y).



ANOVA in Microsoft Excel 2007

In order to access the ANOVA data analysis tool, install the package:

1. Click on the Microsoft Office button (big circle with office logo)
2. Click 'Excel Options'
3. Click 'Add-ins' on the left side
4. In the manage drop-down box at the bottom of the window, select 'Excel Add-ins'
5. Click 'Go...'
6. On the Add-Ins window, check the Analysis ToolPak box and click 'OK'

To use this package:

1. Click on the 'Data' tab and select 'Data Analysis'

2. Choose the desired ANOVA type- 'Anova: Single Factor', 'Anova: Two Factor with Replication', or 'Anova: Two Factor without Replication'(see note below for when to use replication)
3. Select the desired data points including data labels at top of the corresponding columns. Make sure the box is checked for 'Labels in first row' in the ANOVA parameter window.
4. Specify alpha in the ANOVA parameter window. Alpha represents the level of significance.
5. Output the results into a new worksheet.

NOTE: Anova: Two Factor with Replication is used in the cases where there are multiple readings for a single factor. For instance, the input below, there are 2 factors, control architecture and unit. This input shows how there are 3 readings corresponding to each control architecture (FB, MPC, and cascade). In this sense, the control architecture is replicated 3 times, each time providing different data relating to each unit. So, in this case, you would want to use the Anova Two Factor with Replication option.

control architecture			
	unit1	unit2	unit3
FB	52.59	51.48	49.63
FB	48.12	50.56	50.10
FB	50.00	49.22	49.83
MPC	49.67	50.61	50.67
MPC	50.21	51.33	50.44
MPC	49.84	51.81	51.41
cascade	51.29	49.23	50.85
cascade	50.91	48.47	49.95
cascade	50.04	47.70	51.16

Anova: Two Factor without Replication is used in cases where there is only one reading pertaining to a particular factor. For example, in the case below, each sample (row) is independent of the other samples since they are based on the day they were taken. Since multiple readings were not taken within the same day, the "without Replication" option should be chosen.

		Reactor yield is kg of glycerin per run				
sample	Date	Reactor 1	Reactor 2	Reactor 3	Reactor 4	Reactor 5
1	12/1/2007	107.173504	121.537341	94.427813	113.253739	100.40704
2	12/2/2007	111.720747	96.5592501	110.355107	126.534195	115.744327
3	12/3/2007	101.543829	108.285763	113.990532	102.178054	117.698256
4	12/4/2007	108.797073	113.160513	85.2385023	110.027039	96.109622
5	12/5/2007	96.3032079	111.141039	108.192291	118.733219	91.593933
6	12/6/2007	105.847343	91.3955177	88.5376488	102.178735	107.44279
7	12/7/2007	96.2074791	109.437155	120.697309	113.173557	109.97807
8	12/8/2007	111.925223	127.333877	100.12252	99.6354333	92.5296735
9	12/9/2007	97.5502824	97.7502335	108.549928	106.262861	84.4202995
10	12/10/2007	96.3221805	98.0003128	104.460937	115.837804	100.100045
11	12/11/2007	114.13772	87.6361313	109.119963	116.996506	103.318278

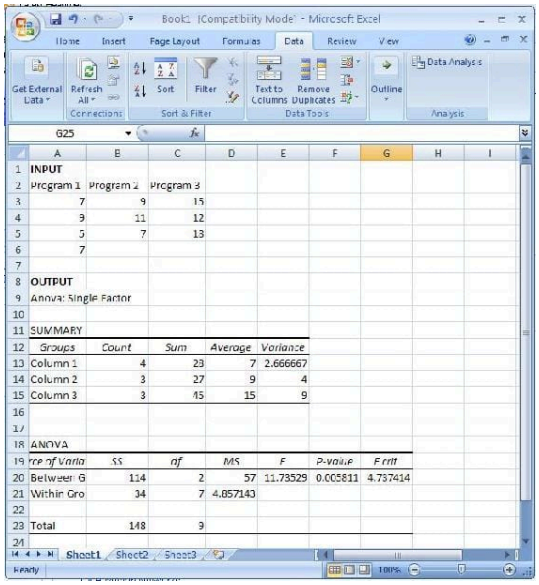
Excel outputs:

Summary:
1. Count- number of data points in a set
2. Sum- sum of the data points in a set
3. Average- mean of the data points in a set
4. Variance- standard deviation of the data points in a set

ANOVA:
1. Sum of squares (SS)
2. The degree of freedom (df)
3. The mean squares (MS)
4. F-statistic (F)

5. P-value
6. $F_{critical}$

See the figure below for an example of the inputs and outputs using Anova: Single Factor. Note the location of the Data Analysis tab. The data was obtained from the dieting programs described in Table 1. Since the F-statistic is greater than $F_{critical}$, the null hypothesis can be rejected at a 95% confidence level (since alpha was set at 0.05). Thus, weight loss was not random and in fact depends on diet type chosen.



INPUT								
Program 1	Program 2	Program 3						
7	9	15						
9	11	12						
5	7	13						
7								
OUTPUT								
Anova: Single Factor								
SUMMARY								
Groups	Count	Sum	Average	Variance				
Column 1	4	28	7	2.666667				
Column 2	3	27	9	4				
Column 3	3	45	15	9				
ANOVA								
rce of Varia	SS	df	MS	F	P-value	F crit		
Between Gr	114	2	57	11.73529	0.005811	4.737414		
Within Gro	34	7	4.857143					
Total	148	9						

Example 13.12.1

Determine the fouling rate of the reboiler at the following parameters:

- $T = 410K$
- $C_C = 16.7g\ /\ L$
- $R_T = 145min$

Which process variable has the greatest effect (per unit) on the fouling rate of the reboiler?

Note that the tables below are made up data. The output data for a single input was gathered assuming that the other input variables provide a negligible output. Although the factors that affect the fouling of the reboiler are similar to the ones found the the article linked in the "ANOVA and Factor Analysis in Process Controls" section, the data is not.

Temperature of Reboiler (K)	400	450	500
Fouling Rate (mg/min)	0.8	0.86	0.95

Catalyst Concentration (g/L)	10	20	30
Fouling Rate (mg/min)	0.5	1.37	2.11

Residence Time (min)	60	120	180
Fouling Rate (mg/min)	0.95	2.3	3.81

Solution

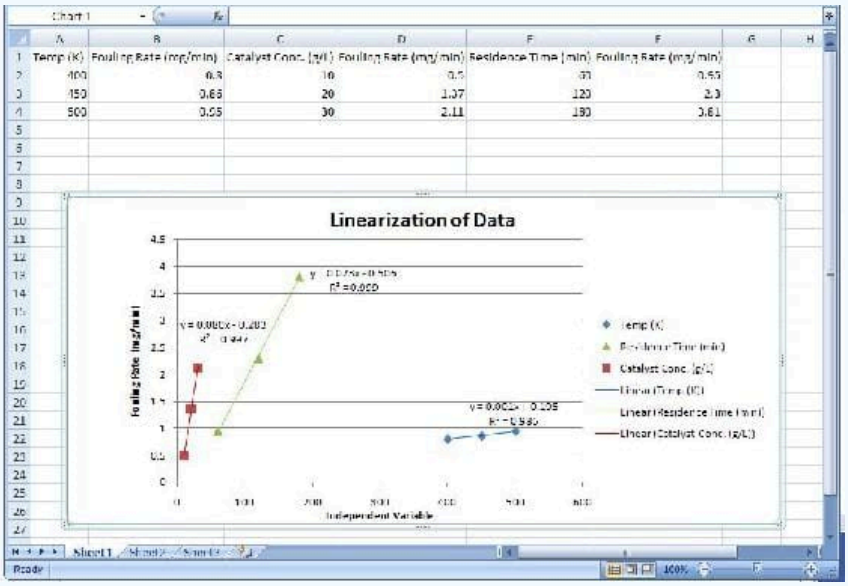
1) Determine the "factor loading" for each variable.

This can be done using any linearization tool. In this case, the factor loading is just the slope of the line for each set of data. Using Microsoft Excel, the equations for each set of data are the following:

Temperature of Reboiler
$y = 0.0015 * x + 0.195$
Factor loading: 0.0015

Catalyst Concentration
$y = 0.0805 * x - 0.2833$
Factor loading: 0.0805

Residence Time
$y = 0.0238 * x - 0.5067$
Factor loading: 0.0238



2) Determine the fouling rate for the given process conditions and which process variable affects the fouling rate the most (per unit). Note that the units of the factor loading value are **always** the units of the output divided by the units of the input.

Plug in the factor loading values into the following equation:
$Output = f_1 * input_1 + f_2 * input_2 + ... + f_n * input_n$

You will end up with:
$FoulingRate = 0.0015 * T + 0.0805 * C_C + 0.0238 * R_T$

Now plug in the process variables:
$FoulingRate = 0.0015 * 410 + 0.0805 * 16.7 + 0.0238 * 145$
$FoulingRate = 5.41 mg / min$

The process variable that affects the fouling rate the most (per unit) is the catalyst concentration because it has the largest factor loading value.

Example 13.12.2

The exit flow rate leaving a tank is being tested for 3 cases. The first case is under the normal operating conditions, while the second (A) and the third (B) cases are for new conditions that are being tested. The flow value of 7 (gallons /hour) is desired with a maximum of 10. A total of 24 runs are tested with 8 runs for each case. The tests are run to determine whether any of the new conditions will result in a more accurate flow rate. First, we determine if the new conditions A and B affect the flow rate. The results are as follows:

	recorded reading							
Case 1	7	4	6	8	6	6	2	9
Case 2	5	5	3	4	4	7	2	2
Case 3	2	4	7	1	2	1	5	5

	x_1	x_1^2	x_2	x_2^2	x_3	x_3^2
	7	49	5	25	2	4
	4	16	5	25	4	16
	6	36	3	9	7	49
	8	64	4	16	1	1
	6	36	4	16	2	4
	6	36	7	49	1	1
	2	4	2	4	5	25
	9	81	2	4	5	25
sum:	48	322	32	148	27	125
sum^2:	2304		1024		729	
M	6		4		3.375	

SS-total= (322+ 148 +125) - [(48 + 32 + 27)^2] / 2 = 117.96

SS-among= (2308/8 + 1024/8 + 729/8) = 30.08

SS-within= SS-total - SS-among : 87.88

Source	SS	df	MS	F
Among	30.08	2	15.04	3.59
Within	87.88	21	4.18	

The recorded values for the 3 cases are tabulated. Following this the values for each case are squared and the sums for all of these are taken. For the 3 cases, the sums are squared and then their means are found.

These values are used to help determine the table above (the equations give an idea as to how they are calculated). In the same way with the help of ANOVA, these values can be determine faster. This can be done using the mathematica explained above.

Conclusion:

$F_{critical}$ equals 3.4668, from an F-table. Since the calculated F value is greater than $F_{critical}$, we know that there is a statistically significant difference between 2 of the conditions. Thus, the null hypothesis can be rejected. However we do not know between which 2 conditions there is a difference. A post-hoc analysis will help us determine this. However we are able to confirmed that there is a difference.

Example 13.12.3

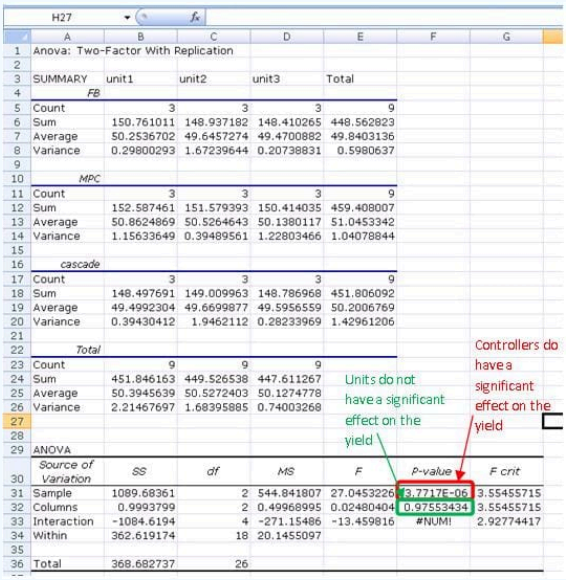
As the new engineer on site, one of your assigned tasks is to install a new control architecture for three different units. You test three units in triplicate, each with 3 different control architecture: feedback (FB), model predictive control (MPC) and cascade control. In each case you measure the yield and organize the data as follows:

control architecture			
	unit1	unit2	unit3
FB	52.59	51.48	49.63
FB	48.12	50.56	50.10
FB	50.00	49.22	49.83
MPC	49.67	50.61	50.67
MPC	50.21	51.33	50.44
MPC	49.84	51.81	51.41
cascade	51.29	49.23	50.85
cascade	50.91	48.47	49.95
cascade	50.04	47.70	51.16

Do the units differ significantly? Do the control architectures differ significantly?

Solution

This problem can be solved using ANOVA Two factor with replication analysis.



Anova: Two-Factor With Replication

SUMMARY	unit1	unit2	unit3	Total
FB				
Count	3	3	3	9
Sum	150.761011	148.937182	148.410265	448.562823
Average	50.2536702	49.6457274	49.4700882	49.8403136
Variance	0.29800293	1.67239644	0.20738831	0.5980637
MPC				
Count	3	3	3	9
Sum	152.587461	151.579393	150.414035	459.408007
Average	50.8624869	50.5264643	50.1380117	51.0453342
Variance	1.15633649	0.39489561	1.22803466	1.04078844
cascade				
Count	3	3	3	9
Sum	148.497691	149.009963	148.786968	451.806092
Average	49.4992304	49.6699877	49.5956559	50.2006769
Variance	0.39430412	1.9462112	0.28233969	1.42961206
Total				
Count	9	9	9	
Sum	451.846163	449.526538	447.611267	
Average	50.3945639	50.5272403	50.1274778	
Variance	2.21467697	1.68395885	0.74003268	

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	1089.68361	2	544.841807	27.0453226	3.7717E-06	3.55455715
Columns	0.9993799	2	0.49968995	0.02480404	0.97553434	3.55455715
Interaction	-1084.6194	4	-271.15486	-13.459816	#NUM!	2.92774417
Within	362.619174	18	20.1455097			
Total	368.682737	26				

Units do not have a significant effect on the yield

Controllers do have a significant effect on the yield

Exercise 13.12.1

ANOVA analysis works better for?

a. Non-linear models
b. Linear models
c. Exponential models
d. All of the above

Answer available at the end of this section.

Exercise 13.12.2

Two-Way ANOVA analysis is used to compare?

a. Any *two* sets of data
b. Two One-Way ANOVA models to each other
c. Two factors on their effect of the output
d. D. B and C

Answer available at the end of this section.

References

- Ogunnaike, Babatunde and W. Harmon Ray. *Process Dynamics, Modeling, and Control*. Oxford University Press. New York, NY: 1994.
- Uts, J. and R. Hekerd. *Mind on Statistics*. Chapter 16 - Analysis of Variance. Belmont, CA: Brooks/Cole - Thomson Learning, Inc. 2004.
- Charles Spearman. Retrieved November 1, 2007, from www.indiana.edu/~intell/spearman.shtml
- Plonsky, M. "One Way ANOVA." Retrieved November 13, 2007, from www.uwsp.edu/psych/stat/12/anova-1w.htm
- Ender, Phil. "Statistical Tables F Distribution." Retrieved November 13, 2007, from www.gseis.ucla.edu/courses/help/dist3.html
- Devore, Jay L. *Probability and Statistics for Engineering and the Sciences*. Chapter 10 - The Analysis of Variance. Belment, CA: Brooks/Cole - Thomson Learning, Inc. 2004.
- "Mood's Median Test (Sign Scores Test)" Retrieved November 29, 2008, from www.micquality.com/six_sigma_glossary/mood_median_test.htm

Answers

Exercise 13.12.1

B

Exercise 13.12.2

C

This page titled 13.12: Factor analysis and ANOVA is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Alexander Voice, Andrew Wilkins, Ibrahim Oraiqat, & Rohan Parambi via source content that was edited to the style and standards of the LibreTexts platform.

13.13: Correlation and Mutual Information

The application of control networks to engineering processes requires an understanding of the relationships between system variables. One would expect, for example, to find a relationship between the steam flow rate to a heat exchanger and the outlet stream temperature. How, then, are engineers to quantify the degree of relation between a pair of variables? Often times, investigations into the correlation and mutual information of data sets can provide the answer. Correlation analysis provides a quantitative means of measuring the strength of a linear relationship between two vectors of data. Mutual information is essentially the measure of how much "knowledge" one can gain of a certain variable by knowing the value of another variable. By utilizing these techniques, engineers can increase the selectivity and accuracy of their control systems.

Correlation

Population Correlation Coefficient

The correlation coefficient is an important measure of the relationship between two random variables. Once calculated, it describes the validity of a linear fit. For two random variables, X and Y, the correlation coefficient, ρ_{xy}, is calculated as follows:

$$\rho_{xy} = \frac{\operatorname{cov}(X,Y)}{\sigma_x \sigma_y}$$

That is, the covariance of the two variables divided by the product of their standard deviations. Covariance serves to measure how much the two random variables vary together. This will be a positive number if both variables consistently lie above the expected value and will be negative if one tends to lie above the anticipated value and the other tends to lie below. For a further description of covariance, as well as how to calculate this value, see Covariance on Wikipedia.

The correlation coefficient will take on values from 1 to -1. Values of 1 and -1 indicate perfect increasing and decreasing linear fits, respectively. As the population correlation coefficient approaches zero from either side, the strength of the linear fit diminishes. Approximate ranges representing the relative strength of the correlation are shown below. See the Correlation article on Wikipedia for more detailed information about the correlation theory establishing these ranges. The ranges also apply for negative values between 0 and -1.

Correlation

- Small: 0.10 to 0.29
- Medium: 0.30 to 0.49
- Large: 0.50 to 1.00

Sample Correlation Coefficient

A modified form of the expression for the correlation coefficient describes the linearity of a data sample. For n measurements of variables X and Y, the sample correlation coefficient is calculated as follows:

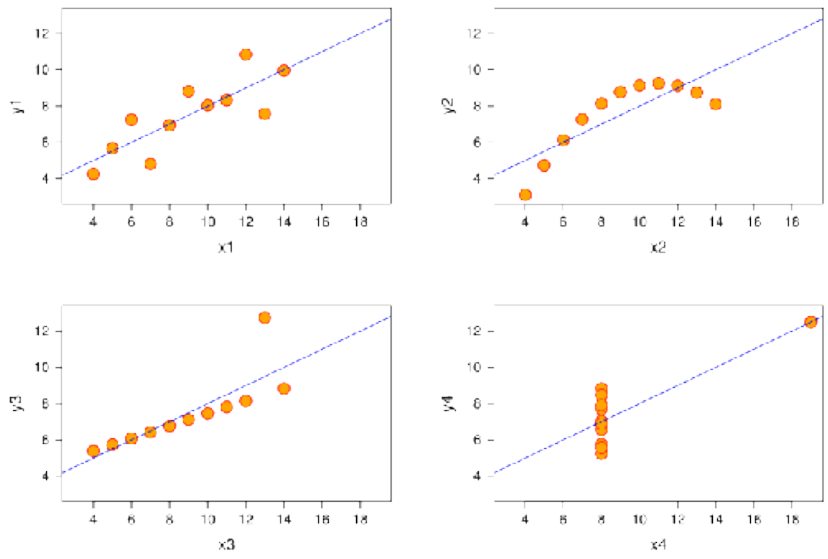
$$\rho_{xy} = \frac{\sum (x_i - \bar{x})\,(y_i - \bar{y})}{(n-1)\,s_x s_y}$$

where $\bar{x}$ and $\bar{y}$ are the sample averages, and s_x and s_y are the samples' standard deviations. This correlation coefficient will still take on values from 1 to -1, corresponding to the strength of linearity in the same way as the population correlation coefficient. This expression is often the more useful for process controls applications; an engineer can measure two variables, calculate their means and standard deviations, and ultimately determine the validity of a linear relationship between the two. This method is extremely useful when applied to a sample of experimental data that can be modeled by a normal distribution function. It is simpler and easier to use compared to the population correlation coefficient because it uses only well defined statistical inputs, mean and standard deviation, rather than the complex covariance function.

Correlation Coefficient Assumptions: Linearity, Normal Distribution

The correlation coefficient indicates the strength of a linear relationship between two variables with random distribution; this value alone may not be sufficient to evaluate a system where these assumptions are not valid.

The image below shows scatterplots of Anscombe's quartet, a set of four different pairs of variables created by Francis Anscombe. The four y variables have the same mean (7.5), standard deviation (4.12), correlation coefficient (0.81) and regression line (y = 3 + 0.5x).



[Image from en.Wikipedia.org/wiki/Correlation Correlation]

However, as can be seen on the plots, the distribution of the variables is very different.

- The scatterplot of y1 (top left) seems to exhibit random distribution. It corresponds to what one would expect when considering two variables linearly correlated and following normal distribution.
- The scatterplot of y2 (top right) does not exhibit random distribution; a plot of its residuals would not show random distribution. However, an obvious non-linear relationship between the two variables can be observed.
- The scatterplot of y3 (bottom left) seems to exhibit a near perfect linear relationship with the exception of one outlier. This outlier exerts enough influence to lower the correlation coefficient from 1 to 0.81. A Q test should be performed to determine if a data point should be retained or rejected.
- The scatterplot of y4 (bottom right) shows another example when one outlier is enough to produce a high correlation coefficient, even though the relationship between the two variables is not linear.

These examples indicate that the correlation coefficient alone should not be used to indicate the validity of a linear fit.

Engineering Applications

Measuring the correlation of system variables is important when selecting manipulated and measured control pairs. Say, for instance, that an engineer takes several measurements of the temperature and pressure within a reactor and notes that the data pairs have a correlation coefficient of approximately one. The engineer now understands that by applying a control system to maintain the temperature, he will also be controlling the reactor pressure by way of the linear relationship between the two variables.

Here we begin to see the role correlations play in defining MIMO versus SISO systems. Generally, engineers seek control systems with independent pairs of manipulated (u) and measured variables (x). If two inputs strongly influence a single output as measured by the correlation value, then the system will be difficult to decouple into a SISO controller. Correlation values closer to 0 for all but one manipulated variable means the system will be easier to decouple because here a single system input controls a single system output and there is limited effect on other process variables.

Correlation in Mathematica

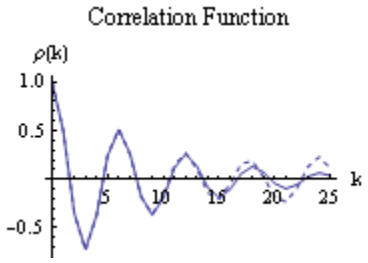
Correlation coefficients can be quickly calculated with the aid of Mathematica. To illustrate by example, suppose that an engineer measures a series of temperatures inside a bacterial chemostat and the corresponding protein concentration in the product. She measures temperatures of 298, 309, 320, 333, and 345 K corresponding to protein concentrations of 0.4, 0.6, 0.7, 0.85, and 0.9 weight percent protein (these values are for illustrative purposes and do not reflect the performance of an actual chemostat). The engineer enters these measurements into Mathematica as a pair of vectors as follows:

```
In[11]:= T = {{298}, {309}, {320}, {333}, {345}}
         W = {{0.4}, {0.6}, {0.7}, {0.85}, {0.9}}
```

where T is a vector containing the temperature measurements and C is a vector containing the concentration measurements. To determine the correlation coefficient, simply use the syntax below.

```
Correlation[T, W]
Out[13]= {{0.978202}}
```

In this case, the correlation coefficient of nearly 1 indicates a strong degree of linearity between the chemostat temperature and the protein concentration, informing the engineer that she can essentially control this concentration through temperature control. The sample correlation function in Mathematica uses an approximate solution with the mean and standard deviation of the data set, but the value is very close to the true correlation function value. This can be seen in the figure below as the dashed line represents the sample correlation function, which deviates slightly from the true correlation function. As the sample becomes larger (i.e. higher n) the approximation gets closer to the true value. This represents a limitation of the sample correlation function and demonstrates how Mathematica can be used to determine multiple correlations between sets of data.



Mutual Information

Mutual information (also referred to as transinformation) is a quantitative measurement of how much one random variable (Y) tells us about another random variable (X). In this case, information is thought of as a reduction in the uncertainty of a variable. Thus, the more mutual information between X and Y, the less uncertainty there is in X knowing Y or Y knowing X. For our purposes, within any given process, several parameters must be selected in order to properly run the process. The relationship between variables is integral to correctly determine working values for the system. For example, adjusting the temperature in a reactor often causes the pressure to change as well. Mutual information is most commonly measured in logarithms of base 2 (bits) but is also found in base e (nats) and base 10 (bans).

Explanation of Mutual Information

The mathematical representation for mutual information of the random variables A and B are as follows:

$$I(A;B) = \sum_{b \in B} \sum_{a \in A} p(a,b) * \log \left(\frac{p(a,b)}{p(a)p(b)} \right)$$

where,

- $p(a,b)$ is the joint probability distribution function of A and B.
- $p(a)$ is the marginal probability distribution function of A.
- $p(b)$ is the marginal probability distribution function of B.

For a review and discussion of probability distribution functions, please refer to Probability Distribution (Wikipedia).

When referring to continuous situations, a double integral replaces the double summation utilized above.

$$I(A;B,c) = \int_B \int_A p(a,b) * \log\left(\frac{p(a,b)}{p(a)p(b)}\right) dadb$$

where,

- $p(a,b)$ is the joint probability distribution function of A and B.
- $p(a)$ is the marginal probability distribution function of A.
- $p(b)$ is the marginal probability distribution function of B.

For a review and discussion of probability density functions, please refer to Probability Density Function (Wikipedia).

By performing the functions above, one measures the distance between the joint distribution/density functions of A and B. Since mutual information is always measuring a distance, the value obtained from this calculation will always be nonnegative and symmetric. Therefore,

$$I(A;B) \geq 0$$

$$I(A;B) = I(B;A)$$

In both cases, the information shared between both A and B is being measured. If A and B are independent of each other, then the difference between the two would be zero. In terms of chemical engineering processes, two independent variables would share no mutual information which means that one variable provides no information about the other variable. If A and B were identical, then all the information derived from obtaining variable A would supply the knowledge needed to get variable B.

By knowing the mutual information between two variables A and B, modeling of a multivariable system can be accomplished. If two or more variables provide the same information or have similar effects on one outcome, then this can be taken into consideration while constructing a model. For example, if the level of a tank was observed to be too high, pressure and temperature could be plausible variables. If manipulating both of these variables causes a change in the fluid level in the tank, then there exists some mutual information between the pressure and temperature. It would then be important to establish the relationship between temperature and pressure to appropriately control the level of fluid in the tank.

Visual Representation of Mutual Information

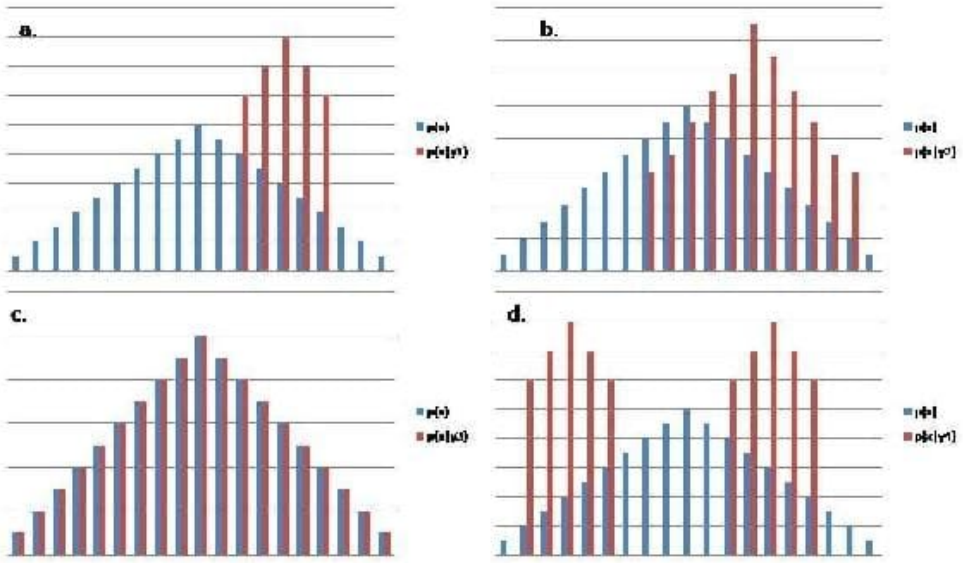


Figure 1. This figure demonstrates how y provides information for x. The blue is p(x) and the red is $p(x|y)$ for a specific y. A. y shares mutual information with x such that given a y1 the values of x are now narrowed down to the five values. B. y shares mutual information with x but it shares less information than A and therefore is not as helpful as A, as it does not eliminate as many potential x values. C. y does not share any mutual information. Given the y3 value, no x values are eliminated as potential values. Since y3 does not provide any information there is no information between the two variables. D. y shares mutual information with x. Because of the bimodal distribution, one more piece of information is required to describe which distribution lies within. Therefore, D gives as much information as part B. B gives 11 potential values while D gives 10 potential values. However, one additional bit of information is required to describe which of the two curves the x values lie on, thus there are 11 needed inputs.

Relating Mutual Information to Other Quantities/Concepts

Mutual information can be related to other concepts to gain a comprehensive understanding of the information obtained from any given parameter. Mutual information discussions are most commonly coupled with a discussion on Shannon entropy (See Shannon Entropy for more information on entropy).

Relative Entropy

In terms of probability theory, the relative entropy of a system measures the distance between two distinct probability functions. In chemical engineering applications, mutual information can be used to determine the entropy inherent in a system. Mutual information in this aspect can now be defined simply as:

$$I(A;B) = H(A) - H(A \mid B) = H(B) - H(BA) = I(B;A) = H(A) + H(B) - H(A,B)$$

where,

- *H*(*A*) = marginal entropy of A.
- *H*(*B*) = marginal entropy of B.
- *H*(*A* | *B*) = conditional entropy of A.
- *H*(*B* | *A*) = conditional entropy of B.
- *H*(*A*,*B*) = joint entropy of A and B.

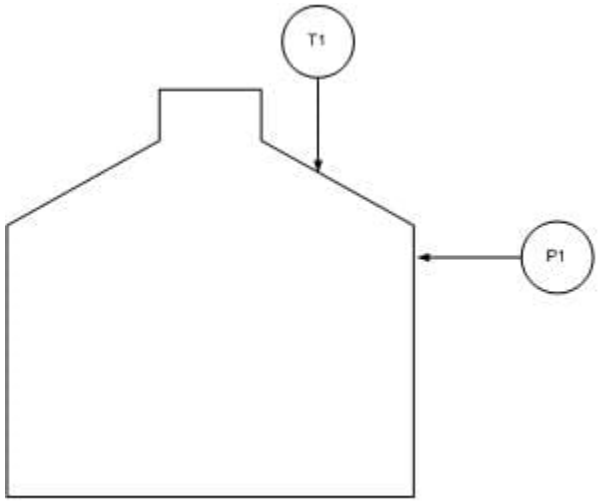
As described above, the marginal entropy still represents the uncertainty around a given variable A and B. Conditional entropy describes the uncertainty in the specified variable that remains after the other variable is known. In terms of mutual information, the conditional entropies of A and B tell us that a certain number of bits needs to be transferred from A in order to determine B and vice versa. Once again, these bits of information are determined by the logarithmic base chosen for analysis. One can determine the uncertainty in several different process parameters with respect to any other process variable.

Mutual Information and Boolean Models

Recall that via Boolean Models, a complex series of sensors and valves can be modeled into a simple control system. A Boolean variable, however, will never return a value because it can only return True or False. This knowledge can be combined to determine a relationship between two different variables (from Boolean model) and then their dependency can be quantified (mutual information). Utilizing the same logic that goes into constructing truth tables, data can be input into tables to determine the mutual information value.

Correlation Example

Suppose an engineer wants to test the applicability of the ideal gas law to a closed tank containing a fixed amount of his company's gaseous product. Given that the vessel is of constant volume, and given that no change in moles will occur in this fixed system, one would hypothesize that the ideal gas law would predict a linear relationship between the temperature and pressure within the tank.



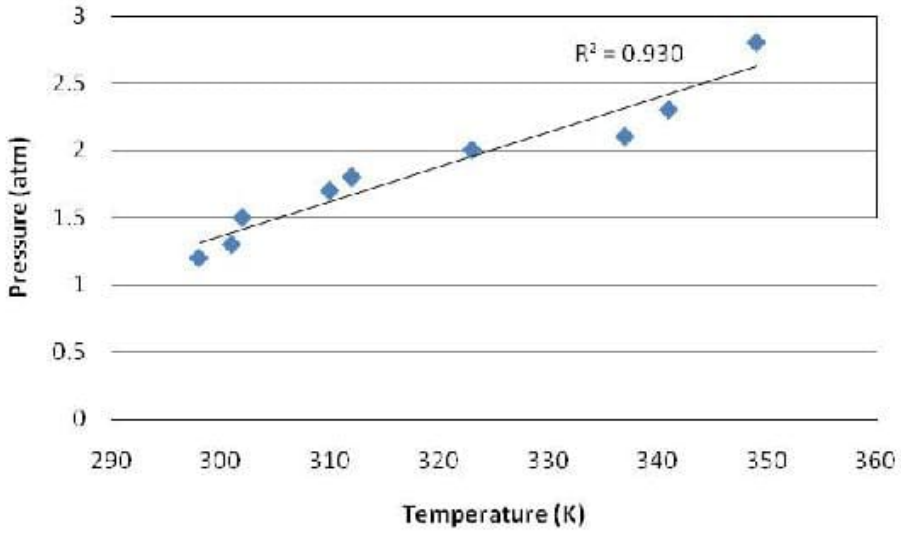
The engineer makes the following measurements and notes his findings in Excel.

T (K)	P (atm)
298	1.2
301	1.3
302	1.5
310	1.7
312	1.8
323	2
337	2.1
341	2.3
349	2.8

Before proceeding with his analysis, the engineer decides to calculate the correlation coefficient for these values in Mathematica.

```
In[14]:= T = {{298}, {301}, {302}, {310}, {312}, {323}, {337}, {341}, {349}}
         P = {{1.2}, {1.3}, {1.5}, {1.7}, {1.8}, {2.0}, {2.1}, {2.3}, {2.8}}
         Correlation[T, P]

Out[14]= {{298}, {301}, {302}, {310}, {312}, {323}, {337}, {341}, {349}}

Out[15]= {{1.2}, {1.3}, {1.5}, {1.7}, {1.8}, {2.}, {2.1}, {2.3}, {2.8}}

Out[16]= {{0.964549}}
```

Noting the correlation coefficient's approximate value of 1, the engineer proceeds by creating a linear fit of the data in Excel.



Indeed, the data are nearly linear.

Summary

If an engineer were trying to determine if two random variables are linearly related, calculating correlation coefficients would be the first step to determine this. If the mean and standard deviation for both variables are known values (i.e. calculated from thousands of measurements), then the population correlation coefficient should be used. To determine the correlation coefficient for a smaller set of data, the sample correlation coefficient equations should be used. If the two variables were shown to have a high degree of correlation, the next step would be to plot the data and graphically check the linear relationship of this data. A high correlation coefficient does not necessarily mean the two variables are linearly related as shown by Anscombe's quartet, so this graphical check is necessary.

Mutual information helps reduce the range of the probability density function (reduction in the uncertainty) for a random variable X if the variable Y is known. The value of I(X;Y) is relative, and the larger its value, the more information that is known of X. It is generally beneficial to try to maximize the value of I(X;Y), thus minimizing uncertainty. The concept of mutual information is quite complex and is the basis of information theory.

References

- Larsen, Richard J. and Marx, Morris L. (2006) *An Introduction to Mathematical Statistics and Its Applications*, New Jersey: Prentice Hall. ISBN 0-13-186793-8
- Wikipedia Correlation.
- Wikipedia Mutual Information.
- Scholarpedia Mutual Information.

- Kojadinovic, Ivan (no date) "On the use of mutual information in data analysis: an overview"

This page titled 13.13: Correlation and Mutual Information is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Daniel McClure, Daniel Wheeler, Jeremy Goldfarb, & Lilian Ngobi via source content that was edited to the style and standards of the LibreTexts platform.

13.14: Random sampling from a stationary Gaussian process

Gaussian distributions are powerful tools for data analysis of randomly generated samples. Gaussian distributions are typically associated with a “Bell curve”, a graphical representation of a probability density function that is characterized by a peak at the data mean and a width determined by the standard deviation of the data set. The distribution of the curve around the mean is symmetric, with certain percentages of the data falling within one standard deviation of the mean. The probability that a data point exists within a certain range of values is equal to the area under the probability density curve between those two values. The total area under the curve is equal to one because all data points can be found to exists somewhere in that data set. For more detailed information on Gaussian distributions, consult the section on Continuous Distributions.

The random number sampler is a powerful tool and is very useful in process modeling. One example of a use for a random number samplers is to generate weather data from a model to simulate changing climate throughout the year, in order to properly maintain the temperature of a storage vessel. Another way it can be used is to create a control chart to monitor how well a process is controlled over time based on some information about normal system operation. The goal of random number generation is to be able to use large amounts of data based on limited experimentation to test a control scheme.

The objective of this article is to explain how random sampling to form a Gaussian data set is useful in process control. First is discussion of how to use a Gaussian distribution to analyze probability and compare the statistical similarities/differences between different data sets. Then is an explanation of how to use the Central Limit Theorem and show how to create a Gaussian distribution from non-Gaussian data using average values. Last are examples of how to use this theory in process controls, including how to generate and use control charts.

Random Number Sampler

The purpose of the random number sampler is to take randomly generated points and turn them into a Gaussian distribution (i.e. a PDF). It is important to use points distributed throughout the range of 0 to 1. This task is accomplished by assuming a random number (R) is equal to the ‘y’ value of the cumulative distribution function (CDF). The CDF is the cumulative area under the curve of the probability density function (PDF). By taking an integral from $-\infty$ to x in the PDF, this ‘y’ value of the CDF can be found. Since the value of R is already known, we can find the x values of the CDF and thus, the PDF. With the x values of the PDF known, we can find $P(x)$ and construct the Gaussian distribution. The relationship between R and x is shown below:

$$R = \int_{-\infty}^{\frac{x}{2}} \frac{1}{\sqrt{2\pi(\sigma)^2}} e^{\frac{-(X-\mu)^2}{2(\sigma)^2}} dX$$

$$R = \frac{1}{2}\left[\mathrm{Erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right) - \mathrm{Erf}\left(\frac{-\infty-\mu}{\sigma\sqrt{2}}\right)\right]$$

$$R = \frac{1}{2}\left[\mathrm{Erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right) + 1\right]$$

$$x = \sigma\sqrt{2}\,\mathrm{InverseErf}(2R-1) + \mu$$

The process of choosing a random number and implementing it into the PDF can be done using Mathematica. Excel is not recommended because it cannot compute the Erf() of a negative number, so we cannot use it for data points below the mean in a CDF.

The first step for random sampling a stationary Gaussian process is to input the mean (μ) and the standard deviation (σ) into the equation below. Then, you can determine the random points either with Random[ ] function in Mathematica or via user input to develop a list of random numbers between 0 and 1. Using a random number input as R, Mathematica can be used to determine the corresponding data point x using the syntax below:

Solve[R == (1/2)*(Erf[(x- μ)/(σ*Sqrt[2])]+1), x]

The syntax can be repeated to determine as many random numbers and their corresponding x values as is necessary for your problem. A CDF plot can be created by plotting all of the random R values versus their corresponding x values.

We will demonstrate how the procedure works in the following example. First, we chose 10 random numbers and found corresponding x values, given in Table 1 with μ = 10 and σ = 5. See Figure 1 for the plot of the CDF.

Table 1: CDF Random Numbers and Corresponding x Values

Random Number, R	x
0	-∞
0.1	3.59224
0.2	5.79189
0.3	7.378
0.4	8.73326
0.5	10
0.6	11.2667
0.7	12.622
0.8	14.2081
0.9	16.4078
1	∞

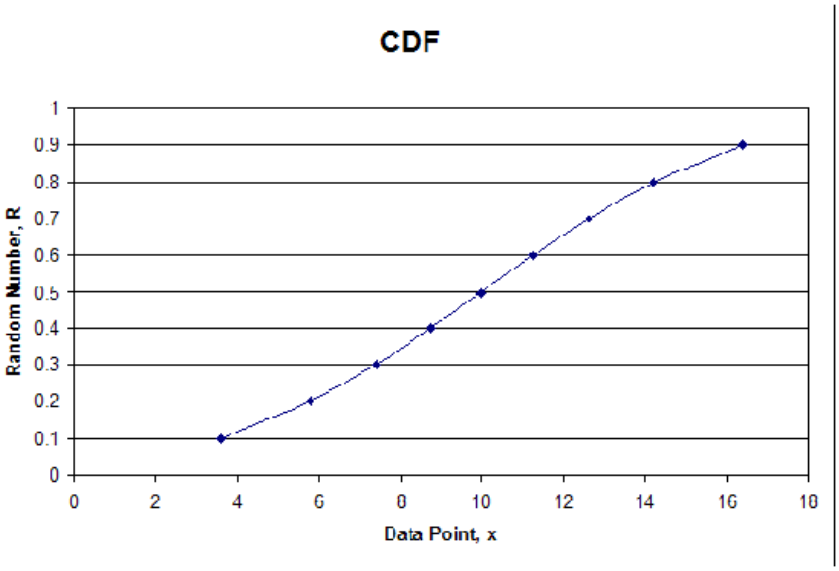


Figure 1: Continuous Distribution Function

In order to develop the PDF you would input the mean and standard deviation into the following PDF equation, and find $P(x)$ for each value of x.

$$P(\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(X-\mu)^2}{2\sigma^2}}$$

The syntax used in Mathematica is:

Solve[P(x) == (1/Sqrt[2*Pi*σ^2])*e^(-(x-μ)^2/(2*σ^2)), P(x)]

The syntax outputs values for $P(x)$ which are then plotted against the x values calculated from the CDF. The following table shows the random numbers generated in Mathematica, the corresponding x values, and the $P(x)$ values for the PDF. The plot of the PDF is also shown below.

Table 2: For Each R, x Values and Corresponding $P(x)$ Values

Random Number, R	x	P(x)
0	-∞	0
0.1	3.59224	0.0351
0.2	5.79189	0.055992
0.3	7.378	0.069539
0.4	8.73326	0.077268
0.5	10	0.079788
0.6	11.2667	0.077269
0.7	12.622	0.069539
0.8	14.2081	0.055992
0.9	16.4078	0.035099
1	∞	0

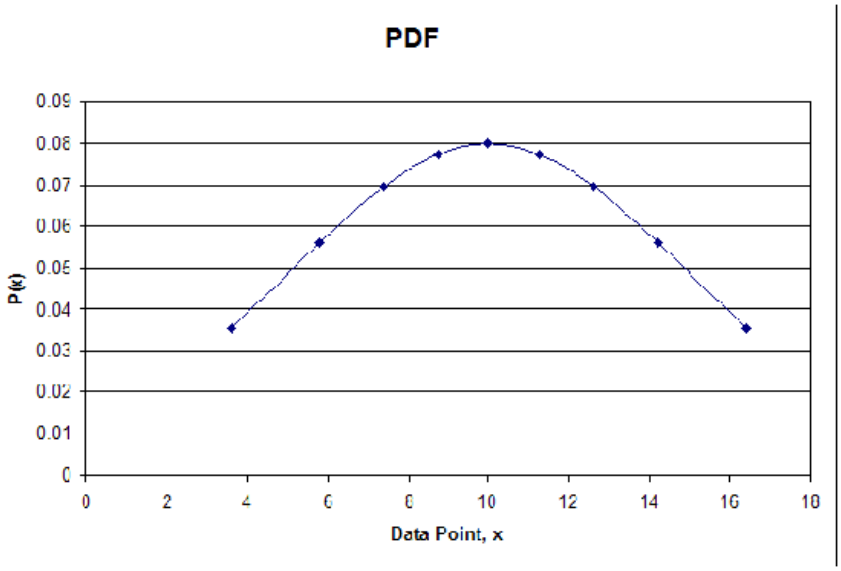


Figure 2: Probability Density Function

As you can see the plot is in fact Gaussian. The distribution will become more Gaussian as more random numbers are used. The take home message from the random number generator is that a data set with numbers ranging between 0 and 1, obtained through random number generation, can be converted into a Gaussian distribution.

Probability Primers

One major advantage of the random number sampler is the ability to generate different data sets without having to actually gather data. These data sets can be used to help you better understand the workings of many of the different statistical comparisons. For instance, if the sample size increases, does that lead to more or less significance in the difference of two means? Most of these analysis tools are the topics of other wikis, so we will revisit a few only briefly with the intent on asking more in-depth questions in example 1.

Probability

Recall that the area under any Gaussian function is related to the probability. Suppose we use our sampler to generate a random set of 100 data points with a mean of 10 and a standard deviation of 5. The probability of creating an additional data point

between ‘a’ and ‘b’ is the integral of the Gaussian distribution function from ‘a’ to ‘b’ as follows:

$$P(a \leq x \leq b) = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$

Further information on probabilities can be found in the section on Continuous Distributions.

Error in the Mean

Assume we have the same data set as described above. Let’s say we wish to add a new data point and compute the probability that the new mean of the data set lies between ‘c’ and ‘d.’ This calculation is very similar to that described above with one difference. The standard deviation, which describes the variance in individual data points, is replaced with the standard error in the mean, which describes the variance in the mean as data sample size changes. The standard error in the mean is calculated as follows:

$$\sigma_\mu = \frac{\sigma}{\sqrt{n}}$$

Once this value is obtained, we can solve for the probability as follows:

$$P(a \leq \mu \leq b) = \int_c^d \frac{1}{\sqrt{2\pi\sigma_\mu^2}} e^{\frac{-(x-\mu)^2}{2\sigma_\mu^2}} dx$$

Further information on standard error can be found in section on the Comparison of Two Means.

Comparison of Two Data Sets

Now suppose we have two distinct data sets with different means and standard deviations and we wish to determine if one data set is statistically different from the second. To do this, we will compute a p-value. The p-value is computed as follows:

$$P\left(\mu_1 \to \mu_2, \mu_1 < \mu_2\right) = \int_m^\infty \frac{1}{\sqrt{2\pi\sigma_{\mu 1}^2}} e^{\frac{-(x-\mu)^2}{2\sigma_{\mu 1}^2}} dx$$

$$P\left(\mu_1 \to \mu_2, \mu_1 > \mu_2\right) = \int_{-\infty}^m \frac{1}{\sqrt{2\pi\sigma_{\mu 1}^2}} e^{\frac{-(x-\mu)^2}{2\sigma_{\mu 1}^2}} dx$$

Further information on p-values and mean comparisons can be found in these wiki articles: Comparison of Two Means and P-Values

Central Limit Theorem

Recall that one property of the random number sampler was that as the set of randomly generated numbers increased, errors became more random and the distribution became more Gaussian. The central limit theorem states that the sampling of a given distribution mean will approach a normal or Gaussian distribution as the sample size increases. This theorem is proven by the following example illustrated in the “Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building.”

Rolling of Dice

The probability that a certain face will land upright when rolling a six-sided die is the same for each face. This is shown below in Figure 3 (a). The mean value of a roll can be calculated to be 3.5 by summing up the value of each face and dividing by 6. As the sample size is increased, this mean value will be shown to have the highest density of occurance. When the sample size is increased to two dice you can notice in Figure 3 (b) that the density distribution of the average score begins to take the shape of a curve. The density distribution of the average score for sample size increases to three, five, and ten dice are shown in Figure 3 (c, d, and e respectively). For each increase it can be noted that the density of the extreme values decreases and the overall density distribution appears more like a Gaussian distribution, as predicted by the central limit theorem.

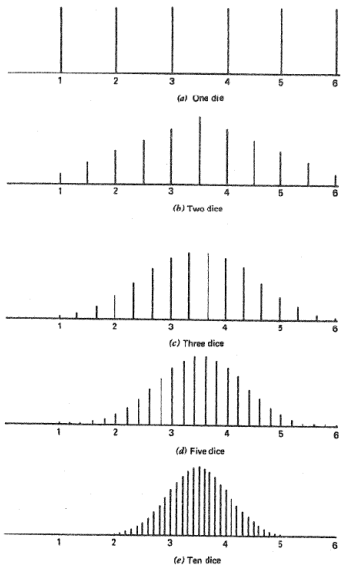


Figure 3: Distribution from Increasing Sample Size (Box, "Statistics for Experimenters")

Random Number Generation

The above dice example is analogous to random number generation. Consider a set of random numbers between 0 and 10. If these numbers were truly chosen randomly, then you would expect an even distribution of numbers between 0 and 10. The average of these numbers would be about 5, with a large standard deviation. However, let's suppose your data set now consists of an average of two random numbers. In this case, the data will be less evenly distributed, as averaging tends to lessen the contribution of terms further away from 5. The shape of the distribution will begin to appear Gaussian.

As you increase the number of random numbers averaged together ('n') to create each data point, the distribution becomes more Gaussian in nature. Please refer to figure 4. When n=1 (one random number is used to compute each data point), the density distribution is relatively even. As 'n' increases, the curves become more Gaussian in nature until we have a very smooth Gaussian distribution when n=100 (100 random numbers were averaged to obtain one data point).

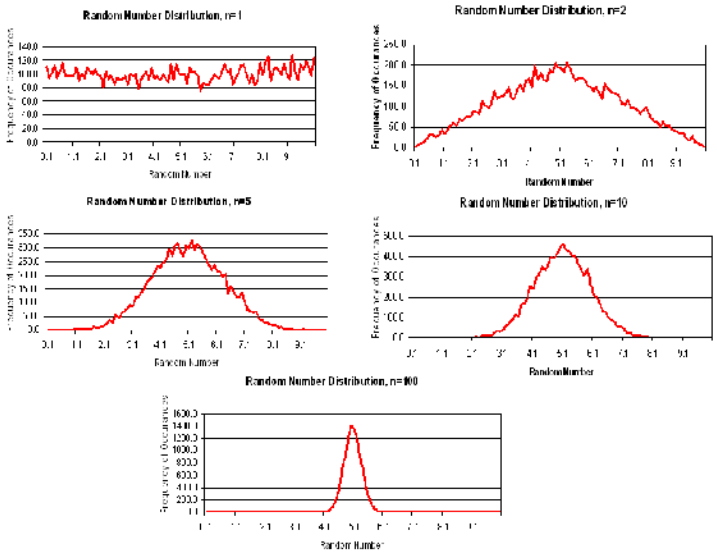


Figure 4: Distribution of data points generated by averaging random numbers

Example 13.14.1

In a healthy adult female, studies have shown that the mean fraction of deoxygenated blood leaving the lungs is 0.05 with a standard deviation 0.01. To obtain this data, 100 healthy females had been sampled using our random number sampler.

a. Assuming this distribution is Gaussian, 95% of all healthy females will have deoxygenated blood fractions in what range (assume this range is centered over the mean of the data)?
b. What is the expected error in the mean to a 95% level of confidence?
c. A second study has been conducted on pregnant adult females. It was found that the mean fraction of deoxygenated blood leaving the lungs was 0.06 with a standard deviation of 0.02. Is the deoxygenated blood concentration in this population of women statistically different?
d. Now assume that only 10 adult females had been sampled instead of 100. Analyze how the answer to part (c) may change.

Solution

a) To solve this problem, recall the equation of a Gaussian distribution. Substitute in the value of the mean (0.05) and standard deviation (0.01). Next, recall that the area under the curve corresponds to probability, so we can set integrate this function to determine the range at which the probability equals 0.95. The bounds on this integral are 0.05 + k and 0.05 – k, since we are assuming this data range is centered about the mean. This integral can be solved in integral form or in error function form, depending on the commands you choose to use in inputting the function into a computer algebra system solver. Maple, Excel, and Mathematics can be used to solve the expression simply by entering the last line of math text shown in the solution below.

$$P(\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$
$$P(\mu,\sigma) = \frac{1}{\sqrt{2\pi(0.01)^2}} e^{\frac{-(x-0.05)^2}{2(0.01)^2}}$$
$$0.95 = \int_{005-k}^{005+k} \frac{1}{\sqrt{2\pi(0.01)^2}} e^{\frac{-(x-0.05)^2}{2(001)^2}} dx$$
$$0.95 = \frac{1}{2}\left[\text{Erf}\left(\frac{0.05+k-0.05}{0.01\sqrt{2}}\right) - \text{Erf}\left(\frac{0.05-k-0.05}{0.01\sqrt{2}}\right)\right]$$
$$k = 0.02$$

Therefore, 95% of healthy adult females have deoxygenated blood levels between 0.03 and 0.07.

b) This problem is solved much in the same way we solved part (a). However, in this case we are looking for the 95% confidence interval of the mean, and not the entire population. Therefore, the standard deviation must be converted to the standard error in the mean. Then, all the above calculations are repeated.

$$\sigma_\mu = \frac{\sigma}{\sqrt{n}} = \frac{0.01}{\sqrt{100}} = 0.001$$
$$P(\mu,\sigma_\mu) = \frac{1}{\sqrt{2\pi(0.001)^2}} e^{\frac{-(x-0.05)^2}{2(0.001)^3}}$$
$$0.95 = \int_{0.05-k}^{0.05+k} \frac{1}{\sqrt{2\pi(0.001)^2}} e^{\frac{-(x-0.09)^2}{2(0.001)^2}} dx$$
$$0.95 = \frac{1}{2}\left[\text{Erf}\left(\frac{0.05+k-0.05}{0.001\sqrt{2}}\right) - \text{Erf}\left(\frac{0.05-k-0.05}{0.001\sqrt{2}}\right)\right]$$
$$k = 0.002$$

Therefore, the 95% confidence interval of the mean is 0.05 ± 0.002.

c) In order to compare the statistical significance of two different data sets, the concept of p-values must be used. Since we are interested comparing the means of these two data sets, standard deviation will be replaced by standard error in the mean. To find the probability (or p-value) that pregnancy results in higher levels of deoxygenated blood, we need to calculate the area under the Gaussian curve for healthy females that is 0.06 or more. Remember, we are evaluating the Gaussian function describing healthy females, so standard deviation data for the pregnant females is not needed.

$$\sigma_\mu = \frac{\sigma}{\sqrt{n}} = \frac{0.01}{\sqrt{100}} = 0.001$$
$$P\left(\mu, \sigma_\mu\right) = \frac{1}{\sqrt{2\pi\sigma_\mu^2}} e^{\frac{-(x-\mu)^2}{2\sigma_k^2}}$$
$$P\left(\mu, \sigma_\mu\right) = \frac{1}{\sqrt{2\pi(0.001)^2}} e^{\frac{-(x-005)^2}{2(0.001)^2}}$$
$$P(x \geq 0.06) = \int_{0.06}^{\infty} \frac{1}{\sqrt{2\pi(0.001)^2}} e^{\frac{-(x-0.05)^2}{2(0.001)^1}} dx$$
$$P(x \geq 0.06) = \frac{1}{2}\left[\text{Erf}\left(\frac{\infty - 0.05}{0.001\sqrt{2}}\right) - \text{Erf}\left(\frac{0.06 - 0.05}{0.001\sqrt{2}}\right)\right]$$
$$P(x \geq 0.06) = \frac{1}{2}\left[1 - \text{Erf}\left(\frac{0.06 - 0.05}{0.001\sqrt{2}}\right)\right]$$
$$P(x \geq 0.06) = 0$$

The p-value is equal to zero. By convention, p-values less than 0.05 are considered to be statistically significant. Therefore, we conclude that pregnancy statistically affects the level of deoxygenated blood in an adult women's body.

d) The size of the sample only has an affect on the standard error in the mean. To solve this problem, recalculate the standard error and repeat the calculations above.

$$\sigma_\mu = \frac{\sigma}{\sqrt{n}} = \frac{0.01}{\sqrt{10}} = 0.003$$
$$P\left(\mu, \sigma_\mu\right) = \frac{1}{\sqrt{2\pi\sigma_\mu^2}} e^{\frac{-(x-\mu)^2}{2\sigma_k^2}}$$
$$P\left(\mu, \sigma_\mu\right) = \frac{1}{\sqrt{2\pi(0.003)^2}} e^{\frac{-(x-0.09)^3}{2(0000)^2}}$$
$$P(x \geq 0.06) = \int_{0.06}^{\infty} \frac{1}{\sqrt{2\pi(0.003)^2}} e^{\frac{-(x-0.00)^2}{2(0.003)^2}} dx$$
$$P(x \geq 0.06) = \frac{1}{2}\left[\text{Erf}\left(\frac{\infty - 0.05}{0.003\sqrt{2}}\right) - \text{Erf}\left(\frac{0.06 - 0.05}{0.003\sqrt{2}}\right)\right]$$
$$P(x \geq 0.06) = \frac{1}{2}\left[1 - \text{Erf}\left(\frac{0.06 - 0.05}{0.003\sqrt{2}}\right)\right]$$
$$P(x \geq 0.06) = 0.0004$$

In this case, the p-value is still less than 0.05, so we still arrive at the same conclusion. We also conclude that as sample size decreases, differences in sample means become less significant because the p-value has slightly increased.

Example 13.14.2: Comparison of Two Data Sets

This example is intended to demonstrate how increased sample size affects the comparison of two data sets.

Using a random number generator, four data sets (A,B,C,D) were generated. Each data set contains 100 total data points. For data sets A and B, two random numbers were averaged to attain each data point, while in C and D, five random numbers were averaged for each point. This difference resulted in smaller standard deviations for data sets C and D. A summary of the four data sets created is shown in the following chart. M is the number of random samples averaged to obtain each data point. N is the total number of data points. The mean and standard deviation for each of the data sets is also given.

Data Set	m	n	Mean	St Dev
A	2	100	5	2.3
B	2	100	5.5	1.8
C	5	100	5	1.1
D	5	100	5.5	1.4

a. Compare the means of data sets A and B by calculating the p-value
b. Compare the means of data sets C and D by calculating the p-value
c. Compare the p-values obtained for parts a) and b) above. Explain why the values make sense.
d. What general effect does increasing sample size have on the comparison of two data sets?

Solution

a) The p-value comparing data sets A and B is computed as follows:

$$\sigma_\mu = \frac{\sigma}{\sqrt{n}} = \frac{2.3}{\sqrt{100}} = 0.23$$
$$P\left(\mu, \sigma_\mu\right) = \frac{1}{\sqrt{2\pi\sigma_\mu^2}} e^{\frac{-(x-\mu)^2}{2\sigma_k^2}}$$
$$P\left(\mu, \sigma_\mu\right) = \frac{1}{\sqrt{2\pi(0.23)^2}} e^{\frac{-(x-5)^2}{2(0.23)^2}}$$
$$P(x \geq 5.5) = \int_{55}^{\infty} \frac{1}{\sqrt{2\pi(0.23)^2}} e^{\frac{-(x-5)^3}{2(0.23)^3}} dx$$
$$P(x \geq 5.5) = \frac{1}{2}\left[\text{Erf}\left(\frac{\infty - 5}{0.23\sqrt{2}}\right) - \text{Erf}\left(\frac{5.5-5}{0.23\sqrt{2}}\right)\right]$$
$$P(x \geq 5.5) = \frac{1}{2}\left[1 - \text{Erf}\left(\frac{5.5-5}{0.23\sqrt{2}}\right)\right]$$
$$P(x \geq 5.5) = 0.015$$

From this p-value, we can see that it is very unlikely that these two data sets are statistically the same. There is only a 1.5% chance of randomly getting a data set with a mean as high as 5.5! It is far more likely that these two data sets are actually statistically different.

b) The p-value comparing data sets C and D is computed as follows:

$$\boldsymbol{\sigma}_\mu = \frac{\boldsymbol{\sigma}}{\sqrt{n}} = \frac{1.1}{\sqrt{100}} = 0.11$$
$$P\left(\mu, \sigma_\mu\right) = \frac{1}{\sqrt{2\pi\sigma_\mu^2}} e^{\frac{-(x-\mu)^2}{2\sigma_F^2}}$$
$$P\left(\mu, \sigma_\mu\right) = \frac{1}{\sqrt{2\pi(0.11)^2}} e^{\frac{-(x-5)^3}{2(0.11)^2}}$$
$$P(x \geq 5.5) = \int_{5.5}^{\infty} \frac{1}{\sqrt{2\pi(0.11)^2}} e^{\frac{-(x-5)^2}{2(0.11)^3}} dx$$
$$P(x \geq 5.5) = \frac{1}{2}\left[\text{Erf}\left(\frac{\infty - 5}{0.11\sqrt{2}}\right) - \text{Erf}\left(\frac{5.5-5}{0.11\sqrt{2}}\right)\right]$$
$$P(x \geq 5.5) = \frac{1}{2}\left[1 - \text{Erf}\left(\frac{5.5-5}{0.11\sqrt{2}}\right)\right]$$
$$P(x \geq 5.5) = 0$$

From this p-value, we can see that these two data sets are statistically different. There is an almost 0 percent chance of randomly getting a data set with a mean as high as 5.5!

c) Comparing the two p-values, we can see that the value for b) is smaller than a), indicating that we are more confident about a statistical difference between sets C and D than between A and B. Going back to the description of the problem, we can find a reason for this result. Because of the larger sampling of random numbers in data sets C and D (five numbers averaged for each data point, compared to just two), these data sets have smaller standard deviations. These smaller standard deviations mean that we have a higher level of confidence that the sample mean is the true mean. Because each data set is more likely to show the true mean, there is also an increased likelihood that one data set is statistically different from the other.

d) In a Gaussian distribution, as sample size increases, standard error decreases. This indicates that the sample mean is closer to the true mean, and two data sets are more likely to be statistically different from each other than if there are smaller sample sizes and higher standard deviations in the data sets.

Control Charts

This section will provide a brief overview on control charts in order to complete example 3, in which our random number sampler is applied to a control chart situation. For a more detailed explanation on control charts, and the methodology behind them, please see this wiki page: Control Charts

Background

Controls charts are tools used to determine whether a particular part of a process is predictable or not. A process is predictable when it is in a state of statistical control and unpredictable when it is not in a state of statistical control. A state of statistical control simply means that we can accurately predict what the output of a process will be in the future based on past measurements; it does not mean that the product is on target or within any limits of consumer acceptability. The random number sampler is a useful tool that can be used to analyze control charts. As previously discussed, a data set is more Gaussian in nature as the amount of data points is increased. Therefore, one will expect that a large set of data generated with our random number sampler will always be in statistical control, but smaller sets may contain an element of unpredictability.

To create a control chart, we begin by looking at historical data measurements on the measurement (variable) of importance, e.g. acetic acid concentration. Once a reasonable amount of data has been gathered, it is used to calculate appropriate limits for the variable. If the historical data and future measurements fall within the range of the limits, it is safe to predict that future measurements will also continue to be within the range of the limits. Conversely, if the historical data does not fall within the range of the limits, it can be safely predicted that future measurements will not fall within the range of the limits.

Constructing Control Charts

The best way to illustrate how to create a control chart is to go through an example. This is done below. The initial assumption in developing a control chart is that the process is stable over period of time where a set of measurements can be made on a particular variable of the process. In the example below, this variable is acetic acid concentration.

Example 13.14.1

Chemical engineers often combine acetic anhydride and water to produce acetic acid in a moderately exothermic reaction. The reaction in this particular process is supposed to produce a product stream of acetic acid at 5.5 wt%. The composition of the product stream is measured four times every hour for ten hours. Each hour's measurements are viewed as a subgroup of data. Table 1 shows the data obtained from the 40 measurements. Table 1 also listed the average concentration and the range of concentrations for each subgroup of measurements.

Table 3. Sample Data - Acetic Acid Concentration

Subgroup	Concentration (wt%)				Average	Range
1	5.31	5.18	5.57	5.05	5.28	0.52
2	5.97	5.61	5.08	5.41	5.52	0.89
3	5.35	5.97	5.82	5.64	5.70	0.63
4	5.78	5.22	5.35	5.50	5.46	0.56
5	5.57	5.08	5.16	5.86	5.42	0.79
6	5.97	5.54	5.09	5.42	5.50	0.87
7	6.00	5.18	5.57	5.71	5.61	0.82
8	5.23	5.72	5.82	5.23	5.50	0.59
9	5.05	5.95	5.97	5.01	5.50	0.95
10	5.88	5.04	5.87	5.62	5.60	0.84

To create a control chart for this process, we must first compute the average concentration for each subset. Additionally, we must determine the range of data for each subset. These steps have already been completed and their values are listed in Table 3 in the last two columns on the right. Next the grand average concentration and average range should be computed by averaging the average concentrations of each subset and by taking the average of each subset's ranges. For the data in table 3, the grand average concentration is 5.51 wt% and the average range is 0.74 wt%. At this point, we are ready to determine our upper and lower control limits for our acetic acid concentration. To determine these limits we will need to use a table of constants that have been mathematically derived for Gaussian distributions. These numbers can then be inserted into equations that find the upper and lower average control limits (UCL_x and LCL_x) and the upper and lower range control limits (UCL_R andLCL_R). The following equations provide the control limits for the average concentrations and for their ranges.

$$UCL_{\overline{X}} = \overline{\overline{X}} + A_2\overline{R}$$

$$LCL_{\overline{X}} = \overline{\overline{X}} - A_2\overline{R}$$

$$UCL_R = D_4\overline{R}$$

$$LCL_R = D_3\overline{R}$$

Table 5. Constants for Average and Range Charts Based on the Average Range

Constants for Average Range Charts Based on the Average Range			
n	A_2	D_3	D_4
2	1.88	0	3.268
3	1.023	0	2.574
4	0.739	0	2.282
5	0.577	0	2.114
6	0.483	0	2.004
7	0.419	0.076	1.924
8	0.373	0.136	1.864
9	0.337	0.184	1.816
10	0.308	0.223	1.777

For this example, our subgroup size, n, is 4, A2 is 0.729, D3 is 0 and D4 is 2.282. When plugged into the above equations, they yield control limits of:

$$UCL_{\bar{X}} = \mathbf{6.05}$$

$$LCL_{\bar{X}} = \mathbf{4.97}$$

$$UCL_{R} = \mathbf{1.70}$$

$$LCL_{R} = \mathbf{0}$$

Graphically the control chart is represented by creating a plot of the average concentration values versus their subgroup number and creating a plot of each range value versus the subgroup number. Horizontal lines should be included on these plots which indicate the control limits. Control charts for this example are plotted in figures 3 and 4.

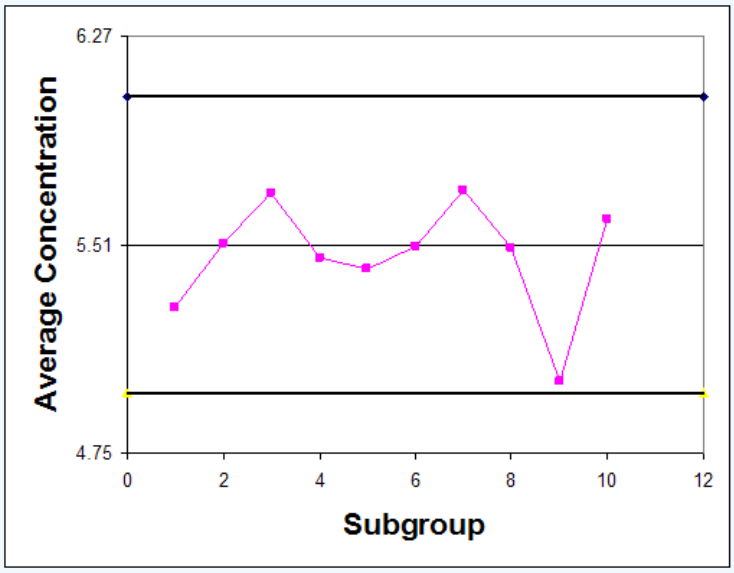


Figure 6. Control Chart for Average Acetic Acid Concentration

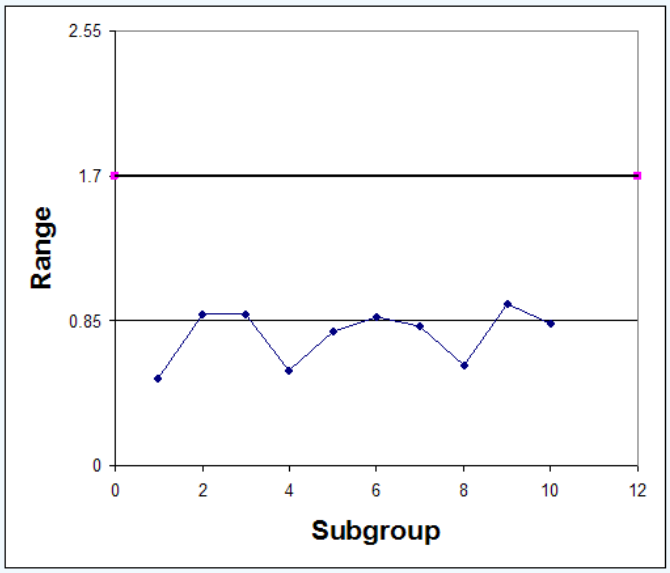


Figure 7. Control Chart for Ranges of Acetic Acid Concentration

Interpreting Control Charts

If any of the following rules hold true for the centering control chart, the process is not in statistical control.

1. One or more points fall outside the control limits.
2. Seven or more consecutive points fall on the same side of the centerline.
3. Ten of 11 consecutive points fall on the same side of the centerline.
4. Three or more consecutive points fall on the same side of the centerline and all are located closer to the control limit than the centerline.

Observe that in both the control chart for the average concentrations, the historical data never exceeds or goes below the control limits. Also, none of the other rules above hold true for this chart. This suggests that the process is stable; however, to confirm this thought, future observations must be made. If the average acetic acid concentration and its range continues to stay within the control limits, the process is said to be stable.

In the event that the historical data did not stay within the control limits, the plots could have looked like figures 5 & 6.

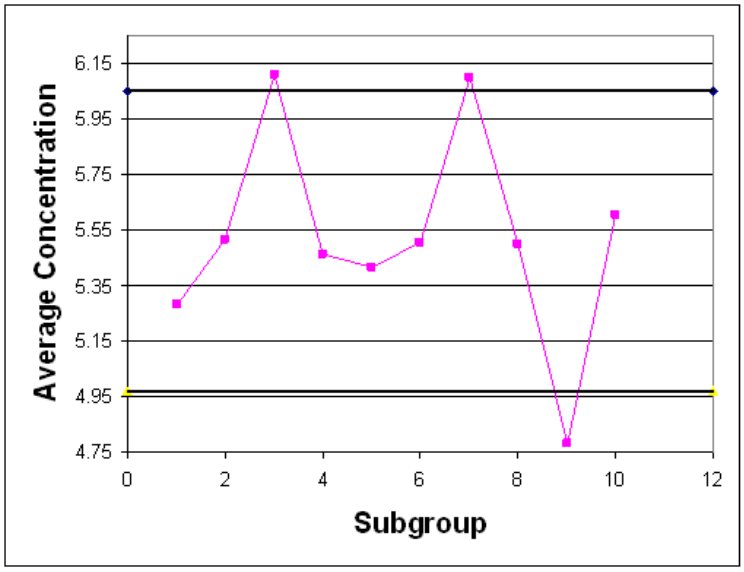


Figure 8. Control Chart for Average Acetic Acid Concentration (Unstable)

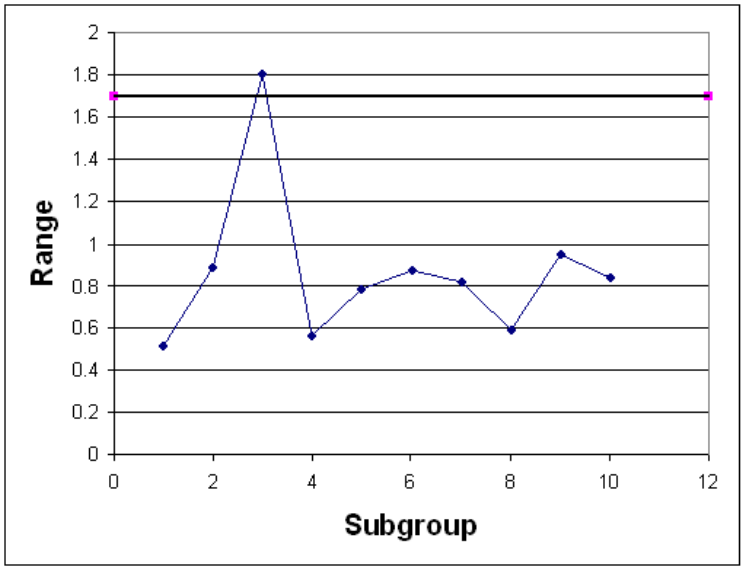


Figure 9. Control Chart for Ranges of Acetic Acid Concentration (Unstable)

Observe that in each of these control charts, the historical data does not stay within the control limits for the process. Based on this observation, it can be predicted that future data will also not stay within the control limits for the process and the process is not stable. In conclusion, control charts allow you to use samples of data to determine the control limits for a process and evaluate if the process is stable or not.

Example 13.14.1: Random Sampling Problem

In an industrial bioengineering process, it is desired to obtain a 75% proline to 25% hydroxyproline ratio in a specialized enzyme. The relative compositions can be analyzed using a fluorescent dye, where a fluorescence of 10 corresponds to the correct ratio of these two amino acids. To monitor the stability of this process, five reading are taken every 30 minutes for the duration of the 10 hour production process. The results were obtained using our random number sampler with a specified mean (10) and standard deviation (in this case 1.6). The results are in the table below

Time (hr)	Sample Flouresence				
	1	2	3	4	5
0.5	11.2	9.8	12.0	12.9	12.5
1	9.7	10.7	9.8	10.6	7.0
1.5	12.4	9.0	12.9	12.6	11.5
2	12.1	9.1	10.9	7.4	7.9
2.5	8.2	11.6	9.7	10.8	10.1
3	9.1	8.1	8.6	11.6	11.0
3.5	7.9	8.3	9.8	12.3	9.6
4	7.7	10.2	8.2	10.4	11.7
4.5	12.0	10.1	11.6	7.0	9.2
5	10.7	12.1	8.2	12.7	10.5
5.5	7.5	9.8	9.8	10.8	7.5
6	12.4	9.3	10.6	10.9	7.9
6.5	9.0	12.1	10.4	11.8	11.3
7	9.2	7.7	10.2	11.5	8.4
7.5	8.7	11.2	11.2	11.0	11.2
8	11.2	7.8	9.1	10.0	10.0
8.5	12.2	8.5	11.5	13.0	9.6
9	10.6	9.2	11.5	10.1	11.2
9.5	7.1	9.4	10.7	10.4	9.1
10	11.5	7.2	12.9	12.2	10.0

a. Calculate the average and range for each data subgroup.
b. Create an average and range chart for the entire time of production.
c. Is this process in control?

Solution

The average and range can be calculated using the =AVERAGE() and =MAX()-MIN() functions of Microsoft excel.

Time (hr)	Sample Flouresence						
	1	2	3	4	5	Avg	R
0.5	11.2	9.8	12.0	12.9	12.5	11.7	3.0
1	9.7	10.7	9.8	10.6	7.0	9.6	3.7
1.5	12.4	9.0	12.9	12.6	11.5	11.7	3.9
2	12.1	9.1	10.9	7.4	7.9	9.5	4.7
2.5	8.2	11.6	9.7	10.8	10.1	10.1	3.4
3	9.1	8.1	8.6	11.6	11.0	9.7	3.6
3.5	7.9	8.3	9.8	12.3	9.6	9.6	4.4
4	7.7	10.2	8.2	10.4	11.7	9.6	4.0
4.5	12.0	10.1	11.6	7.0	9.2	10.0	5.0
5	10.7	12.1	8.2	12.7	10.5	10.8	4.5
5.5	7.5	9.8	9.8	10.8	7.5	9.1	3.3
6	12.4	9.3	10.6	10.9	7.9	10.2	4.5
6.5	9.0	12.1	10.4	11.8	11.3	10.9	3.0
7	9.2	7.7	10.2	11.5	8.4	9.4	3.8
7.5	8.7	11.2	11.2	11.0	11.2	10.7	2.5
8	11.2	7.8	9.1	10.0	10.0	9.6	3.5
8.5	12.2	8.5	11.5	13.0	9.6	11.0	4.4
9	10.6	9.2	11.5	10.1	11.2	10.5	2.3
9.5	7.1	9.4	10.7	10.4	9.1	9.3	3.6
10	11.5	7.2	12.9	12.2	10.0	10.7	5.7

b) The grand average is 10.2 and the average range is 3.8. Using Table 2, A2 = 0.577, D3 = 0, and D4 = 2.114. Therefore:

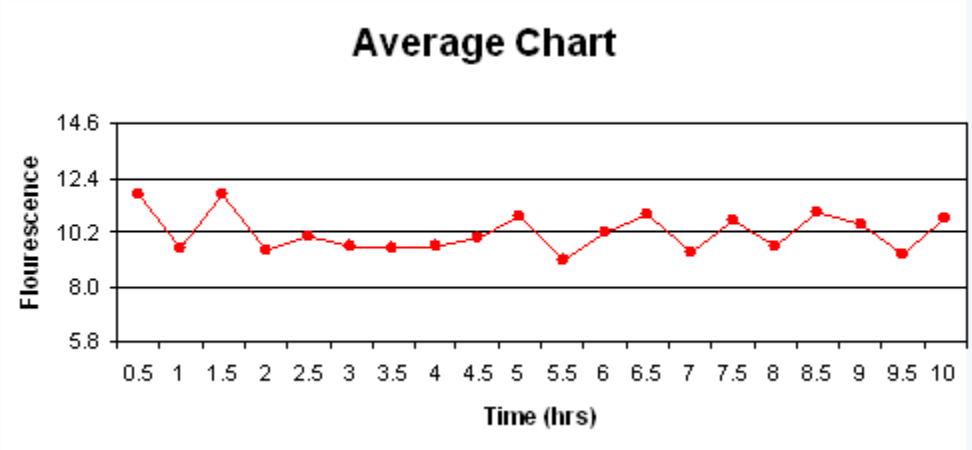
$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R} = 10.2 + (0.577)(3.8) = 12.4$$

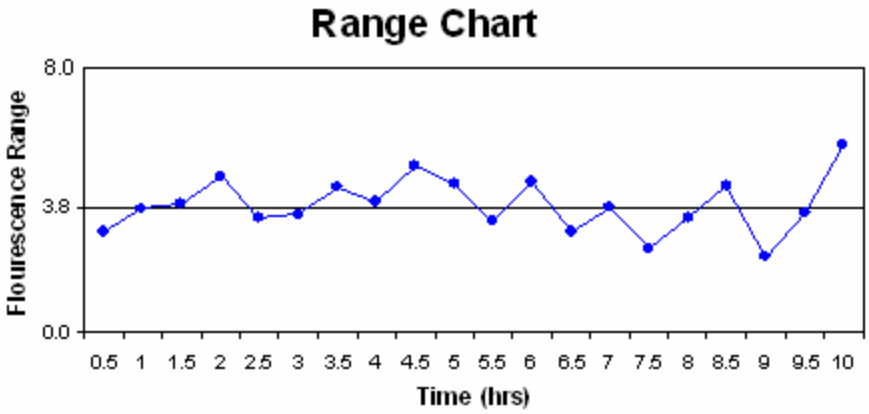
$$LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R} = 10.2 - (0.577)(3.8) = 8.0$$

$$UCL_R = D_4\bar{R} = (2.114)(3.8) = 8.0$$

$$LCL_R = D_3\bar{R} = (0)(3.8) = 0$$

The charts are as follows:





c) The first rule is not violated as none of the points fall outside the upper and lower control limits. Seven or more points do not fall on one side of the centerline (the maximum was six), so rule two is not violated. Rule three was not violated, as 10 of 11 points did not fall on the same side of the centerline (in ouir case, eight was the mx). Finally, rule four was not violated as none of the points were closer to the control limits then the centerline. Therefore, this process is under statistic control since none of the rules were violated.

Exercise 13.14.1

Our random number sampler was designed based on which key principle:

a. Data points obtained by taking the average of several random numbers are more likely to fit a Gaussian distribution.
b. The area under a Gaussian distribution is always between zero and one and equal to the 'y' value of the CDF.
c. The distribution of data points in a Gaussian distribution is caused by random uncertainties.
d. A process under Gaussian statistical control will not have a data sample exceed the upper or lower control limit

Answer available at the end of this section.

Exercise 13.14.2

Which of the following would decrease the p-value if comparing two different data sets:

a. Smaller sample size
b. Smaller count of averaged random numbers
c. Smaller standard deviation
d. Smaller difference in mean values

Answer available at the end of this section.

References

- Box, George E., William G. Hunter, and J S. Hunter. Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building. New York: John Wiley & Sons. 43-45.
- Liptak, Bela G. "Process Control and Optimization." Instrument Engineers' Handbook 4: 405-413.
- Wheeler, Donald J., and David S. Chambers. Understanding Statistical Process Control. 2nd ed. Knoxville: SPC P. 37-88.
- Woolf, Peter, Amy Keating, Christopher Burge, and Michael Yaffe. Statistics and Probability Primer for Computational Biologists. Massachusetts Institute of Technology. 2004.

Answers

Exercise 13.14.1

A

Exercise 13.14.2

C

This page titled 13.14: Random sampling from a stationary Gaussian process is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Andrew Laskowski, Emily Yatch, Halley Crast, & Maurice Telesford via source content that was edited to the style and standards of the LibreTexts platform.

14: DESIGN OF EXPERIMENTS

14.1: Design of Experiments via Taguchi Methods - Orthogonal Arrays
14.2: Design of experiments via factorial designs
14.3: Design of Experiments via Random Design
14.4: Summary- Summary on Control Architectures' philosophies, advantages, and disadvantages.

This page titled 14: Design of Experiments is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

14.1: Design of Experiments via Taguchi Methods - Orthogonal Arrays

The **Taguchi method** involves reducing the variation in a process through robust design of experiments. The overall objective of the method is to produce high quality product at low cost to the manufacturer. The Taguchi method was developed by Dr. Genichi Taguchi of Japan who maintained that variation. Taguchi developed a method for designing experiments to investigate how different parameters affect the mean and variance of a process performance characteristic that defines how well the process is functioning. The experimental design proposed by Taguchi involves using orthogonal arrays to organize the parameters affecting the process and the levels at which they should be varies. Instead of having to test all possible combinations like the factorial design, the Taguchi method tests pairs of combinations. This allows for the collection of the necessary data to determine which factors most affect product quality with a minimum amount of experimentation, thus saving time and resources. The Taguchi method is best used when there is an intermediate number of variables (3 to 50), few interactions between variables, and when only a few variables contribute significantly.

The Taguchi arrays can be derived or looked up. Small arrays can be drawn out manually; large arrays can be derived from deterministic algorithms. Generally, arrays can be found online. The arrays are selected by the number of parameters (variables) and the number of levels (states). This is further explained later in this article. Analysis of variance on the collected data from the Taguchi design of experiments can be used to select new parameter values to optimize the performance characteristic. The data from the arrays can be analyzed by plotting the data and performing a visual analysis, ANOVA, bin yield and Fisher's exact test, or Chi-squared test to test significance.

In this article, the specific steps involved in the application of the Taguchi method will be described and examples of using the Taguchi method to design experiments will be given.

Summary of Taguchi Method

Philosophy of the Taguchi Method

1. **Quality should be designed into a product, not inspected into it.** Quality is designed into a process through system design, parameter design, and tolerance design. Parameter design, which will be the focus of this article, is performed by determining what process parameters most affect the product and then designing them to give a specified target quality of product. Quality "inspected into" a product means that the product is produced at random quality levels and those too far from the mean are simply thrown out.
2. **Quality is best achieved by minimizing the deviation from a target. The product should be designed so that it is immune to uncontrollable environmental factors.** In other words, the signal (product quality) to noise (uncontrollable factors) ratio should be high.
3. **The cost of quality should be measured as a function of deviation from the standard and the losses should be measured system wide.** This is the concept of the loss function, or the overall loss incurred upon the customer and society from a product of poor quality. Because the producer is also a member of society and because customer dissatisfaction will discourage future patronage, this cost to customer and society will come back to the producer.

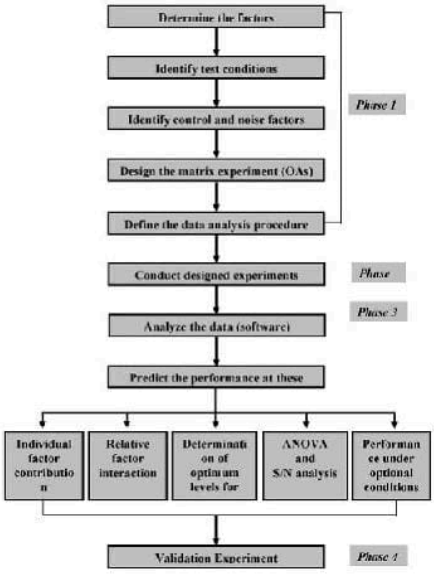
Taguchi Method Design of Experiments

The general steps involved in the Taguchi Method are as follows:

1. Define the process objective, or more specifically, a target value for a performance measure of the process. This may be a flow rate, temperature, etc. The target of a process may also be a minimum or maximum; for example, the goal may be to maximize the output flow rate. The deviation in the performance characteristic from the target value is used to define the loss function for the process.
2. Determine the design parameters affecting the process. Parameters are variables within the process that affect the performance measure such as temperatures, pressures, etc. that can be easily controlled. The number of levels that the parameters should be varied at must be specified. For example, a temperature might be varied to a low and high value of 40 C and 80 C. Increasing the number of levels to vary a parameter at increases the number of experiments to be conducted.

3. Create orthogonal arrays for the parameter design indicating the number of and conditions for each experiment. The selection of orthogonal arrays is based on the number of parameters and the levels of variation for each parameter, and will be expounded below.
4. Conduct the experiments indicated in the completed array to collect data on the effect on the performance measure.
5. Complete data analysis to determine the effect of the different parameters on the performance measure.

See below for a pictorial depiction of these and additional possible steps, depending on the complexity of the analysis.



A detailed description of the execution of these steps will be discussed next.

Taguchi Loss Function

The goal of the Taguchi method is to reduce costs to the manufacturer and to society from variability in manufacturing processes. Taguchi defines the difference between the target value of the performance characteristic of a process, τ, and the measured value, y, as a loss function as shown below.

$$l(y) = k_c(y - \tau)^2$$

The constant, k_C, in the loss function can be determined by considering the specification limits or the acceptable interval, delta.

$$k_c = \frac{C}{\Delta^2}$$

The difficulty in determining k_C is that τ and C are sometimes difficult to define.

If the goal is for the performance characteristic value to be minimized, the loss function is defined as follows:

$$l(y) = k_c y^2$$

where $\tau = 0$.

If the goal is for the performance characteristic value to maximized, the loss function is defined as follows:

$$l(y) = \frac{k_c}{y^2}$$

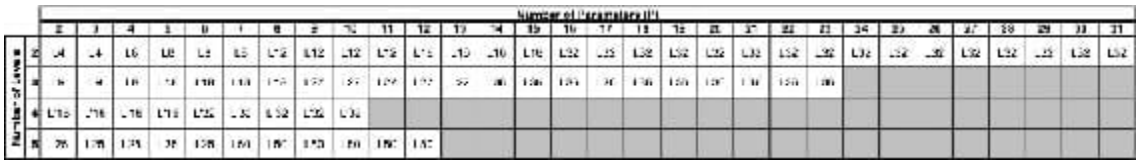
The loss functions described here are the loss to a customer from one product. By computing these loss functions, the overall loss to society can also be calculated.

Determining Parameter Design Orthogonal Array

The effect of many different parameters on the performance characteristic in a condensed set of experiments can be examined by using the orthogonal array experimental design proposed by Taguchi. Once the parameters affecting a process that can be controlled have been determined, the levels at which these parameters should be varied must be determined. Determining what levels of a variable to test requires an in-depth understanding of the process, including the minimum, maximum, and current value of the parameter. If the difference between the minimum and maximum value of a parameter is large, the values being tested can be further apart or more values can be tested. If the range of a parameter is small, then less values can be tested or the values tested can be closer together. For example, if the temperature of a reactor jacket can be varied between 20 and 80 degrees C and it is known that the current operating jacket temperature is 50 degrees C, three levels might be chosen at 20, 50, and 80 degrees C. Also, the cost of conducting experiments must be considered when determining the number of levels of a parameter to include in the experimental design. In the previous example of jacket temperature, it would be cost prohibitive to do 60 levels at 1 degree intervals. Typically, the number of levels for all parameters in the experimental design is chosen to be the same to aid in the selection of the proper orthogonal array.

Knowing the number of parameters and the number of levels, the proper orthogonal array can be selected. Using the array selector table shown below, the name of the appropriate array can be found by looking at the column and row corresponding to the number of parameters and number of levels. Once the name has been determined (the subscript represents the number of experiments that must be completed), the predefined array can be looked up. Links are provided to many of the predefined arrays given in the array selector table. These arrays were created using an algorithm Taguchi developed, and allows for each variable and setting to be tested equally. For example, if we have three parameters (voltage, temperature, pressure) and two levels (high, low), it can be seen the proper array is L4. Clicking on the link L4 to view the L4 array, it can be seen four different experiments are given in the array. The levels designated as 1, 2, 3 etc. should be replaced in the array with the actual level values to be varied and P1, P2, P3 should be replaced with the actual parameters (i.e. voltage, temperature, etc.)

Array Selector



Links to Orthogonal Arrays

The following links are connected to images of the orthogonal array named in the link title:

L4 Array

L8 Array

L9 Array

L12 Array

L16 Array

L'16 Array

L18 Array

L25 Array

L27 Array

L32 Array

L'32 Array

L36 Array

L50 Array

Important Notes Regarding Selection + Use of Orthogonal Arrays

Note 1

The array selector assumes that each parameter has the same number of levels. Sometimes this is not the case. Generally, the highest value will be taken or the difference will be split.

The following examples offer insight on choosing and properly using an orthogonal array. Examples 1 and 2 focus on array choice, while Example 3 will demonstrate how to use an orthogonal array in one of these situations.

Example 1:

parameter: A, B, C, D = 4 # levels: 3, 3, 3, 2 = ~3 array: L9

Example 2:

parameter: A, B, C, D, E, F = 6 # levels: 4, 5, 3, 2, 2, 2 = ~3 array: modified L16

Example 3:

A reactor's behavior is dependent upon impeller model, mixer speed, the control algorithm employed, and the cooling water valve type. The possible values for each are as follows:

Impeller model: A, B, or C

Mixer speed: 300, 350, or 400 RPM

Control algorithm: PID, PI, or P

Valve type: butterfly or globe

There are 4 parameters, and each one has 3 levels with the exception of valve type. The highest number of levels is 3, so we will use a value of 3 when choosing our orthogonal array.

Using the array selector above, we find that the appropriate orthogonal array is L9:

Experiment	P1	P2	P3	P4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

When we replace P1, P2, P3, and P4 with our parameters and begin filling in the parameter values, we find that the L9 array includes 3 levels for valve type, while our system only has 2. The appropriate strategy is to fill in the entries for P4=3 with 1 or 2 in a random, balanced way. For example:

Experiment	Impeller	Motor Speed	Control	Valve
1	A	300	PID	BF
2	A	350	PI	G
3	A	400	P	G
4	B	300	PI	BF
5	B	350	P	BF
6	B	400	PID	G
7	C	300	P	G
8	C	350	PID	BF
9	C	400	PI	BF

Here, the third value was chosen twice as butterfly and once as global.

Note 2

If the array selected based on the number of parameters and levels includes more parameters than are used in the experimental design, ignore the additional parameter columns. For example, if a process has 8 parameters with 2 levels each, the L12 array should be selected according to the array selector. As can be seen below, the L12 Array has columns for 11 parameters (P1-P11). The right 3 columns should be ignored.

Experiment	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	2	2	2	2	2	2
3	1	1	2	2	2	1	1	1	2	2	2
4	1	2	1	2	2	1	2	2	1	1	2
5	1	2	2	1	2	2	1	2	1	2	1
6	1	2	2	1	2	2	1	2	1	2	1
7	1	2	2	2	1	2	2	1	2	1	1
8	2	1	2	1	2	2	2	1	1	1	2
9	2	1	1	2	2	2	1	2	2	1	1
10	2	2	2	1	1	1	1	2	2	1	2
11	2	2	1	2	1	2	1	1	1	2	2
12	2	2	1	1	2	1	2	1	2	2	1

Analyzing Experimental Data

Once the experimental design has been determined and the trials have been carried out, the measured performance characteristic from each trial can be used to analyze the relative effect of the different parameters. To demonstrate the data analysis procedure, the following L9 array will be used, but the principles can be transferred to any type of array.

In this array, it can be seen that any number of repeated observations (trials) may be used. Ti,j represents the different trials with i = experiment number and j = trial number. It should be noted that the Taguchi method allows for the use of a noise matrix including external factors affecting the process outcome rather than repeated trials, but this is outside of the scope of this article.

Experiment Number	P1	P2	P3	P4	T1	T2	...	TN
1	1	1	1	1	$T_{1,1}$	$T_{1,2}$	...	$T_{1,N}$
2	1	2	2	2	$T_{2,1}$	$T_{2,2}$	...	$T_{2,N}$
3	1	3	3	3	$T_{3,1}$	$T_{3,2}$	...	$T_{3,N}$
4	2	1	2	3	$T_{4,1}$	$T_{4,2}$	...	$T_{4,N}$
5	2	2	3	1	$T_{5,1}$	$T_{5,2}$	...	$T_{5,N}$
6	2	3	1	2	$T_{6,1}$	$T_{3,2}$	...	$T_{6,N}$
7	3	1	3	2	$T_{7,1}$	$T_{7,2}$	...	$T_{7,N}$
8	3	2	1	3	$T_{8,1}$	$T_{3,2}$	...	$T_{8,N}$
9	3	3	2	1	$T_{9,1}$	$T_{9,2}$	...	$T_{9,N}$

To determine the effect each variable has on the output, the signal-to-noise ratio, or the SN number, needs to be calculated for each experiment conducted. The calculation of the SN for the first experiment in the array above is shown below for the case of a specific target value of the performance characteristic. In the equations below, yi is the mean value and si is the variance. yi is the value of the performance characteristic for a given experiment.

$$SN_i = 10 \log \frac{\bar{y_i}^2}{{s_i}^2}$$

$$SN_i = 10 \log \frac{\bar{y}_i^2}{s_i^2}$$

where

$$\bar{y}_i = \frac{1}{N_i} \sum_{u=1}^{N_i} y_{i,u}$$

$$s_i^2 = \frac{1}{N_i - 1} \sum_{u=1}^{N_i} (y_{i,u} - \bar{y}_i)$$

and

- i is the experimental number,
- u is the trial number, and
- N_i is the numb er of trials for experiment i

For the case of minimizing the performance characteristic, the following definition of the SN ratio should be calculated:

$$SN_i = -10 \log \left(\sum_{u=1}^{N_i} \frac{y_u^2}{N_i} \right)$$

For the case of maximizing the performance characteristic, the following definition of the SN ratio should be calculated:

$$SN_i = -10 \log \left[\frac{1}{N_i} \sum_{u=1}^{N_i} \frac{1}{y_u^2} \right]$$

After calculating the SN ratio for each experiment, the average SN value is calculated for each factor and level. This is done as shown below for Parameter 3 (P3) in the array:

Experiment Number	P1	P2	P3	P4	SN
1	1	1	1	1	SN1
2	1	2	2	2	SN2
3	1	3	3	3	SN3
4	2	1	2	3	SN4
5	2	2	3	1	SN5
6	2	3	1	2	SN6
7	3	1	3	2	SN7
8	3	2	1	3	SN8
9	3	3	2	1	SN9

$$SN_{P3,1} = \frac{(S_{N1} + S_{N6} + S_{N8})}{3}$$

$$SN_{P3,2} = \frac{(S_{N2} + S_{N4} + S_{N9})}{3}$$

$$SN_{P3,3} = \frac{(S_{N3} + S_{N5} + S_{N7})}{3}$$

Once these SN ratio values are calculated for each factor and level, they are tabulated as shown below and the range R (R = high SN - low SN)of the SN for each parameter is calculated and entered into the table. The larger the R value for a parameter, the larger the effect the variable has on the process. This is because the same change in signal causes a larger effect on the output variable being measured.

Level	P1	P2	P3	P4
1	$SN_{P1,1}$	$SN_{P2,1}$	$SN_{P3,1}$	$SN_{P4,1}$
2	$SN_{P1,2}$	$SN_{P2,2}$	$SN_{P2,2}$	$SN_{P4,2}$
3	$SN_{P1,3}$	$SN_{P2,3}$	$SN_{P3,3}$	$SN_{P4,3}$
Δ	R_{P1}	R_{P2}	R_{P3}	R_{P4}
Rank	...	...	...	...

Please refer to the Worked Out Example for a numeric example of how the data analysis procedure described here is applied.

Advantages and Disadvantages

An advantage of the Taguchi method is that it emphasizes a mean performance characteristic value close to the target value rather than a value within certain specification limits, thus improving the product quality. Additionally, Taguchi's method for experimental design is straightforward and easy to apply to many engineering situations, making it a powerful yet simple tool. It can be used to quickly narrow down the scope of a research project or to identify problems in a manufacturing process from data already in existence. Also, the Taguchi method allows for the analysis of many different parameters without a prohibitively high amount of experimentation. For example, a process with 8 variables, each with 3 states, would require 6561 (3^8) experiments to test all variables. However using Taguchi's orthogonal arrays, only 18 experiments are necessary, or less than .3% of the original number of experiments. In this way, it allows for the identification of key parameters that have the most effect on the performance characteristic value so that further experimentation on these parameters can be performed and the parameters that have little effect can be ignored.

The main disadvantage of the Taguchi method is that the results obtained are only relative and do not exactly indicate what parameter has the highest effect on the performance characteristic value. Also, since orthogonal arrays do not test all variable combinations, this method should not be used with all relationships between all variables are needed. The Taguchi method has been criticized in the literature for difficulty in accounting for interactions between parameters. Another limitation is that the Taguchi methods are offline, and therefore inappropriate for a dynamically changing process such as a simulation study. Furthermore, since Taguchi methods deal with designing quality in rather than correcting for poor quality, they are applied most effectively at early stages of process development. After design variables are specified, use of experimental design may be less cost effective.

Other Methods of Experimental Design

Two other methods for determining experimental design are factorial design and random design. For scenarios with a small number of parameters and levels (1-3) and where each variable contributes significantly, factorial design can work well to determine the specific interactions between variables. However, factorial design gets increasingly complex with an increase in the number of variables. For large systems with many variables (50+) where there are few interactions between variables, random design can be used. Random design assigns each variable a state based on a uniform sample (ex: 3 states = 0.33 probability) for the selected number of experiments. When used properly (in a large system), random design usually produces an experimental design that is desired. However, random design works poorly for systems with a small number of variables.

To obtain a even better understanding of these three different methods, it's good to get a visual of these three methods. It will

illustrate the degree of efficiency for each experimental design depending on the number of variables and the number of states for each variable. The following will have the three experimental designs for the same scenario.

Scenario. You have a CSTR that has four(4) variables and each variable has three or two states. You are to design an experiment to systematically test the effect of each of the variables in the current CSTR.

These variables can take the following values:
Impellers: model A, B, or C
Motor speed for mixer: 300, 350, or 400 RPM
Control algorithm: PID, PI, or P only
Cooling water valve type: butterfly or globe

Experimental Design #1: Factorial Design By looking at the # variables and # states, there should be a total of 54 experiments because (3impellers)(3speeds)(3controllers)(2valves)=54. Here's a list of these 54 experiments:

A, 300,PID, B	A, 300,PI, B	A, 300,P, B	A, 300,PI, G	A, 300,PID, G	A, 300,P, G
B, 300,PID, B	B, 300,PI, B	B, 300,P, B	B, 300,PI, G	B, 300,PID, G	B, 300,P, G
C, 300,PID, B	C, 300,PI, B	C, 300,P, B	C, 300,PI, G	C, 300,PID, G	C, 300,P, G
A, 350,PID, B	A, 350,PI, B	A, 350,P, B	A, 350,PI, G	A, 350,PID, G	A, 350,P, G
B, 350,PID, B	B, 350,PI, B	B, 350,P, B	B, 350,PI, G	B, 350,PID, G	B, 350,P, G
C, 350,PID, B	C, 350,PI, B	C, 350,P, B	C, 350,PI, G	C, 350,PID, G	C, 350,P, G
A, 400,PID, B	A, 400,PI, B	A, 400,P, B	A, 400,PI, G	A, 400,PID, G	A, 400,P, G
B, 400,PID, B	B, 400,PI, B	B, 400,P, B	B, 400,PI, G	B, 400,PID, G	B, 400,P, G
C, 400,PID, B	C, 400,PI, B	C, 400,P, B	C, 400,PI, G	C, 400,PID, G	C, 400,P, G

Experimental Design #2: Taguchi Method Since you know the # of states and variables, you can refer to the table above in this wiki and obtain the correct Taguchi array. It turns out to be a L9 array.

L9

Experiment	P1	P2	P3	P4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

With the actual variables and states, the L9 array should look like the following:

Experiment	Impeller	Moter Speed	Control	Valve
1	A	300	PID	BF
2	A	350	PI	G
3	A	400	P	G
4	B	300	PI	BF
5	B	350	P	BF
6	B	400	PID	G
7	C	300	P	G
8	C	350	PID	BF
9	C	400	PI	BF

Experimental Design #3: Random Design

Since we do not know the number of signal recoveries we want and we don't know the probabilities of each state to happen, it will be difficult to construct a random design table. It will mostly be used for extreme large experiments. Refer to the link below to help you obtain a better grasp on the random design concept.

http://groups.csail.mit.edu/drl/journal_club/papers/CS2-Candes-Romberg-05.pdf

Example 14.1.1

A microprocessor company is having difficulty with its current yields. Silicon processors are made on a large die, cut into pieces, and each one is tested to match specifications. The company has requested that you run experiments to increase processor yield. The factors that affect processor yields are temperature, pressure, doping amount, and deposition rate.

Determine the Taguchi experimental design orthogonal array. The operating conditions for each parameter and level are listed below:

- **A: Temperature**

- A1 = 100°C
- A2 = 150°C (current)
- A3 = 200°C

- **B: Pressure**

- B1 = 2 psi
- B2 = 5 psi (current)
- B3 = 8 psi

- **C: Doping Amount**

- C1 = 4%
- C2 = 6% (current)
- C3 = 8%

- **D: Deposition Rate**

- D1 = 0.1 mg/s
- D2 = 0.2 mg/s (current)
- D3 = 0.3 mg/s

Conducting three trials for each experiment, the data below was collected. Compute the SN ratio for each experiment for the target value case, create a response chart, and determine the parameters that have the highest and lowest effect on the processor yield.

Experiment Number	Temperature	Pressure	Doping Amount	Deposition Rate	Trial 1	Trial 2	Trial 3	Mean
1	100	2	4	0.1	87.3	82.3	70.7	80.1
2	100	5	6	0.2	74.8	70.7	63.2	69.6
3	100	8	8	0.3	56.5	54.9	45.7	52.1
4	150	2	6	0.3	79.8	78.2	62.3	73.4
5	150	5	8	0.1	77.3	76.5	54.9	69.3
6	150	8	4	0.2	89	87.3	63.2	86.5
7	200	2	8	0.2	64.8	62.3	55.7	60.9
8	200	5	4	0.3	89	83.2	87.3	93.2
9	200	8	6	0.1	75.7	71	63.2	71

Solution

a

The L9 orthogonal array should be used. The filled in orthogonal array should look like this:

Experiment Number	Temperature	Pressure	Doping Amount	Deposition Rate
1	100	2	4	0.1
2	100	5	6	0.2
3	100	8	8	0.3
4	150	2	6	0.3
5	150	5	8	0.1
6	150	8	4	0.2
7	200	2	8	0.2
8	200	5	4	0.3
9	200	8	6	0.1

This setup allows the testing of all four variables without having to run 81 [=3^4=(3 Temperatures)(3 Pressures)(3 Doping Amounts)(3 Deposition rates)] separate trials.b) Question:

b) Shown below is the calculation and tabulation of the SN ratio.

$$S_{m1} = \frac{(87.3 + 82.3 + 70.7)^2}{3} = 19248.0$$

$$S_{T1} = 87.3^2 + 82.3^2 + 70.7^2 = 19393.1$$

$$S_{e1} = S_{T1} - S_{m1} = 19393.1 - 19248.0 = 145.0$$

$$V_{e1} = \frac{S_{e1}}{N-1} = \frac{145.1}{2} = 72.5$$

Experiment Number	A (temp)	B (pres)	C (dop)	D (dep)	T1	T2	T3	SN
1	1	1	1	1	87.3	82.3	70.7	19.5
2	1	2	2	2	74.8	70.7	63.2	21.4
3	1	3	3	3	56.5	54.0	45.7	19.3
4	2	1	2	3	79.8	78.2	62.3	17.6
5	2	2	3	1	77.3	76.5	54.0	14.3
6	2	3	1	2	89.0	87.3	83.2	29.2
7	3	1	3	2	64.8	62.3	55.7	22.2
8	3	2	1	3	99.0	93.2	87.3	24.0
9	3	3	2	1	75.7	74.0	63.2	20.4

Shown below is the response table. This table was created by calculating an average SN value for each factor. A sample calculation is shown for Factor B (pressure):

Experiment Number	A (temp)	B (pres)	C (dop)	D (dep)	SN
1	1	1	1	1	19.5
2	1	2	2	2	21.4
3	1	3	3	3	19.3
4	2	1	2	3	17.6
5	2	2	3	1	14.3
6	2	3	1	2	29.2
7	3	1	3	2	22.2
8	3	2	1	3	24.0
9	3	3	2	1	20.4

$$SN_{B1} = \frac{(19.5 + 17.6 + 22.2)}{3} = 19.8$$

$$SN_{B2} = \frac{(21.4 + 14.3 + 24.0)}{3} = 19.9$$

$$SN_{B3} = \frac{(19.3 + 29.2 + 20.4)}{3} = 23.0$$

The effect of this factor is then calculated by determining the range:

$$\Delta = Max - Min = 23.0 - 19.8 = 3.2$$

Level	A (temp)	B (pres)	C (dop)	D (dep)
1	20.1	19.8	24.2	18.1
2	20.4	19.9	19.8	24.3
3	22.2	23.0	18.6	20.3
Δ	2.2	3.2	5.6	6.3
Rank	4	3	2	1

It can be seen that deposition rate has the largest effect on the processor yield and that temperature has the smallest effect on the processor yield.

NOTE: *The data for the following example is not real, and details pertaining to microprocessor fabrication may not be completely accurate.*

a) Solution:

Example 14.1.2: Sesame Seed Suffering

You have just produced one thousand 55 gallon drums of sesame oil for sale to your distributors. However, just before you are to ship oil, one of your employees remembers that one of the oil barrels was temporarily used to store insecticide and is almost surely contaminated. Unfortunately, all of the barrels look the same.

One barrel of sesame oil sells for $1000, while each assay for insecticide in food oil costs $1200 and takes 3 days. Tests for insecticide are extremely expensive. What do you do?

Solution

Extreme multiplexing. This is similar to using a Taguchi method but optimized for very sparse systems and specific cases. For example, instead of 1000 barrels, let us consider 8 barrels for now, one of which is contaminated. We could test each one, but that would be highly expensive. Another solution is to mix samples from each barrel and test the mixtures.

- Mix barrels 1,2,3,4 ---> Sample A
- Mix barrels 1,2,5,6 ---> Sample B
- Mix barrels 1,3,5,7 ---> Sample C

We claim that from testing only these three mixtures, we can determine which of the 8 barrels was contaminated. Let us consider some possible results of these tests. We will use the following label scheme: +/-,+/-,+/- in order of A, B, C. Thus, +,-,+ indicates A and C showed contamination but not B.

- Possible Result 1: -,-,- The only barrel not mixed in was #8, so it is contaminated.
- Possible Result 2: +,-,- Barrel #4 appears in A, but not in B and C. Since only A returned positive, barrel #4 was contaiminated.
- Possible Result 3: -,+,- Barrel #6 appears in B, but not in A and C. Since only B returned positive, barrel #6 was contaminated.

We can see that we have 2^3 = 8 possible results, each of which corresponds to a particular barrel being contaminated. We leave the rest of the cases for the reader to figure out.

Solution with 1,000 barrels: Mix samples from each barrel and test mixtures. Each mixture will consist of samples from a unique combination of 500 barrels. Experiments required = $log_2(1000)$=~10.

Solution with 1,000,000 barrels: Experiments required = $log_2(1000000)$=~20.

Thus, by using extreme multiplexing, we can greatly reduce the # of experiments needed, since the # of experiments scales with log_2(# of barrels) instead of # of barrels.

Exercise 14.1.1

You are producing decaffeinated coffee using supercritical carbon dioxide as a solvent. To optimize the effectiveness of caffeine removal, you would like to test 2 different possible pressures of CO2, 3 possible temperatures, 3 ratios of CO2 to coffee beans, 3 residence times of supercritical CO2 with beans, and 2 different procedures for preroasting the beans prior to caffeine extraction. What is the most appropriate method to design your optimization experiments?

a. Factorial design
b. Taguchi methods
c. Random design
d. None of the above. Coffee should never be decaffeinated

Answer available at the end of this section.

Exercise 14.1.2

Which of the following is inconsistent with the Taguchi philosophy of quality control?

a. Variation is the opposite of quality
b. Interactions between parameters affecting product quality are unimportant and should never be considered
c. Customer dissatisfaction is the most important measure for process variation, and should be minimized
d. A high signal to noise ratio (SN) should be present in a process to make it robust against uncontrollable factors that would cause variation

Answer available at the end of this section.

References

- Vuchkov, I.N. and Boyadjieva, L.N. *Quality Improvement with Design of Experiments: A Response Surface Approach*. Kluwer Academic Publishers. Dordrecht, 2001.
- http://www.ee.iitb.ac.in/~apte/CV_PRA_TAGUCHI_INTRO.htm
- http://www.york.ac.uk/depts/maths/tables/orthogonal.htm
- www.freequality.org/sites/www_freequality_org/documents/tools/Tagarray_files/tamatrix.htm
- Roy, R. K., A Primer on the Taguchi Method, Van Nostrand Reinhold, New York 1990.
- Ravella Sreenivas Rao, C. Ganesh Kumar, R. Shetty Prakasham, Phil J. Hobbs (2008) The Taguchi methodology as a statistical tool for biotechnological applications: A critical appraisal Biotechnology Journal 3:510–523.

Answers

Exercise 14.1.1

To perform a full factorial experiment, you'd need to account for: (2 pressures)(3 temperatures)(3 ratios)(3 times)(2 roasts) = 108 experiments! Thus, factorial design is not a practical choice: a good rule of thumb is 1-3 variables with few states for a manageable factorial analysis. However, selecting 3 for the number of levels and consulting the array selector, we see that an L18 array will suffice for a Taguchi analysis. 18 is a much more feasible number of experiments than 108. Finally, random design is only practical when >50 variables are involved and few will have significant impact on the system, which is not the case here. Thus, although the author is of the opinion that choice 'D' has significant merit, the correct answer is 'B'.

Exercise 14.1.2

Choice 'A' is another way of expressing point 3 of the Taguchi philosophy above. Similarly, choice 'C' is another way of expressing the cost function, and choice 'D' is consistent with philosophy 2. Although it is true that the Taguchi methods do not account well for interactions between variables, choice 'B' is incorrect for two reasons. 1) Pair-wise interactions between variables are accounted for in Taguchi orthogonal arrays. 2) Negligibility of interactions between parameters is not so much a central tenet of Taguchi philosophy as a necessary simplifying assumption. An analog is decoupling of MIMO systems. Since applying this assumption yields good results, it is consistent with Taguchi thinking to say that makes it valid.

This page titled 14.1: Design of Experiments via Taguchi Methods - Orthogonal Arrays is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Ben Terrien, John Zalewski, Mike Oom, & Stephanie Fraley via source content that was edited to the style and standards of the LibreTexts platform.

14.2: Design of experiments via factorial designs

Factorial design is an important method to determine the effects of multiple variables on a response. Traditionally, experiments are designed to determine the effect of ONE variable upon ONE response. R.A. Fisher showed that there are advantages by combining the study of multiple variables in the same factorial experiment. Factorial design can reduce the number of experiments one has to perform by studying multiple factors simultaneously. Additionally, it can be used to find both main effects (from each independent factor) and interaction effects (when both factors must be used to explain the outcome). However, factorial design can only give relative values, and to achieve actual numerical values the math becomes difficult, as regressions (which require minimizing a sum of values) need to be performed. Regardless, factorial design is a useful method to design experiments in both laboratory and industrial settings.

Factorial design tests all possible conditions. Because factorial design can lead to a large number of trials, which can become expensive and time-consuming, factorial design is best used for a small number of variables with few states (1 to 3). Factorial design works well when interactions between variables are strong and important and where every variable contributes significantly.

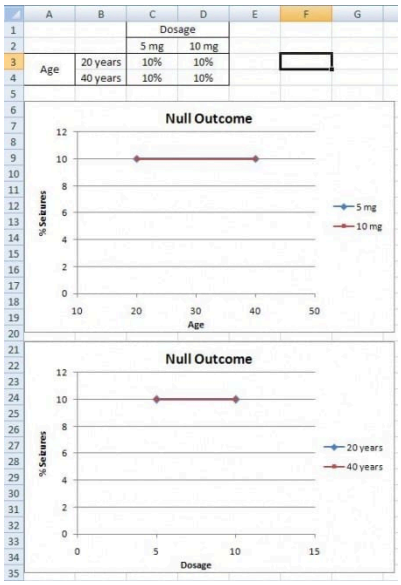
What is Factorial Design?

Factorial Design Example

The easiest way to understand how factorial design works is to read an example. Suppose that you, a scientist working for the FDA, would like to study and measure the probability of patients suffering from seizures after taking a new pharmaceutical drug called CureAll. CureAll is a novel drug on the market and can cure nearly any ailment of the body. You along with your co-workers at the FDA have decided to test two dosage levels: 5 mg and 10 mg. You are also interested in determining whether the drug side-effects differ between younger adults (age 20) and older adults (age 40). Based on the given information, you see that there are two factors: dosage and age. Factors are the main categories to explore when determining the cause of seizures in patients. Under each of these factors, there are different levels: 5 and 10 mg for the dosage; 20 and 40 years for age. A level is basically one of the subdivisions that make up a factor. From this information, we can see that we have a 2 x 2 factorial design, which means that we will have 2 * 2 = 4 groups. A group is set of conditions that will make up that particular experiment.

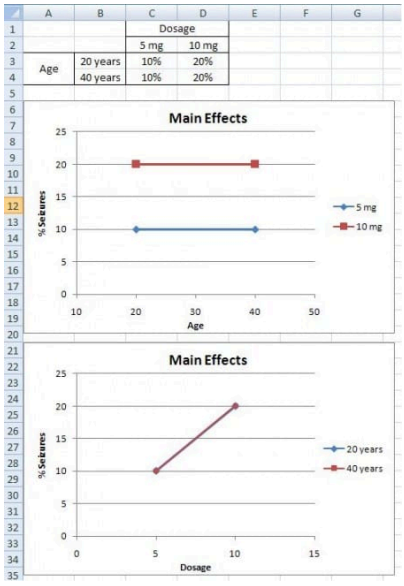
Null Outcome

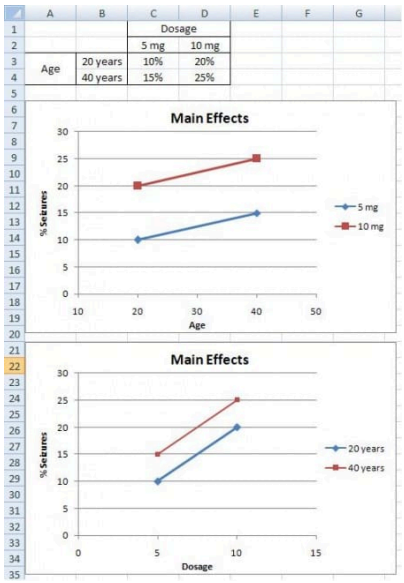
A null outcome situation is when the outcome of your experiment is the same regardless of how the levels within your experiment were combined. From the example above, a null outcome would exist if you received the same percentage of seizures occurring in patients with varying dose and age. The graphs below illustrate no change in the percentage of seizures for all factors, so you can conclude that the chance of suffering from a seizure is not affected by the dosage of the drug or the age of the patient.



Main Effects

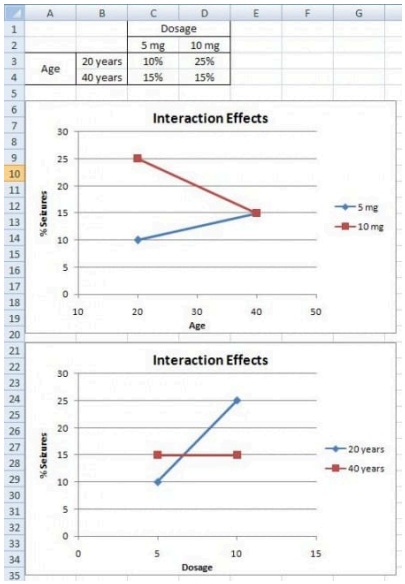
A main effects situation is when there exists a consistent trend among the different levels of a factor. From the example above, suppose you find that as dosage increases, the percentage of people who suffer from seizures increases as well. You also notice that age does not play a role; both 20 and 40 year olds suffer the same percentage of seizures for a given amount of CureAll. From this information, you can conclude that the chance of a patient suffering a seizure is minimized at lower dosages of the drug (5 mg). The second graph illustrates that with increased drug dosage there is an increased percentage of seizures, while the first graph illustrates that with increased age there is no change in the percentage of seizures. Both of these graphs only contain one main effect, since only dose has an effect the percentage of seizures. Whereas, graphs three and four have two main effects, since dose and age both have an effect on the percentage of seizures.





Interaction Effects

The interaction effects situation is the last outcome that can be detected using factorial design. From the example above, suppose you find that 20 year olds will suffer from seizures 10% of the time when given a 5 mg CureAll pill, while 20 year olds will suffer 25% of the time when given a 10 mg CureAll pill. When 40 year olds, however, are given a 5 mg pill or a 10 mg pill, 15% suffer from seizures at both of these dosages. This correlation can be seen in the graphs below. There is an increasing chance of suffering from a seizure at higher doses for 20 year olds, but no difference in suffering from seizures for 40 year olds. Thus, there must be an interaction effect between the dosage of CureAll, and the age of the patient taking the drug. When you have an interaction effect it is impossible to describe your results accurately without mentioning both factors. You can always spot an interaction in the graphs because when there are lines that are not parallel an interaction is present. If you observe the main effect graphs above, you will notice that all of the lines within a graph are parallel. In contrast, for interaction effect graphs, you will see that the lines are not parallel.



Mathematical Analysis Approach

In the previous section, we looked at a qualitative approach to determining the effects of different factors using factorial design. Now we are going to shift gears and look at factorial design in a quantitative approach in order to determine how much influence

the factors in an experiment have on the outcome.

How to Deal with a 2^n Factorial Design

Suppose you have two variables A and B and each have two levels a_1, a_2 and b_1, b_2. You would measure combination effects of A and B (a_1b_1, a_1b_2, a_2b_1, a_2b_2). Since we have two factors, each of which has two levels, we say that we have a 2 x 2 or a 2^2 factorial design. Typically, when performing factorial design, there will be two levels, and *n* different factors. Thus, the general form of factorial design is 2^n.

In order to find the main effect of A, we use the following equation:

$$A = (a_2b_1 - a_1b_1) + (a_2b_2 - a_1b_2)$$

Similarly, the main effect of B is given by:

$$B = (b_2a_1 - b_1a_1) + (b_2a_2 - b_1a_2)$$

By the traditional experimentation, each experiment would have to be isolated separately to fully find the effect on B. This would have resulted in 8 different experiments being performed. Note that only four experiments were required in factorial designs to solve for the eight values in A and B. This shows how factorial design is a timesaver.

By taking the coefficients in A and B, the table below was created.

	Coefficient Effect			
	a_1b_1	a_2b_1	a_1b_2	a_2b_2
A	-1	+1	-1	+1
B	-1	-1	+1	+1
AB	+1	-1	-1	+1

AB is found by multiplying the coefficients of a_xb_x to get the new coefficient effect.

The additional complication is the fact that more than one trial/replication is required for accuracy, so this requires adding up each sub-effect (e.g adding up the three trials of a_1b_1). By adding up the coefficient effects with the sub-effects (multiply coefficient with sub-effect), a **total factorial effect** can be found. This value will determine if the factor has a significant effect on the outcome. For larger numbers, the factor can be considered extremely important and for smaller numbers, the factor can be considered less important. The sign of the number also has a direct correlation to the effect being positive or negative.

To get a mean factorial effect, the totals needs to be divided by 2 times the number of **replicates**, where a replicate is a repeated experiment.

$$\text{mean factorial effect} = \frac{\text{total factorial effect}}{2r}$$

By adding a third variable (C), the process of obtaining the coefficients becomes significantly complicated. The main factorial effect for A:

$$A = (a_2b_1c_1 - a_1b_1c_1) + (a_2b_2c_1 - a_1b_2c_1) + (a_2b_1c_2 - a_1b_1c_2) + (a_2b_2c_2 - a_1b_2c_2)$$

The table of coefficients is listed below

	Coefficient Effect							
	$a_1b_1c_1$	$a_2b_1c_1$	$a_1b_2c_1$	$a_2b_2c_1$	$a_1b_1c_2$	$a_2b_1c_2$	$a_1b_2c_2$	$a_2b_2c_2$
A	-1	+1	-1	+1	-1	+1	-1	+1
B	-1	-1	+1	+1	-1	-1	+1	+1
AB	+1	-1	-1	+1	+1	-1	-1	+1
C	-1	-1	-1	-1	+1	+1	+1	+1
AC	+1	-1	+1	-1	-1	+1	-1	+1
BC	+1	+1	-1	-1	-1	-1	+1	+1
ABC	-1	+1	+1	-1	+1	-1	-1	+1

It is clear that in order to find the total factorial effects, you would have to find the main effects of the variable and then the coefficients. Yates Algorithm can be used to simplify the process.

Yates Algorithm

Frank Yates created an algorithm to easily find the total factorial effects in a 2^n factorial that is easily programmable in Excel. While this algorithm is fairly straightforward, it is also quite tedious and is limited to 2^n factorial designs. Thus, modern technology has allowed for this analysis to be done using statistical software programs through regression.

Steps:

1. In the first column, list all the individual experimental combinations

According to the yates order, such as follows for a 2^3 factorial design

- - -

+ - -

- + -

+ + -

- - +

+ - +

- + +

+ + +

2. In the second column, list all the totals for each combination
3. The 1st four entries in the third column (Stage 1) are obtained by adding pairs together from the "totals list" (previous column). The next four numbers are obtained by subtracting the top number from the bottom number of each pair.

Combination	Total		(1)
$a_1b_1c_1$	52	+	119.5
$a_2b_1c_1$	67.5	+	121
$a_1b_2c_1$	56	+	148
$a_2b_2c_1$	65	+	136
$a_1b_1c_2$	72	-	15.5
$a_2b_1c_2$	76	-	9
$a_1b_2c_2$	73	-	4
$a_2b_2c_2$	63	-	-10

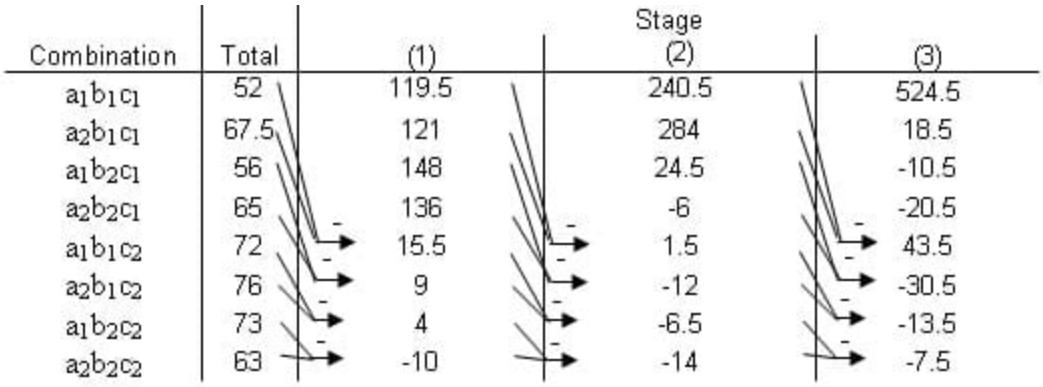
Combination	Total	(1)
$a_1b_1c_1$	52	119.5
$a_2b_1c_1$	67.5	121
$a_1b_2c_1$	56	148
$a_2b_2c_1$	65	136
$a_1b_1c_2$	72	15.5
$a_2b_1c_2$	76	9
$a_1b_2c_2$	73	4
$a_2b_2c_2$	63	-10

4. The fourth column (Stage 2) is obtained in the same fashion, but this time adding and subtracting pairs from Stage 1.

Combination	Total	(1)	Stage (2)
$a_1b_1c_1$	52	119.5	240.5
$a_2b_1c_1$	67.5	121	284
$a_1b_2c_1$	56	148	24.5
$a_2b_2c_1$	65	136	-6
$a_1b_1c_2$	72	15.5	1.5
$a_2b_1c_2$	76	9	-12
$a_1b_2c_2$	73	4	-6.5
$a_2b_2c_2$	63	-10	-14

5. The fifth column (Stage 3) is obtained in the same fashion, but this time adding and subtracting pairs from Stage 2.

Combination	Total	(1)	Stage (2)	(3)
$a_1b_1c_1$	52	119.5	240.5	524.5
$a_2b_1c_1$	67.5	121	284	10.5
$a_1b_2c_1$	56	148	24.5	-10.5
$a_2b_2c_1$	65	136	-6	-20.5
$a_1b_1c_2$	72	15.5	1.5	43.5
$a_2b_1c_2$	76	9	-12	-30.5
$a_1b_2c_2$	73	4	-6.5	-13.5
$a_2b_2c_2$	63	-10	-14	-7.5



6. Continue with Stages until reaching *n*, or the number of factors. This final column is the Effect Total. A positive value means a positive correlation, and a negative values means a negative correlation. These values are all relative, however, so there is no way to compare different effect totals from different experiments.

Ignoring the first row, look in the last stage and find the variable that has the largest relative number, then that row indicates the **MAIN TOTAL EFFECT**. The Main Total Effect can be related to input variables by moving along the row and looking at the first column. If the row in the first column is $a_2b_1c_1$ then the main total effect is A. The row for $a_1b_2c_1$ would be for B. The row for $a_2b_1c_2$ would be for AC.

This main total effect value for each variable or variable combination will be some value that signifies the relationship between the output and the variable. For instance, if your value is positive, then there is a positive relationship between the variable and the output (i.e. as you increase the variable, the output increases as well). A negative value would signify a negative relationship. Notice, however, that the values are all relative to one another. So while the largest main total effect value in one set of experiments may have a value of 128, another experiment may have its largest main total effect value be 43. There is no way to determine if a value of 128 in one experiment has more control over its output than a value of 43 does, but for the purposes of comparing variables within an experiment, the main total effect does allow you to see the relative control the variables have on the output.

Factorial Design Example Revisited

Recall the example given in the previous section What is Factorial Design? In the example, there were two factors and two levels, which gave a 2^2 factorial design. The Yates Algorithm can be used in order to quantitatively determine which factor affects the percentage of seizures the most. For the use of the Yates algorithm, we will call age factor A with a_1 = 20 years, and a_2 = 40 years. Likewise, we will call dosage factor B, with b_1 = 5 mg, and b_2 = 10 mg. The data for the three outcomes is taken from the figures given in the example, assuming that the data given resulted from multiple trials.

Null Outcome

The following Yates algorithm table using the data for the null outcome was constructed. As seen in the table, the values of the main total factorial effect are 0 for A, B, and AB. This proves that neither dosage or age have any effect on percentage of seizures.

Combination	Total	Stage 1	Stage 2	Main Total Factorial Effect
a_1b_1	10	20	40	Doesn't Matter
a_2b_1	10	20	0	A
a_1b_2	10	0	0	B
a_2b_2	10	0	0	AB

Main Effect

The following Yates algorithm table using the data from the first two graphs of the main effects section was constructed. Besides the first row in the table, the row with the largest main total factorial effect is the B row, while the main total effect for A is 0. This means that dosage (factor B) affects the percentage of seizures, while age (factor A) has no effect, which is also what was seen graphically.

Combination	Total	Stage 1	Stage 2	Main Total Factorial Effect
a_1b_1	10	20	60	Doesn't Matter
a_2b_1	10	40	0	A
a_1b_2	20	0	20	B
a_2b_2	20	0	0	AB

The following Yates algorithm table using the data from second two graphs of the main effects section was constructed. Besides the first row in the table, the main total effect value was 10 for factor A and 20 for factor B. This means that both age and dosage affect percentage seizures. However, since the value for B is larger, dosage has a larger effect on percentage of seizures than age. This is what was seen graphically, since the graph with dosage on the horizontal axis has a slope with larger magnitude than the graph with age on the horizontal axis.

Combination	Total	Stage 1	Stage 2	Main Total Factorial Effect
a_1b_1	10	25	70	Doesn't Matter
a_2b_1	15	45	10	A
a_1b_2	20	5	20	B
a_2b_2	25	5	0	AB

Interaction Effect

The following Yates algorithm table was constructed using the data from the interaction effects section. Since the main total factorial effect for AB is non-zero, there are interaction effects. This means that it is impossible to correlate the results with either one factor or another; both factors must be taken into account.

Combination	Total	Stage 1	Stage 2	Main Total Factorial Effect
a_1b_1	10	25	65	Doesn't Matter
a_2b_1	15	40	-5	A
a_1b_2	25	5	15	B
a_2b_2	15	-10	-15	AB

Chemical Engineering Applications

It should be quite clear that factorial design can be easily integrated into a chemical engineering application. Many chemical engineers face problems at their jobs when dealing with how to determine the effects of various factors on their outputs. For example, suppose that you have a reactor and want to study the effect of temperature, concentration and pressure on multiple outputs. In order to minimize the number of experiments that you would have to perform, you can utilize factorial design. This will allow you to determine the effects of temperature and pressure while saving money on performing unnecessary experiments.

A 2007 study on converting wheat straw to fuel utilized factorial design to study the effect of four factors on the composition and susceptibility to enzyme hydrolysis of the final product. (Perez, et.al.). The table below shows the full factorial design for

the study. The four factors that were studied all had only two levels and dealt with pretreatment parameters. They were: Water temperature, residence time, solid fraction and overpressure in the reactor. It is not necessary to understand what each of these are to understand the experimental design. As seen in the table below, there were sixteen trials, or 2^4 experiments.

Table 1. Full factorial design conditions for Liquid Hot Water pretreatment of wheat straw

Full factorial design				
Run	A: Temp (?C)	B: Time (min)	C: Overpressure (bar)	D: Subs.loading (% w/)
1	170	0	0	5
2	200	0	0	5
3	170	40	0	5
4	200	40	0	5
5	170	0	30	5
6	200	0	30	5
7	170	40	30	5
8	200	40	30	5
9	170	0	0	10
10	200	0	0	10
11	170	40	0	10
12	200	40	0	10
13	170	0	30	10
14	200	0	30	10
15	170	40	30	10
16	200	40	30	10

Source: Perez, et. al.

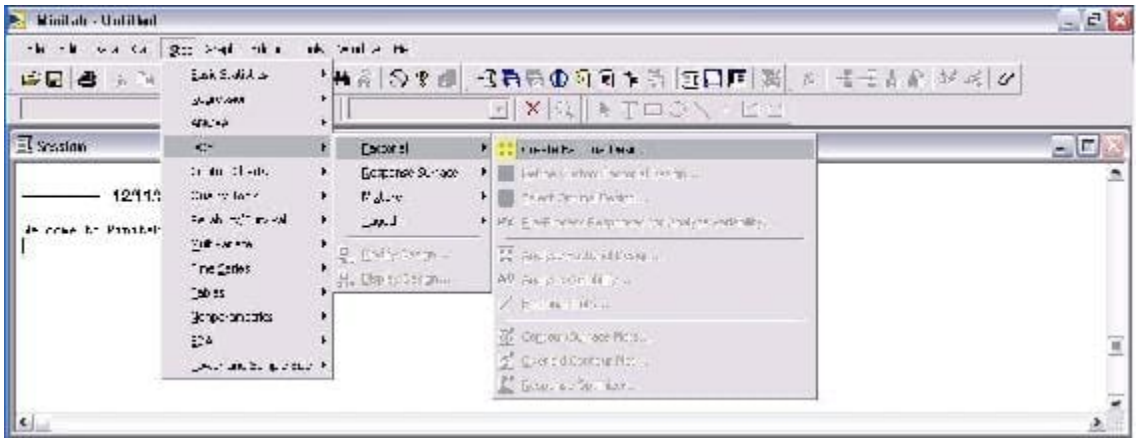
Minitab DOE Example

Minitab 15 Statistical Software is a powerful statistics program capable of performing regressions, ANOVA, control charts, DOE, and much more. Minitab is especially useful for creating and analyzing the results for DOE studies. It is possible to create factorial, response surface, mixture, and taguchi method DOEs in Minitab. The general method for creating factorial DOEs is discussed below.

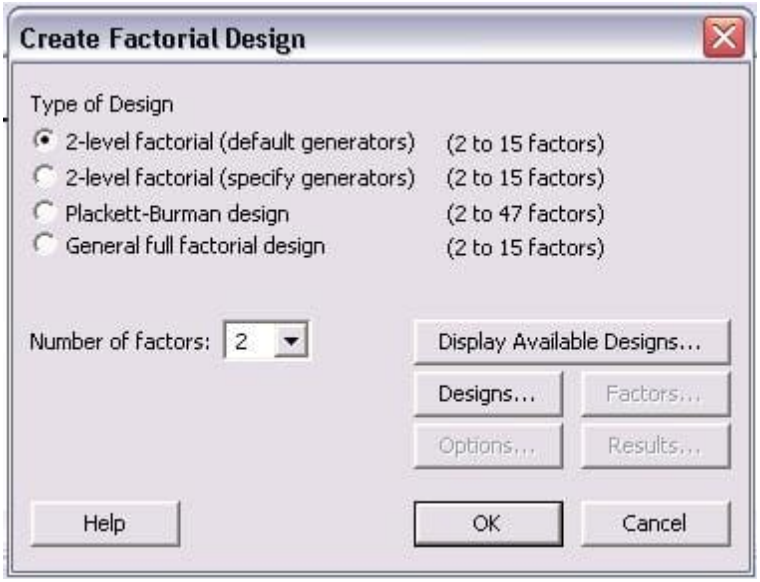
Creating Factorial DOE

Minitab provides a simple and user-friendly method to design a table of experiments. Additionally, analysis of multiple responses (results obtained from experimentation) to determine which parameters significantly affect the responses is easy to do with Minitab. Minitab 15 Statistical Software can be used via Virtual CAEN Labs by going to Start>All Programs>Math and Numerical Methods>Minitab Solutions>Minitab 15 Statistical Software.

The first step is creating the DOE by specifying the number of levels (typically 2) and number of responses. To do this, go to Stat>DOE>Factorial>Create Factorial Design as shown in the image below.

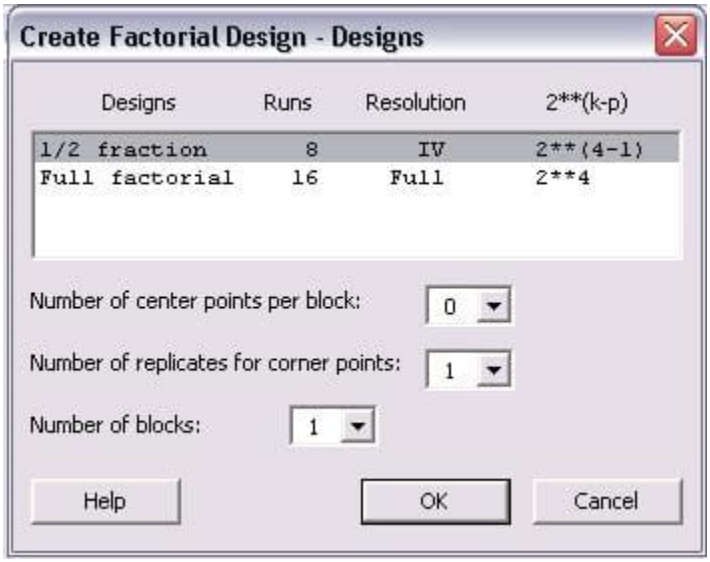


The next image is the "Create Factorial Design" options menu.



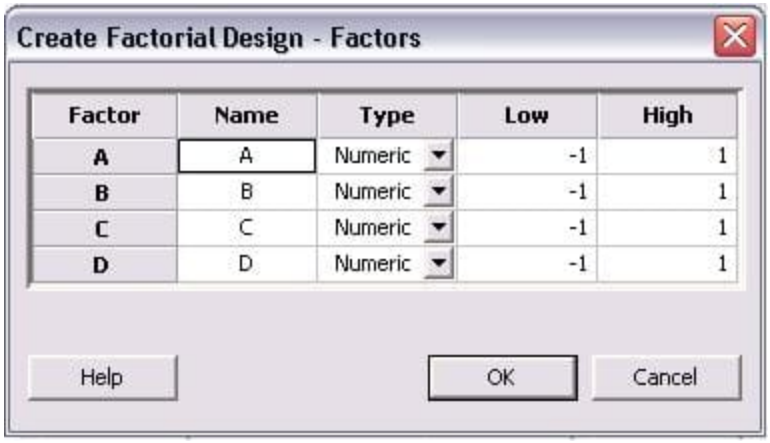
For a 2 level design, click the "2-level factorial (default generators)" radio button. Then specify the number of factors between 2 and 15. Other designs such as Plackett-Burman or a General full factorial design can be chosen. For information about these designs, please refer to the "Help" menu.

After the number of factors is chosen, click on the "Designs..." option to see the following menu.



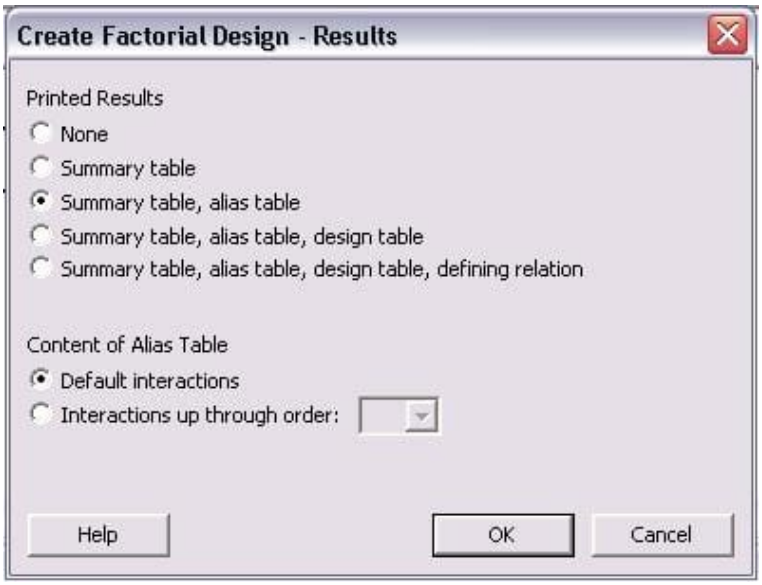
In this menu, a 1/2 fraction or full factorial design can be chosen. Although the full factorial provides better resolution and is a more complete analysis, the 1/2 fraction requires half the number of runs as the full factorial design. In lack of time or to get a general idea of the relationships, the 1/2 fraction design is a good choice. Additionally, the number of center points per block, number of replicates for corner points, and number of blocks can be chosen in this menu. Consult the "Help" menu for details about these options. Click "Ok" once the type of design has been chosen.

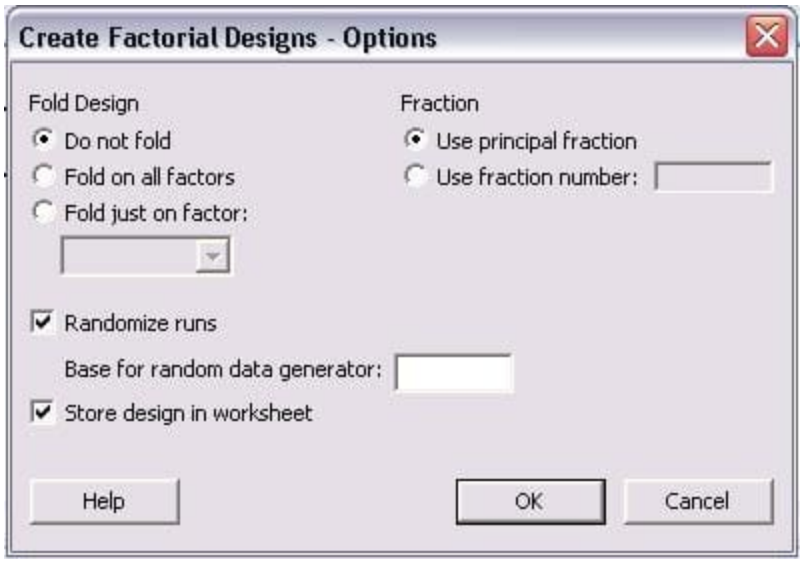
Once the design has been chosen, the "Factors...", "Options..." and "Results..." buttons become active in the "Create Factorial Designs" option menu. Click on "Factors..." button to see the following menu.



The above image is for a 4 factor design. Factors A - D can be renamed to represent the actual factors of the system. The factors can be numeric or text. Additionally, a low and high value are initially listed as -1 and 1, where -1 is the low and 1 is the high value. The low and high levels for each factor can be changed to their actual values in this menu. Click "OK" once this is completed.

The necessary steps for creating the DOE are complete, but other options for "Results..." and "Options..." can be specified. The menus for "Results..." and "Options..." are shown below.





In the main "Create Factorial Design" menu, click "OK" once all specifications are complete. The following table is obtained for a 2-level, 4 factor, full factorial design. None of the levels were specified as they appear as -1 and 1 for low and high levels, respectively.

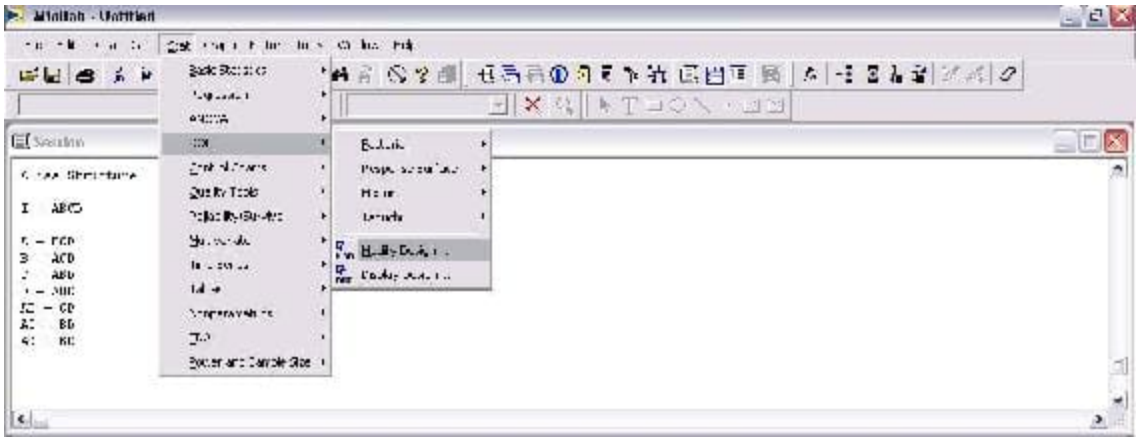
Worksheet 6 ***

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	StdOrder	RunOrder	CenterPt	Blocks	A	B	C	D		
1	11	1	1	*	1	1	1	*		
2	6	2	1	*	1	-1	1	*		
3	2	3	1	*	1	-1	-1	*		
4	16	4	1	*	1	1	1	*		
5	5	5	1	*	-1	-1	1	*		
6	7	6	1	*	1	1	-1	*		
7	1	7	1	*	-1	-1	-1	*		
8	13	8	1	*	-1	-1	1	*		
9	16	9	1	*	1	1	1	*		
10	12	10	1	*	1	1	-1	*		
11	3	11	1	*	-1	1	-1	*		
12	7	12	1	*	-1	1	1	*		
13	14	13	1	*	1	-1	1	*		
14	10	14	1	*	1	-1	-1	*		
15	8	15	1	*	1	1	1	*		
16	9	16	1	*	-1	-1	-1	*		
17										

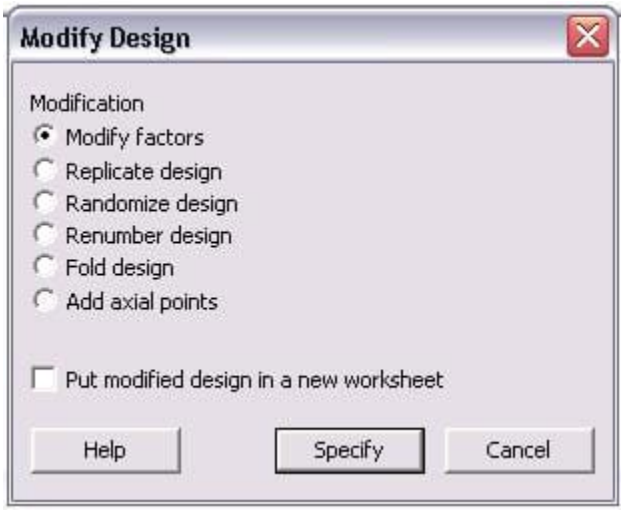
The above table contains all the conditions required for a full factorial DOE. Minitab displays the standard order and randomized run order in columns C1 and C2, respectively. Columns A-D are the factors. The first run (as specified by the random run order) should be performed at the low levels of A and C and the high levels of B and D. A total of 16 runs are required to complete the DOE.

Modifying DOE Table

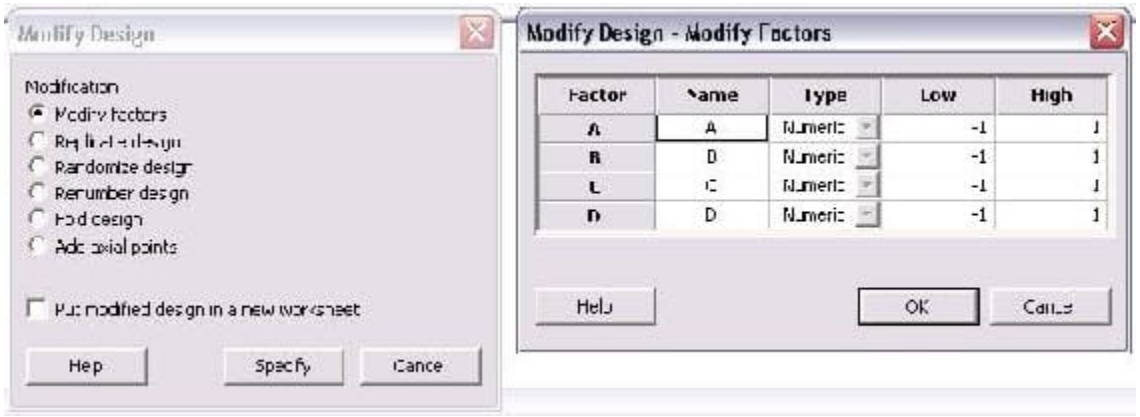
Once a table of trials for the DOE has been created, additional modifications can be made as needed. Some typical modifications include modifying the name of each factors, specifying the high and low level of each factor, and adding replicates to the design. To being modifications of a current design, go to Stat>DOE>Modify Design... as seen in the figure below.



The following menu is displayed for modifying the design.

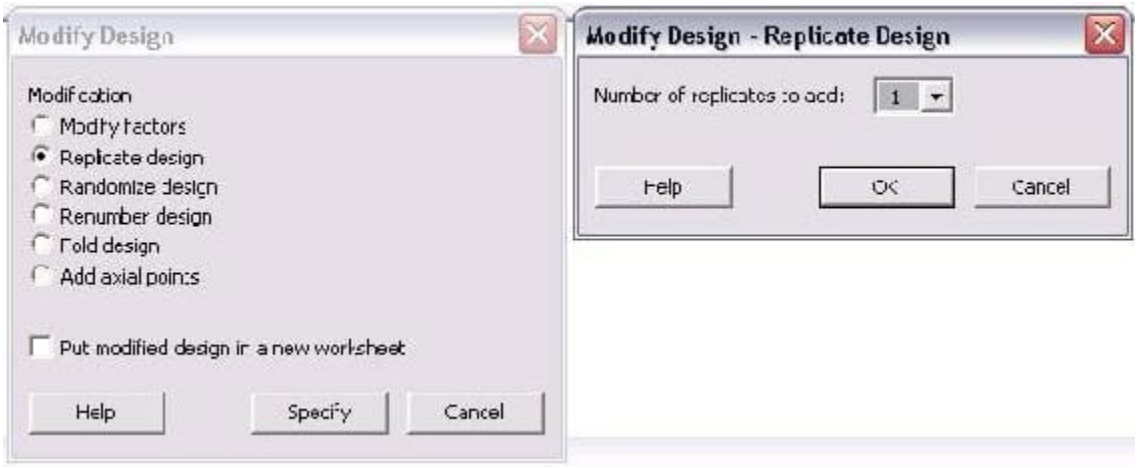


In the "Modify Design" menu, users can modify factors, replicate design, randomize design, renumber design, fold design, and add axial points. Additionally, any changes made can be put into a new worksheet. To change the factors, click the "Modify factors" radio button and then "Specify" to see the following options menu.



The default factors are named "A", "B", "C", and "D" and have respective high and low levels of 1 and -1. The name of the factors can be changed by simply clicking in the box and typing a new name. Additionally, the low and high levels for each factor can be modified in this menu. Since the high and low levels for each factor may not be known when the design is first created, it is convenient to be able to define them later. Click "OK" after modifications are complete.

Another typical modification is adding replicates to a design. Replicates are repeats of each trial that help determine the reproducibility of the design, thus increasing the number of trials and accuracy of the DOE. To add replicates, click the "Replicate design" radio button in the "Modify Design" menu. The following menu will be displayed.

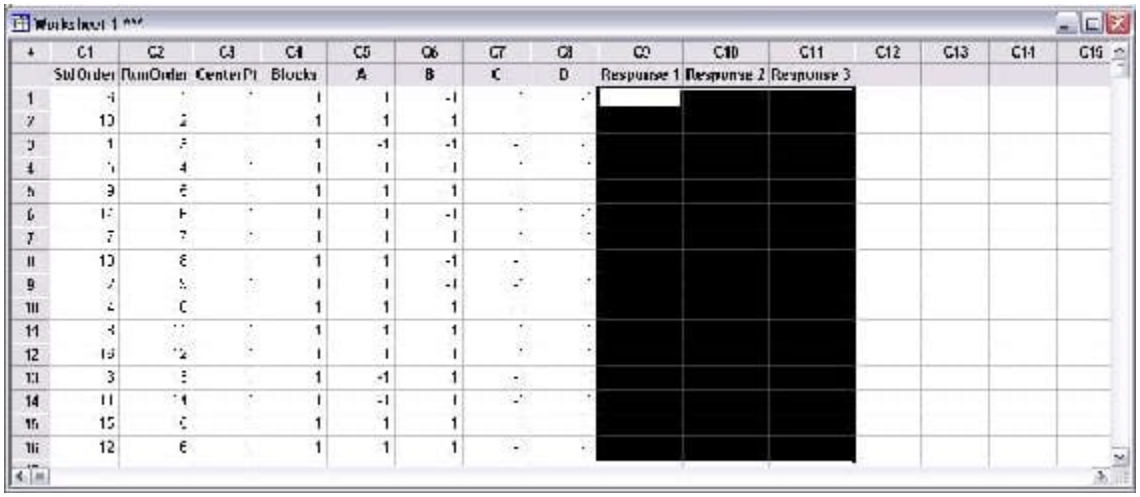


The only option in this menu is the number of replicates to add. The number ranges between 1 and 10. To have a total of 3 trials of each, the user should add 2 replicates in this menu. If 4 replicates are added, there will be a total of 5 trials of each. Typically, if the same experimentation will occur for 3 lab periods, 2 replicates will be added.

Additional modifications to the design include randomizing and renumbering the design. These are very straightforward modifications which affect the ordering of the trials. For information about the "Fold design" and "Add axial points", consult the "Help" menu.

Analyzing DOE Results

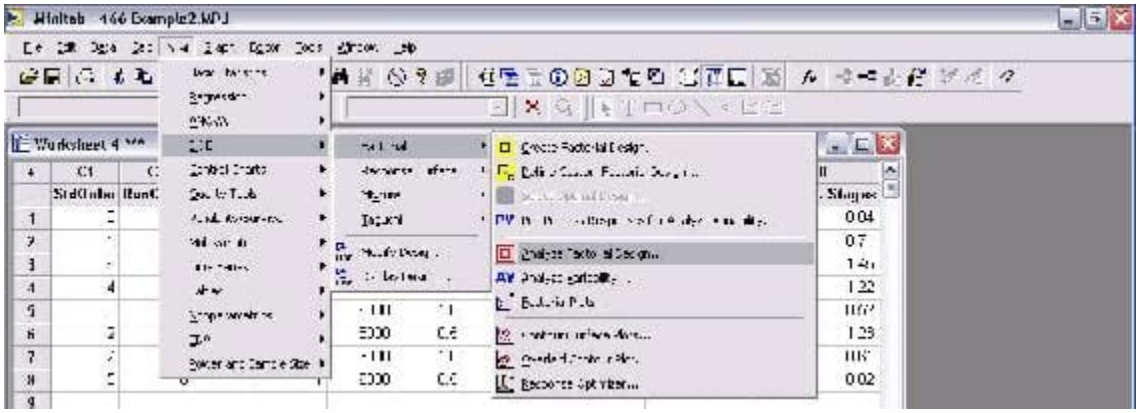
After the complete DOE study has been performed, Minitab can be used to analyze the effect of experimental results (referred to as responses) on the factors specified in the design. The first step in analyzing the results is entering the responses into the DOE table. This is done much like adding data into an Excel data sheet. In the columns to the right of the last factor, enter each response as seen in the figure below.



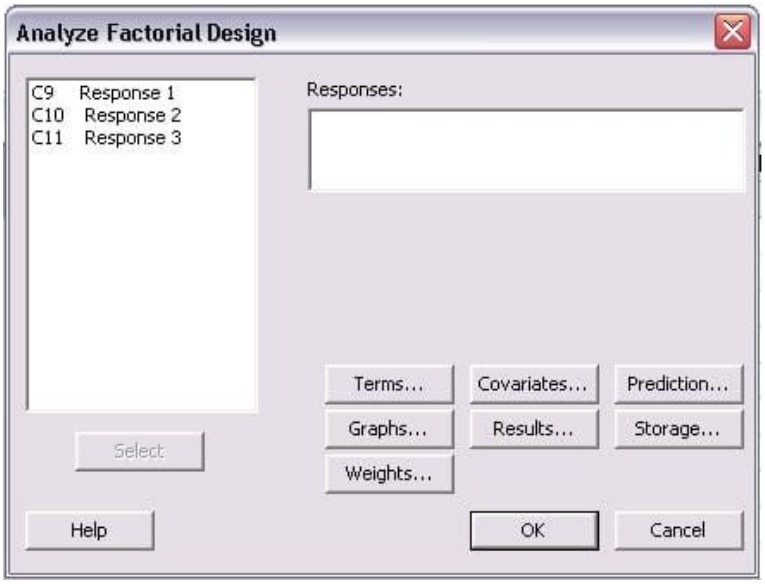
The above figure contains three response columns. The names of each response can be changed by clicking on the column name and entering the desired name. In the figure, the area selected in black is where the responses will be inputted. For instance, if the purity, yield, and residual amount of catalyst was measured in the DOE study, the values of these for each trial would be entered in the columns.

Once the responses are entered, statistical analysis on the data can be performed. Go to Stat>DOE>Factorial>Analyze Factorial

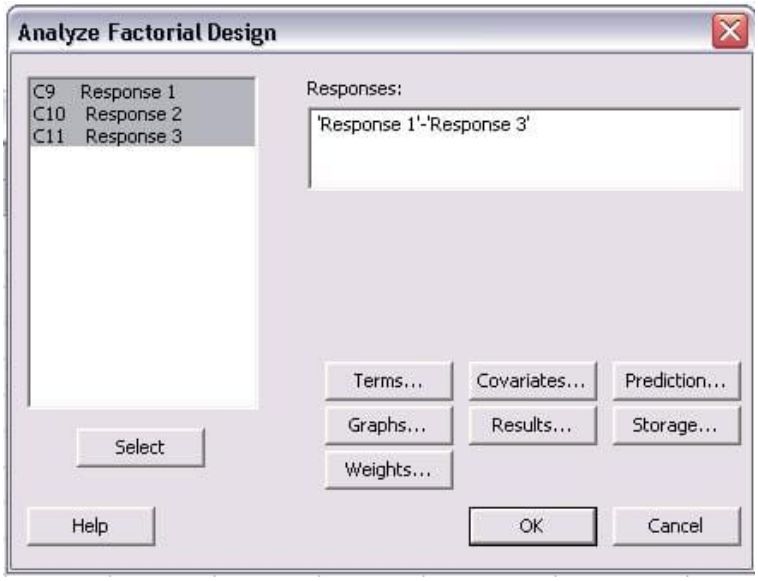
Design... as seen in the following image.



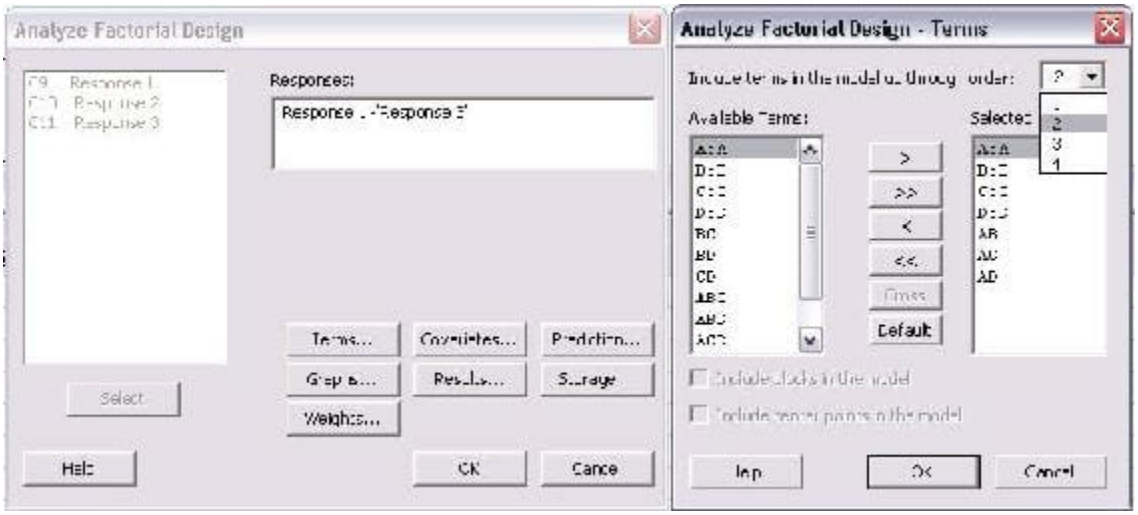
The menu that appears for analyzing factorial design is shown below.



In the "Analyze Factorial Design" menu, the responses are shown on the left of the screen. The first step is to choose the responses to be analyzed. All of the responses can be chosen at once or individually. To choose them, click (or click and drag to select many) and then click "Select" to add them into the "Responses:" section as seen below.

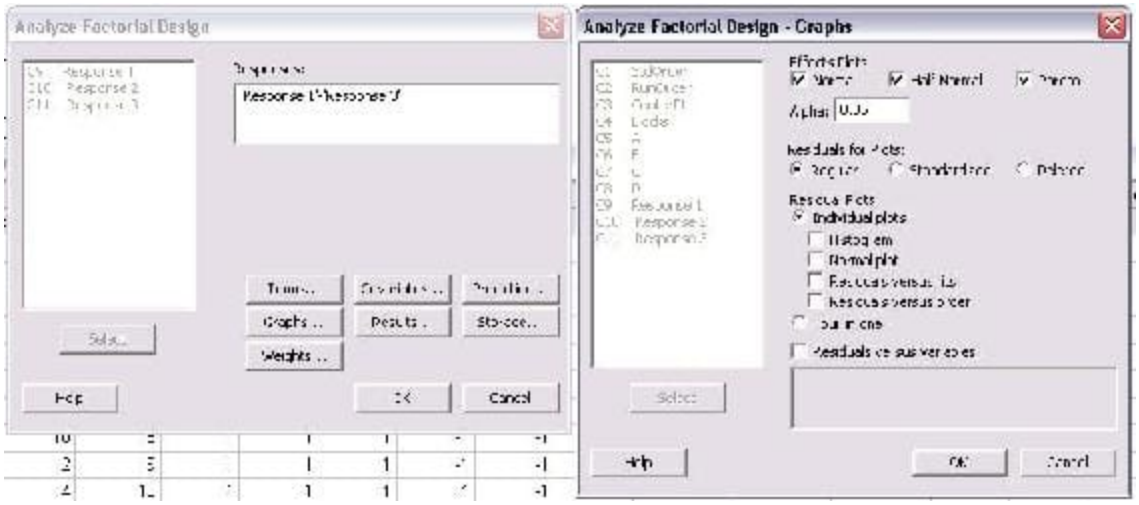


The next step is selecting which terms will be analyzed for the responses. To do this, click on "Terms..." and the following menu will appear.



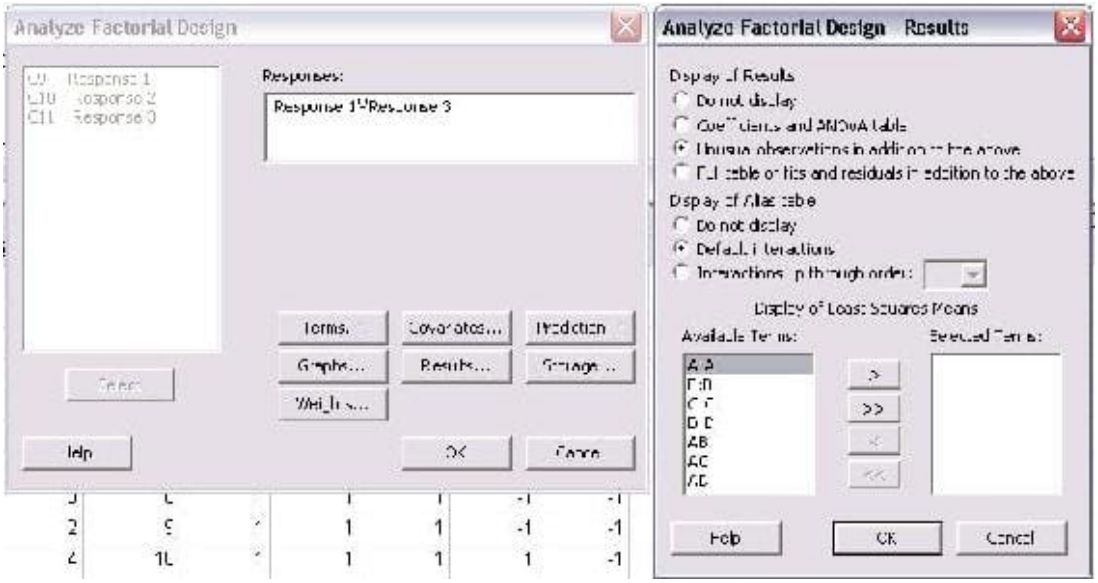
The types of interactions between factors are chosen in this menu. For a first order model which excludes all factor-to-factor interactions, "1" should be chosen from the drop-down menu for "Include terms in the model up through order:". To include higher order terms and account for factor interactions, choose 2, 3, or 4 from the drop-down menu. Unless significant factor-to-factor interactions are expected, it is recommended to use a first order model which is a linear approximation.

Once the terms have been chosen, the next step is determining which graphs should be created. The types of graphs can be selected by clicking on "Graphs..." in the main "Analyze Factorial Design" menu.



In the Graphs menu shown above, the three effects plots for "Normal", "Half Normal", and "Pareto" were selected. These plots are different ways to present the statistical results of the analysis. Examples of these plots can be found in the Minitab Example for Centrifugal Contactor Analysis. The alpha value, which determines the limit of statistical significance, can be chosen in this menu also. Typically, the alpha value is 0.05. The last type of plots that can be chosen is residual plots. A common one to select is "Residuals versus fits" which shows how the variance between the predicted values from the model and the actual values.

The final option that must be specified is results. Click "Results..." from the "Analyze Factorial Design" menu to see the following screen.



In this menu, select all of the "Available Terms" and click the ">>" button to move them to the "Selected Terms". This will ensure that all the terms will be included in the analysis. Another feature that can be selected from this menu is to display the "Coefficients and ANOVA table" for the DOE study.

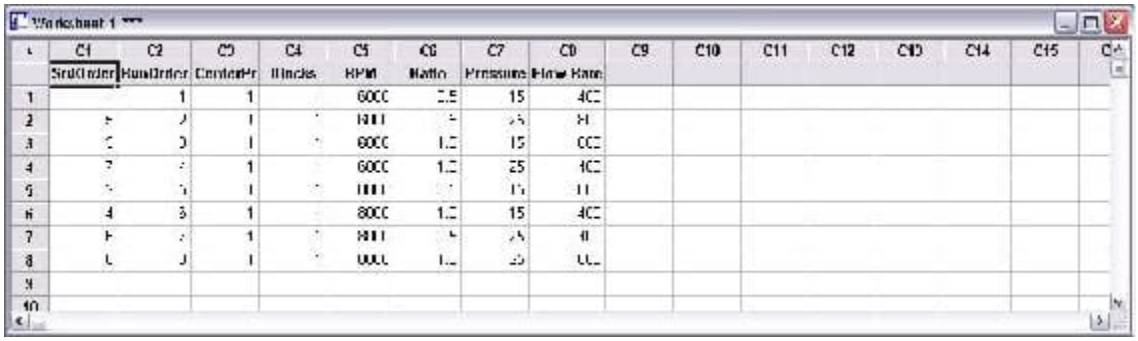
Other options can be selected from the "Analyze Factorial Design" menu such as "Covariates...", "Prediction...", "Storage...", and "Weights...". Consult the "Help" menu for descriptions of the other options. Once all desired changes have been made, click "OK" to perform the analysis. All of the plots will pop-up on the screen and a text file of the results will be generated in the session file.

Minitab Example for Centrifugal Contactor Analysis

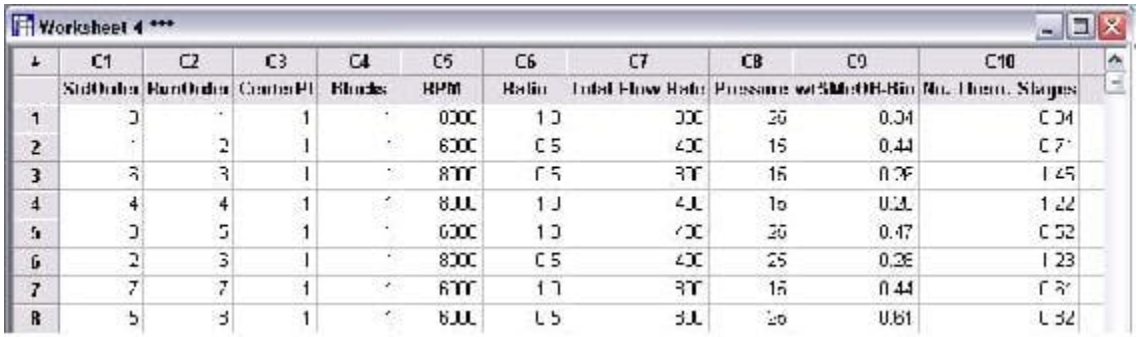
Centrifugal Contactors, also known as Podbielniak (POD) centrifugal contactors, are used to purify a contaminated stream by counter-current, liquid-liquid extraction. Two immiscible fluids with different specific gravities are contacted counter-currently

and the solute from the dirty stream is extracted by the clean stream. A common use for PODs methanol removal from biodiesel by contacting the stream with water. The amount of methanol remaining in the biodiesel (wt% MeOH) after the purification and the number of theoretical stages (No. Theor. Stages) obtained depend on the operating conditions of the POD. The four main operating parameters of the POD are rotational speed (RPM), ratio of biodiesel to water (Ratio), total flow rate of biodiesel and water (Flow Rate), and pressure (Pressure). A DOE study has been performed to determine the effect of the four operating conditions on the responses of wt% MeOH in biodiesel and number of theoretical stages achieved. (NOTE: The actual data for this example was made-up)

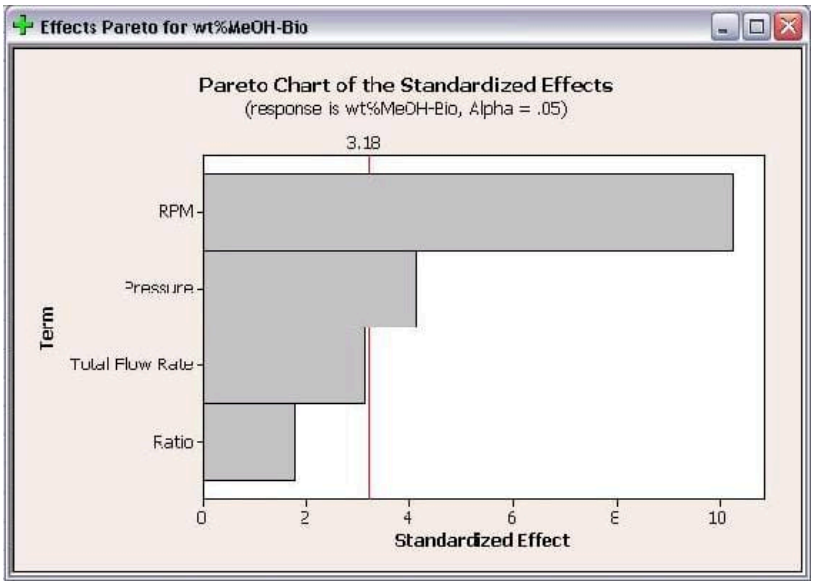
A 4-factor, 2-level DOE study was created using Minitab. Because experiments from the POD are time consuming, a half fraction design of 8 trial was used. The figure below contains the table of trials for the DOE.

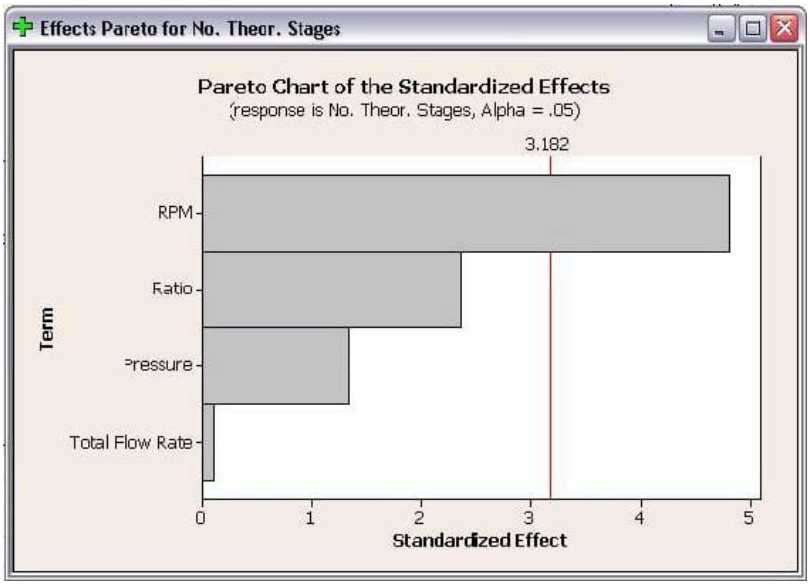


After all the trials were performed, the wt% methanol remaining in the biodiesel and number of theoretical stages achieved were calculated. The figure below contains the DOE table of trials including the two responses.



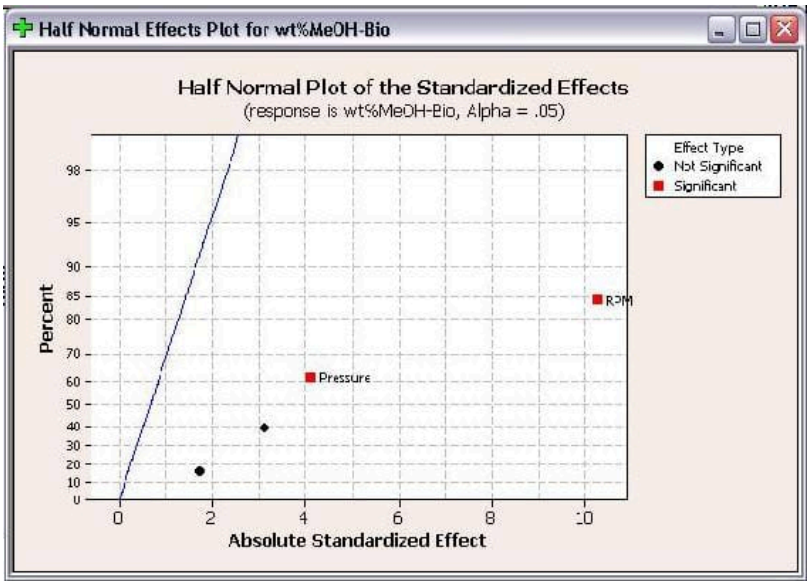
Analysis was performed on the DOE study to determine the effects of each factor on the responses. Only first order terms were included in the analysis to create a linear model. Pareto charts for both wt% MeOH in biodiesel and number of theoretical stages are shown below.

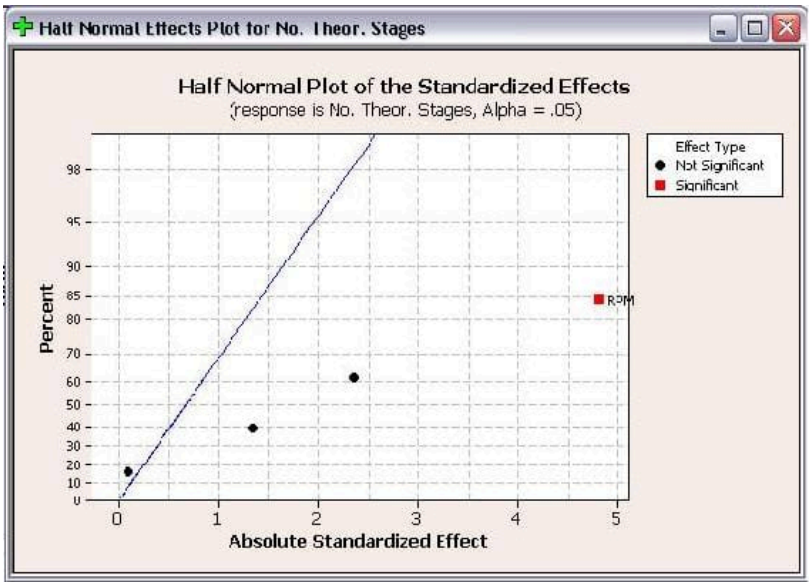




The Pareto charts show which factors have statistically significant effects on the responses. As seen in the above plots, RPM has significant effects for both responses and pressure has a statistically significant effect on wt% methanol in biodiesel. Neither flow rate or ratio have statistically significant effects on either response. The Pareto charts are bar charts which allow users to easily see which factors have significant effects.

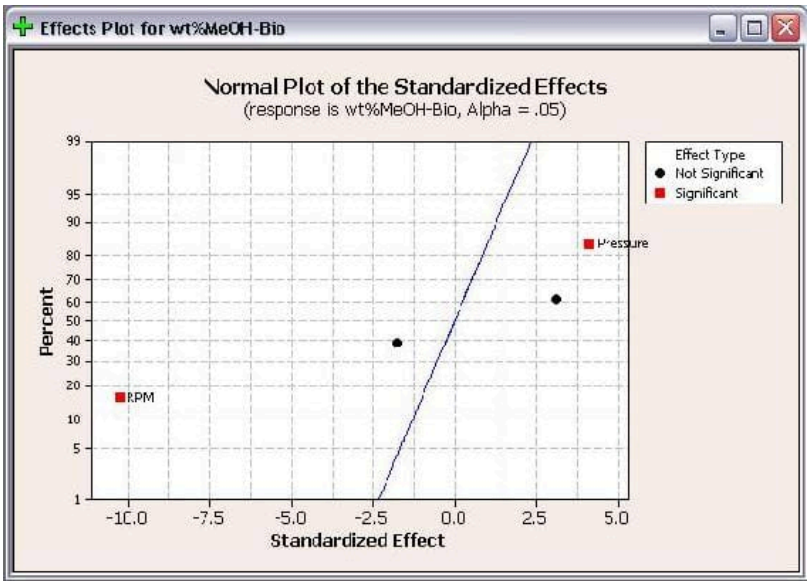
Half Normal Plots for wt% methanol in biodiesel and number of theoretical stages are shown below.

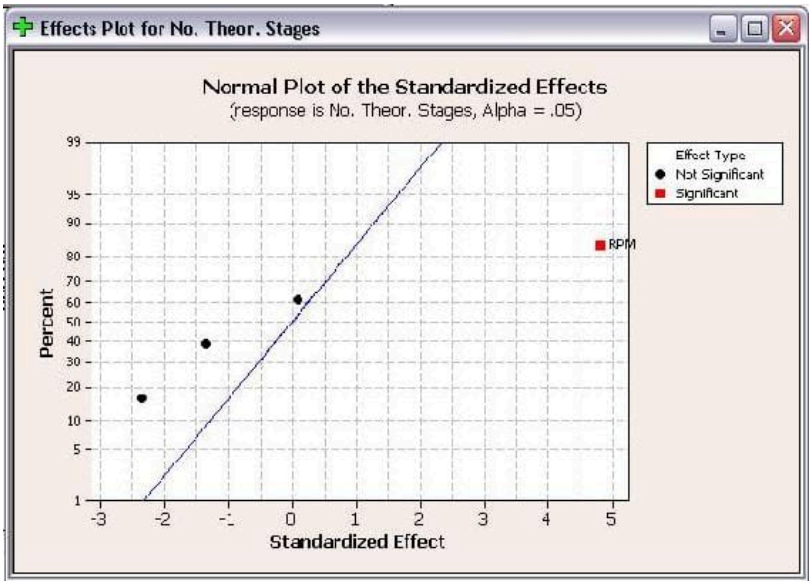




Like Pareto plots, Half Normal plots show which factors have significant effects on the responses. The factors that have significant effects are shown in red and the ones without significant effects are shown in black. The further a factor is from the blue line, the more significant effect it has on the corresponding response. For wt% methanol in biodiesel, RPM is further from the blue line than pressure, which indicates that RPM has a more significant effect on wt% methanol in biodiesel than pressure does.

The final plot created is the Normal Effect Plot. The Normal Plot is similar to the Half Normal plot in design. However, the Normal Plot displays whether the effect of the factor is positive or negative on the response. The Normal Plots for the responses are shown below.





As seen above, RPM is shown with a positive effect for number of theoretical stages, but a negative effect for wt% methanol in biodiesel. A positive effect means that as RPM increases, the number of theoretical stages increases. Whereas a negative effect indicates that as RPM increases, the wt% methanol in biodiesel decreases. Fortunately for operation with the POD, these are desired results. When choosing operating conditions for the POD, RPM should be maximized to minimize the residual methanol in biodiesel and maximize the number of theoretical stages achieved.

In addition to the above effects plots, Minitab calculates the coefficients and constants for response equations. The response equations can be used as models for predicting responses at different operating conditions (factors). The coefficients and constants for wt% methanol in biodiesel and number of theoretical stages are shown below.

Estimated Coefficients for wt%MeOH-Bio using data in uncoded units

Term	Coef
Constant	0.926833
RPM	-1.09592E-04
Ratio	-0.0748768
Total Flow Rate	0.000165648
Pressure	0.00879626

Estimated Coefficients for No. Theor. Stages using data in uncoded units

Term	Coef
Constant	-0.239346
RPM	0.000260115
Ratio	-0.509538
Total Flow Rate	0.000025577
Pressure	-0.0144769

Since this is a first order, linear model, the coefficients can be combined with the operating parameters to determine equations. The equations from this model are shown below.

$$wt\%MeOH = 0.927 - 1.096 \cdot 10^{-4}(RPM) - 0.075(Ratio) + 1.656 \cdot 10^{-4}(FlowRate) + 8.796 \cdot 10^{-3}(Pressure)$$

$$NoStages = -0.229 + 2.601 \cdot 10^{-4}(RPM) - 0.510(Ratio) + 2.558 \cdot 10^{-5}(FlowRate) - 0.0144(Pressure)$$

These equations can be used as a predictive model to determine wt% methanol in biodiesel and number of theoretical stages achieved at different operating conditions without actually performing the experiments. However, the limits of the model should be tested before the model is used to predict responses at many different operating conditions.

Example 14.2.1

You have been employed by SuperGym, a local personal training gym, who want an engineer's perspective on how to offer the best plans to their clients. SuperGym currently categorizes her clients into 4 body types to help plan for the best possible program.

- Type 1 - Very healthy
- Type 2 - Needs tone
- Type 3 - Needs strength
- Type 4 - Needs tone and strength

In addition, SuperGym offers 4 different workout plans, A through D, none of which are directly catered to any of the different types. Create an experimental factorial design that could be used to test the effects of the different workout plans on the different types of people at the gym.

Solution

In order to solve this problem, we need to determine how many different experiments would need to be performed. In order to solve this, we can see that we have two different factors, body type and workout plan. For each factor, there exist four different levels. Thus, we have a 4^2 factorial design, which gives us 16 different experimental groups. Creating a table of all of the different groups, we arrive at the following factorial design:

Solution

A1	B1	C1	D1
A2	B2	C2	D2

A3	B3	C3	D3
A4	B4	C4	D4

Where A-D is the workout plan and 1-4 is the types

Example 14.2.2

Suppose that you are looking to study the effects of hours slept (A), hours spent with significant other (B), and hours spent studying (C) on a students exam scores. You are given the following table that relates the combination of these factors and the students scores over the course of a semester. Use the Yates method in order to determine the effect each variable on the students performance in the course.

Given Information

1	17	24	19	21	22	28	25	24
2	18.5	21	20	19	26	22	27	19
3	16.5	22.5	22	25	24	26	21	20
Total	52	67.5	61	65	72	76	73	63

Solution

Using the approach introduced earlier in this article, we arrive at the following Yates solution.

			Stage		Main Total
Combination	Total	1	2	3	Factorial Effect
$a_1b_1c_1$	52	119.5	245.5	529.9	Doesn't matter
$a_2b_1c_1$	67.5	126	284	13.5	A
$a_1b_2c_1$	61	148	19.5	-5.5	B
$a_2b_2c_1$	65	136	-6	-25.5	AB
$a_1b_1c_2$	72	15.5	6.5	38.5	C
$a_2b_1c_2$	76	4	-12	-25.5	AC
$a_1b_2c_2$	73	4	-11.5	-18.5	BC
$a_2b_2c_2$	63	-10	-14	-2.5	ABC

From this table, we can see that there is positive correlation for factors A and C, meaning that more sleep and more studying leads to a better test grade in the class. Factor B, however, has a negative effect, which means that spending time with your significant other leads to a worse test score. The lesson here, therefore, is to spend more time sleeping and studying, and less time with your boyfriend or girlfriend.

Example 14.2.3

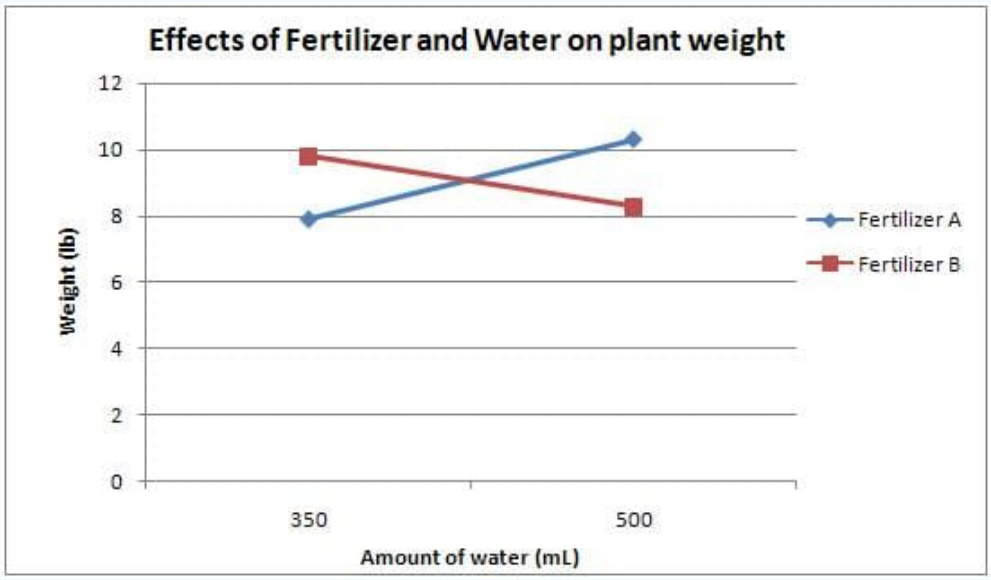
Your mom is growing a garden for the state fair and has done some experiments to find the ideal growing condition for her vegetables. She asks you for help interpreting the results and shows you the following data:

Plant No.	Fertilizer Type	Amount of water (mL)	Weight (lb)
1	A	350	7.9
2	B	350	9.8
3	A	500	10.3
4	B	500	8.3

Make plots to determine the main or interaction effects of each factor.

Solution

Here is the plot you should have gotten for the given data.



From this one can see that there is an interaction effect since the lines cross. One cannot discuss the results without speaking about both the type of fertilizer and the amount of water used. Using fertilizer A and 500 mL of water resulted in the largest plant, while fertilizer A and 350 mL gave the smallest plant. Fertilizer B and 350 mL gave the second largest plant, and fertilizer B and 500 mL gave the second smallest plant. There is clearly an interaction due to the amount of water used and the fertilizer present. Perhaps each fertilizer is most effective with a certain amount of water. In any case, your mom has to consider both the fertilizer type and amount of water provided to the plants when determining the proper growing conditions.

Exercise 14.2.1

Which of the following is not an advantage of the use of factorial design over one factor design?

a. More time efficient
b. Provides how each factor effects the response
c. Does not require explicit testing
d. Does not require regression

Answer available at the end of this section.

Exercise 14.2.2

In a 2^2 factorial design experiment, a total main effect value of -5 is obtained. This means that

a. there is a relative positive correlation between the two factors
b. there is no correlation between the two factors
c. there is a relative negative correlation between the two factors
d. there is either a positive or negative relative correlation between the two factors

Answer available at the end of this section.

References

- Box, George E.P., et. al. "Statistics for Engineers: An Introduction to Design, Data Analysis, and Model Building." New York: John Wiley & Sons.
- Trochim, William M.K. 2006. "Factorial Designs." Research Methods Knowledge Base. <http://www.socialresearchmethods.net/kb/expfact.htm>
- Perez, Jose A., et. al. "Effect of process variables on liquid hot water pretreatment of wheat straw for bioconversion to fuel-ethanol in a batch reactor." Journal of Chemical Technology & Biotechnology. Volume 82, Issue 10, Pages 929-938. Published Online Sep 3, 2007.

Answers

Exercise 14.2.1
TBA

Exercise 14.2.2
TBA

This page titled 14.2: Design of experiments via factorial designs is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Andrew Wang, Diane Feldkamp, Heeral Sheth, Jason Bourgeois, Jocelyn Anleitner, Michael Kravchenko, Nicholas Parsons, & Stephanie Combs via source content that was edited to the style and standards of the LibreTexts platform.

14.3: Design of Experiments via Random Design

Introduction

Random design is an approach to designing experiments. As the name implies, random experimental design involves randomly assigning experimental conditions. However, numbers should not be picked without any thought. This type of experimental design is surprisingly powerful and often results in a high probability to create a near optimal design.

The simplified steps for random design include the following:

1. Choose a number of experiments to run (NOTE: This may be tricky to pick a number because it is dependent upon the amount of signal recovery you want.)
2. Assign to each variable a state based on a uniform sample. For instance, if there are 5 states, each state has a probability of 20%.

Random designs typically work well for large systems with many variables, 50 or more. There should be few interactions between variables and very few variables that contribute significantly. Random design does not work very well with relatively smaller systems. Generally speaking, Taguchi and random designs often perform better than factorial designs depending on size and assumptions. When choosing the design for an experiment, it is important to determine an efficient design that helps optimize the process and determines factors that influence variability.

There is more than one type of random design, randomized block design and completely randomized design. Randomized block design involves blocking, which is arranging experimental units into groups so they have a common similarity. The blocking factor is usually not a primary source of variability. An example of a blocking factor may include eye color of a patient, so if this variability source is controlled, greater precision is achieved. Completely randomized design is where the groups are chosen at random.

In various technological fields, it is important to design experiments where a limited number of experiments is required. Random design is practical for many design applications. Extensive mathematical theory has been used to explore random experimental design. Examples of random design include areas of data compression and medical imaging. The research conducted to support the practical application of random design can be found at <http://groups.csail.mit.edu/drl/journal_club/papers/CS2-Candes-Romberg-05.pdf>.

Other research has been conducted recently on random design, and more information can be found at: ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1614066

Completely Randomized Design (CRD)

Description of Design

Completely randomized design (CRD) is the simplest type of design to use. The most important requirement for use of this design is homogeneity of experimental units.

Procedure for Randomization

1. Assign treatments to experimental units completely at random.
2. Verify that every experimental unit has the same probability of receiving any treatment.
3. Perform randomization by using a random number table, computer, program, etc.

Example of CRD

If you have 4 treatments (I, II, III, IV) and 5 replicates, how many experimental units do you have?

{I} {IV} {III} {II} {II} {III} {III} {II} {I} {III} {I} {IV} {III} {IV} {I} {IV} {II} {I} {II} {IV} **=20 randomized experimental units**

Randomized Block Design (RBD)

Description of Design

Randomized block design (RBD) takes advantage of grouping similar experimental units into blocks or replicates. One requirement of RBd is that the blocks of experimental units be as uniform as possible. The reason for grouping experimental units is so that the observed differences between treatments will be largely due to “true” differences between treatments and not random occurrences or chance.

Procedure for Randomization

1. Randomize each replicate separately.
2. Verify that each treatment has the same probability of being assigned to a given experimental unit within a replicate.
3. Have each treatment appear at least once per replicate.

Advantages of RBD

1. Generally more precise than the CRD.
2. Some treatments may be replicated more times than others.
3. Missing plots are easily estimated.
4. Whole treatments or entire replicates may be deleted from the analysis.

This page titled 14.3: Design of Experiments via Random Design is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.

14.4: Summary- Summary on Control Architectures' philosophies, advantages, and disadvantages.

Summary on Control Architectures

A summary of the philosophies, advantages, and disadvantages of the different control architectures is shown below in Table 1.

Table 1. Control architectures' philosophy, advantages, and disadvantages.

	Philosophy	Advantages	Disadvantages
Feedback	Control adjusts errors as they occur	• simple to design • no process model needed	• connects for error after they have already occurred • usually only takes input from one sensor
Feed Forward	Control connects errors they occur	• corrects potential errors before they occur • Ideally can produce a perfect control where there is never any offset	• infinitely accurate models are needed • infinitely accurate measurements are needed
Ratio	Controls connects two flow to maintain a constant ratio	• Generate defined ratio • Simple Model	• The flow you control will never truly measured and ratio model is a assumed to be accurate • does not take into consideration of varying pressure
Cascade	Controls set points of sensors by controlling other sensors to integrate multiple information	• Responds quickly to high frequency changes • Incorporate multiple sensor readings together (incidence graphs)	• can oe very complex model • tuning model is difficult due to multiple set points and parameters
Mixed	Control integrating combinations of control architectures	• pick the right combination of controls to accomplish main control in zo objective • incorporate any number of sensors in a logical manner	• even more complex model than cascade alone • tuning model is even more complex than cascade alone, Zeigler Nichols tuning won't work by itself

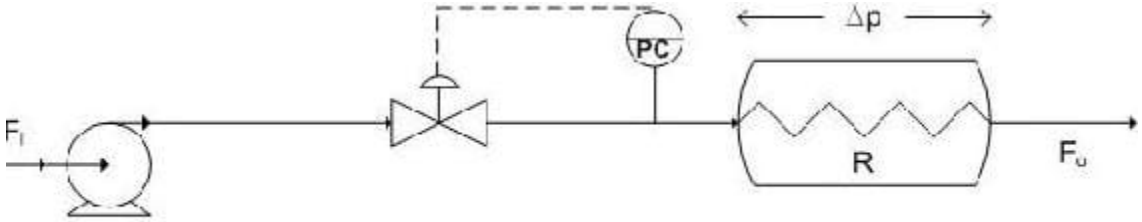
Introduction

Liquid level and liquid pressure control are fundamental aspects of many control processes because of the reliance on the stability and convenience of liquid usage. The following article will introduce the main concepts behind liquid pressure control, liquid level control, examining the loop characteristics, response, tuning and limitations. Two models are included to illustrate the application of a liquid pressure control loop and a level control loop in idealized systems.

Pressure Control Basics

Like its counterparts of temperature, level and flow, pressure is one of the most common process variables. Pressure is a key process variable as it provides a critical condition for processes such as any chemical reaction, extrusion, boiling, distillation, air conditioning and vacuuming. Poor pressure control may likely lead to problems associated with quality, safety and productivity. For example, the presence of high-pressure conditions in a sealed vessel may result in an explosion and thus it is imperative that pressure be kept under good control within its safety limits.

Liquid pressure control is one of the easier control loops in the sense that it is shares many of the same characteristics of a common flow loop. The aim of the liquid pressure loop is to control the pressure at the needed set point pressure by controlling the flow with changing process needs. Because dealing with a liquid, the flow present is that of an incompressible fluid and as a result the pressure changes very quickly leading to a fast-responding process with small dead time and capacitance. This process behaves as a fixed restriction where the change in pressure is a function of the flow through the process. Illustrated below is a diagram with liquid pressure control on the pipe leading to the process.



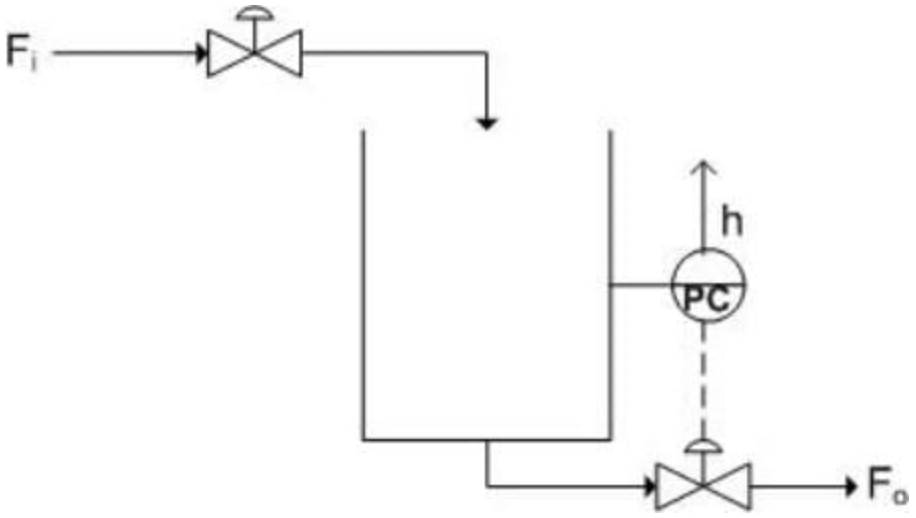
There are a few additional considerations for the aforementioned liquid pressure loop. Firstly the controller can be proportional plus integral (PI) or integral only (I-only) if $K_c < 2$ otherwise a proportional only (P-only) controller is to be used. This controller is tuned similarly to the flow controller. Another consideration is that process gain is not constant, a square-root extractor or the highest loop gain should be used in tuning the controller. This highest loop gain is employed to prevent the process loop from ever becoming unstable. Finally, this liquid pressure loop is noisy like its flow loop counterpart and as a result it is recommended that derivative action in the controller not be used.

If you wish to see the various issues associated with pressure controllers please refer to previous wiki articles.

Level Control Basics

Controlling the level of liquid in a tank is one of the most basic requirements for a chemical process to run effectively. Most chemical engineering processes require some sort of liquid to be held for usage, and the most convenient way of ensuring the proper amount of liquid is available is through level control. A specific example of a process that requires level control is seen around the world. Water towers must have a designated amount of water within them to ensure sufficient water pressure to surrounding neighborhoods and hence liquid level control (LLC) is used.

Fortunately Liquid level control is one of the easier control loops available. The systems are fairly simple, a sensor reads the level of liquid within a tank, which then is conveyed to the controller. Subsequently, the controller changes a valve to increase or decrease the flowrate into or out of the tank, depending on the required action and controlled valve placement. Below is a figure with LLC on the pipe leaving the tank.



The ease of LLC also lies in its large capacitance and nominal dead time. Usually hold-up times vary from 5-15 minutes. Problems may arise when signal noise becomes a factor, as it usually does with level controllers. Refer to previous wiki articles to the issues associated with level controllers(), however it is important to note that a measurement of weight instead of directly measuring level can be a effective way of eliminating noise.

P-only Controllers

Oftentimes using a P-only controller is the best way to ensure proper level control, in most cases only a small percent error will result, and it reduces problems associated with noise. P-only controllers should only be used when gain is very small and the tank has a large capacity. This will be the basis for the rest of the text.

A P-only controller works off the following principle of control:

$$\text{Output} = \text{Gain} \times \text{Error} + \text{Bias}$$

The output directly effects a valve to control the flowrate into or out of the tank. Refer to the previous wiki articles to learn more about P-only control.

Level Measurement Noise

When controlling liquid levels in drums or vessels, an important aspect to consider is that there may be noise in the level measurement due to disturbances such surface turbulence, boiling of the liquid, or agitation. Due to the existence of this noise, using a derivative action controller may not be appropriate. However, if this type of controller action is used, there are various methods that may be employed to minimize the noise.

Method 1: Use a displacer in a stilling well

- *Advantage:* Filters high frequency noise due to turbulence in the tank
- *Disadvantage:* Bobbing level due to the low-frequency movement of tank liquid into the well resulting from the formation of a U-tube between tank and well

Method 2: Use an ultrasonic level measurement with electronic filtering of the signal

- *Advantage:* Works well when the level response period is much lower than the noise frequency

Method 3: Use a tank weighing method

In the this method, a loading device is placed beneath the tank supports to measure the mass of the tank. Output weight values are averaged by a transmitter which are then sent to a converter which outputs the corresponding level to the controller.

- *Advantage:* Effectively eliminates noise in the level measurement due to turbulence

Models

The following are Excel models of Pressure and Level control in simple systems. Proportional only control is used primarily for these models. Controller noise has been eliminated for sake of simplicity.

Liquid Pressure Control Model

The model at hand illustrates both the feed backwards system of the liquid pressure in a pipline of a process and its process gain relationships. This liquid pressure control loop is regulated by the position of the valve on the pipeline leading to the process. A P-only control is used here to change the valve position resulting in the need response in attaining the set point value. The following are any assumptions made and the equations used in modeling this process. An assumption is made that the process behaves like a fixed restriction such as an orifice plate whose Dp is a function of flow through the process. Another assumption that should be made is that the valve responds linearly to the flow-rate through it. In determining the process gain the following equations where used.

$$P = \Delta p + P_0$$

where P_0 is the downstream pressure at zero flow

$$\Delta p = \frac{F^2}{R^2}$$

where R is the process flow resistance. As a result we obtain:

$$P = \frac{F^2}{R^2} + P_0$$

The gain of the process is then determined by the following expression:

$$K_p = \frac{dP}{dF} = \frac{2F}{R^2}$$

A model for the feed backwards system for liquid pressure control introduced in this article is very complicated and might possibly be out of the scope of our discussion. A possible alternative model would be very similar to the liquid level model presented below with the head in the tank determining the liquid pressure in the pipeline leaving the reactor.

Liquid Level Control Model

The following model represents a feed backward system of a liquid holding tank. The control loop is regulated by the exiting valve position on the pipe leaving the tank. A maximal flowrate out is assumed and the height of water in the tank does not have any effect on the flowrate out of the tank. This assumption can be made in tanks which are sufficiently wider than the pipe exiting operating with little liquid level change, as often is the case. Another assumption made is the linear response of the valve to flowrate out.

The pipe entering the tank is set by the user at a fixed flowrate. The P-only control is used to change the exiting valve position in order to cause the desired response to obtain the setpoint value. The P control follows the equation listed above and the bias is designated by the point at which the valve is open to allow a steady state to occur. Gain in Level control systems is usually low, however this can be changed by the user as well. The setpoint for the tank is determined by the user, as are the entering liquid flowrates and the maximal exiting flowrate.

The user can input a set point change in order to view the corresponding response in the controller.

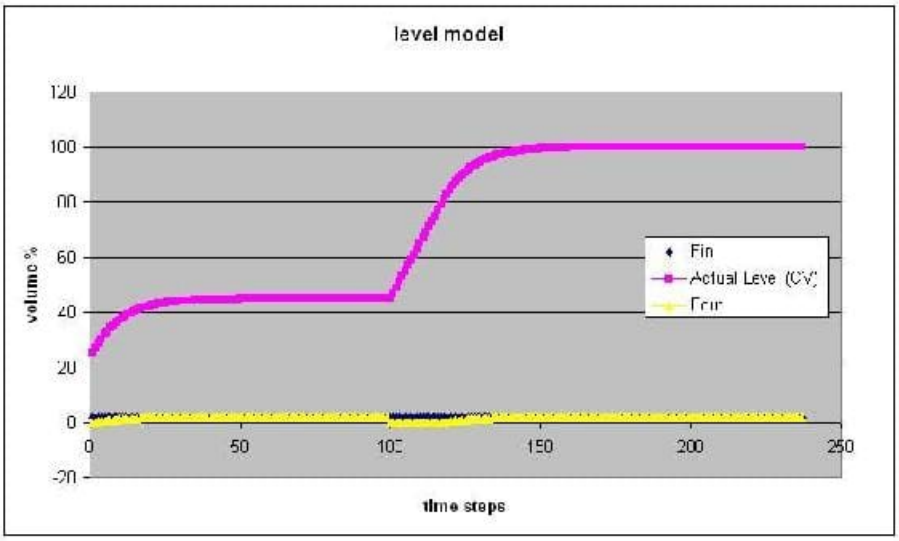
Example 14.4.1

Using the Liquid Level control Excel Model, what do you find when you change the level set point at a given point in time?

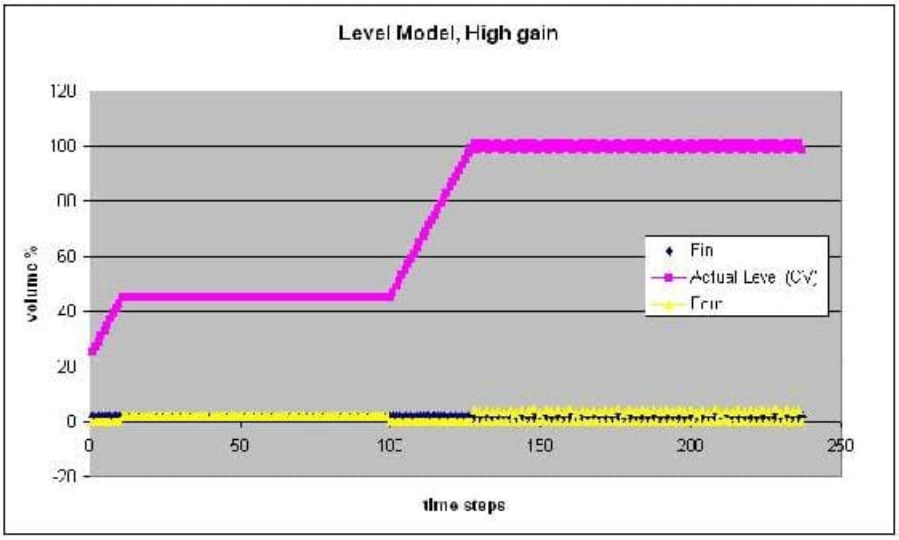
What happens if you change the gain to an large number?

Solution

When you change the level set point, you should notice a corresponding increase or decrease in the level as the P-control kicks in and beings to compensate for the error. The rate at which this set point is obtained is determined by the controller gain. At high controller gains the system responds extremely rapidly in an unreal fashion, this is why controller gain is usually small in large capacity tanks with flow control.



The above figure changes the setpoint to 100 @ t=100



The above figure uses a very high controller gain. Note the unreal response.

Example 14.4.2

What is the primary difference between the model for the process gain of the pressure flow loop introduced earlier in the article and a typical head flow response curve?

Solution

The response of pressure to flow is shaped similarly to the response of the head flow device and as a result the rules that govern the liquid pressure process gain model are in essence the same for the head flow device response. The only difference between the two response curves is that whereas the pressure varies from Po to 100 per cent for the process gain curve for the pressure flow loop, the pressure varies from 0 to 100 percent for the head flow device.

Below is the response for the pressure loop followed by that of head flow device.

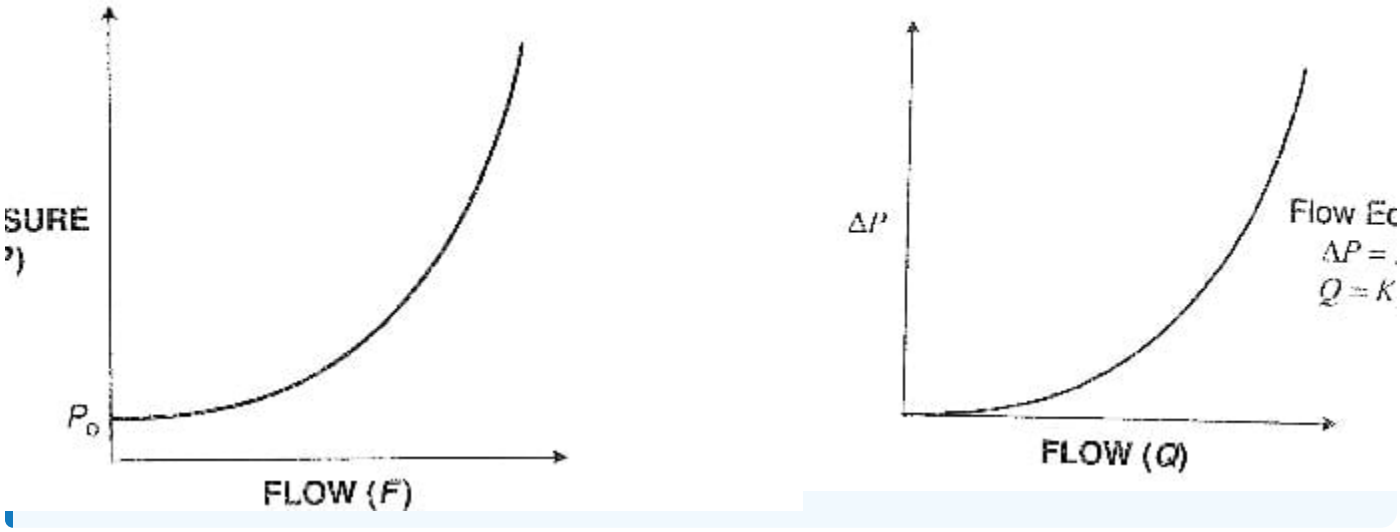


Figure 14.4.1: esponse for the pressure loop followed by that of head flow device.

Multiple Choice Question 1

When should P-only controllers be used in Liquid Level control?

a. Whenever, because it's really easy.
b. Never because it is too simple.
c. Sometimes, when the gain is small and the tank capacity is large.
d. None of the above

Multiple Choice Question 2

Various methods may be used to minimize the noise in a level measurement when using which type of controller action:

a. None, you can never minimize noise in a level measurement
b. Integral
c. Derivative
d. PI-control

This page titled 14.4: Summary- Summary on Control Architectures' philosophies, advantages, and disadvantages. is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by George Cater, Jennifersue Bowker, & Kibwe John Date via source content that was edited to the style and standards of the LibreTexts platform.

INDEX

A

acid
3.7: pH and Viscosity Sensors

activity coefficient
3.7: pH and Viscosity Sensors

Analysis of variance (ANOVA)
13.12: Factor analysis and ANOVA

attractors
10.5: Phase Plane Analysis - Attractors, Spirals, and Limit cycles

autonumheader:yes
5.1: Boolean Models - Truth Tables and State Transition Diagrams

autonumheader:yes2

1: Overview
1.1: Introduction to Controls- Background and design methodology
1.2: Introduction to DCS- Control system hardware
1.3: Current Significance- Process controls and you
1.4: Failures in Process Control- Bhopal, Three Mile Island
1.5: Process Controls in Everyday Life- Applying process control thinking to everyday situations
10: Dynamical Systems Analysis
10.1: Finding fixed points in ODEs and Boolean models
10.2: Linearizing ODEs
10.3: Eigenvalues and Eigenvectors
10.4: Using Eigenvalues and Eigenvectors to Find Stability and Solve ODEs
10.5: Phase Plane Analysis - Attractors, Spirals, and Limit cycles
10.6: Root Locus Plots - Effect of Tuning
10.7: Routh Stability - Ranges of Parameter Values that are Stable
11: Control Architectures
11.1: Feedback Control
11.2: Feed Forward Control
11.3: Cascade Control
11.4: Ratio Control
11.5: Common control loops and model for liquid pressure and liquid level
11.6: Common Control Loops and Model for Temperature Control
11.7: Common control architectures and model for reactors
12: Multiple Input, Multiple Output (MIMO) Control
12.1: Determining if a system can be decoupled
12.2: MIMO control using RGA
12.3: MIMO using Model Predictive Control
12.4: Neural Networks for automatic model construction
12.5: Understanding MIMO Control Through Two Tanks Interaction
13: Statistics and Probability Background
13.1: Basic statistics- mean, median, average, standard deviation, z-scores, and p-value
13.10: Multinomial Distributions
13.11: Comparisons of two means
13.12: Factor analysis and ANOVA
13.13: Correlation and Mutual Information
13.14: Random sampling from a stationary Gaussian process
13.2: SPC- Basic Control Charts- Theory and Construction, Sample Size, X-Bar, R charts, S charts
13.3: Six Sigma- What is it and what does it mean?
13.4: Bayes Rule, Conditional Probability and Independence
13.5: Bayesian Network Theory
13.6: Learning and analyzing Bayesian networks with Genie
13.7: Occasionally Dishonest Casino? Markov Chains and Hidden Markov Models
13.8: Continuous Distributions- normal and exponential
13.9: Discrete Distributions - Hypergeometric, Binomial, and Poisson
14: Design of Experiments
14.1: Design of Experiments via Taguchi Methods - Orthogonal Arrays
14.2: Design of experiments via factorial designs
14.3: Design of Experiments via Random Design
14.4: Summary- Summary on Control Architectures' philosophies, advantages, and disadvantages.
2: Modeling Basics
2.1: Verbal Modeling- process description, control specifications, and connections
2.2: Degrees of Freedom- importance, calculation procedure, and examples
2.3: Incidence Graphs -Interpretations, Consistency, and Inconsistency
2.4: Excel Modeling - logical models, optimization with solver for nonlinear regression, sampling random numbers
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles
2.6: Numerical ODE solving in Excel- Euler's method, Runge Kutta, Dead time in ODE solving
2.7: Solving ODEs with Mathematica- How to find numerical and analytical solutions to ODEs with Mathematica
2.8: Fitting ODE parameters to data using Excel- Using regression to fit complex models in Excel
2.9: Helpful Mathematica Syntax- Hints on how to use Mathematica to model chemical processes
3: Sensors and Actuators
3.1: Control Systems - Measurement Devices
3.10: Valves - Modeling Dynamics
3.2: Temperature Sensors
3.3: Pressure Sensors
3.4: Level Sensors
3.5: Flow Sensors
3.6: Composition Sensors
3.7: pH and Viscosity Sensors
3.8: Miscellaneous Sensors
3.9: Valves - Types, Kinds, and Selection
4: Piping and Instrumentation Diagrams
4.1: PandID General Information
4.2: Piping and Instrumentation Diagram Standard Notation
4.3: Piping and Instrumentation Diagrams - Location of Controls and Standard Control Structures
4.4: Piping and Instrumentation Diagrams - Standard Pitfalls
4.5: Safety Features in Piping and Instrumentation Diagrams
4.6: Regulatory Agencies and Compliance
5: Logical Modeling
5.2: Logical Control Programs - IF... THEN... WHILE...
6: Modeling Case Studies
6.1: Surge Tank Model
6.2: Heated Surge Tank
6.3: Bacterial Chemostat
6.4: ODE and Excel CSTR model with heat exchange
6.5: ODE and Excel model of a Simple Distillation Column
6.6: ODE and Excel model of a Heat Exchanger
6.7: ODE and Excel Model of an Adiabatic PFR
6.8: Modeling and PID Controller Example - Cruise Control for an Electric Vehicle
6.9: Blood Glucose Control in Diabetic Patients
7: Mathematics for Control Systems
7.1: Dirac delta (impulse) function
7.2: First-order Differential Equations
7.3: Second-order Differential Equations
7.4: Taylor Series
7.5: Laplace Transforms
8: Optimization
8.1: Introduction to Optimization
8.2: Linear Optimization
8.3: Non-linear Optimization
9: Proportional-Integral-Derivative (PID) Control
9.1: Constructing Block Diagrams- Visualizing control measurements
9.2: P, I, D, PI, PD, and PID control
9.3: PID Tuning via Classical Methods
9.4: PID tuning via Frequency Responses with Bode Plots
9.5: PID tuning via Optimization
9.6: PID Downsides and Solutions
Chemical Process Dynamics and Controls (Woolf)

B

ball valves
3.9: Valves - Types, Kinds, and Selection
base
3.7: pH and Viscosity Sensors
Bayes factor
13.5: Bayesian Network Theory
Bayes rule
13.4: Bayes Rule, Conditional Probability and Independence
Bayes' Theorem
13.4: Bayes Rule, Conditional Probability and Independence
Bayesian Information Criterion (BIC)
13.5: Bayesian Network Theory
Bayesian network theory
13.5: Bayesian Network Theory
Bayesian networks
13.6: Learning and analyzing Bayesian networks with Genie
belief network
13.5: Bayesian Network Theory
Bergman minimal model
6.9: Blood Glucose Control in Diabetic Patients
binomial distribution
13.9: Discrete Distributions - Hypergeometric, Binomial, and Poisson
blend stations
11.4: Ratio Control
blue noise
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles
Bode plot
9.4: PID tuning via Frequency Responses with Bode Plots
boolean
5.1: Boolean Models - Truth Tables and State Transition Diagrams
Boolean modeling
10.1: Finding fixed points in ODEs and Boolean models
Boolean network
5.1: Boolean Models - Truth Tables and State Transition Diagrams
Bourdon tubes
3.3: Pressure Sensors
brown noise
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles
butterfly valves
3.9: Valves - Types, Kinds, and Selection

C

completely randomized design (CRD)
14.3: Design of Experiments via Random Design
conditional probability
13.4: Bayes Rule, Conditional Probability and Independence
control chart
13.2: SPC- Basic Control Charts- Theory and Construction, Sample Size, X-Bar, R charts, S charts
correlation
13.13: Correlation and Mutual Information

correlation coefficient
13.13: Correlation and Mutual Information
crackling (noise)
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles
CSTR
9.4: PID tuning via Frequency Responses with Bode Plots
cumulative distribution function (CDF)
13.10: Multinomial Distributions

D

dead time
9.2: P, I, D, PI, PD, and PID control
degrees of freedom analysis
4.3: Piping and Instrumentation Diagrams - Location of Controls and Standard Control Structures
determinant
10.3: Eigenvalues and Eigenvectors
Dirac delta function
7.1: Dirac delta (impulse) function
directed acyclic graph
13.5: Bayesian Network Theory
distillation column
6.5: ODE and Excel model of a Simple Distillation Column
doppler meters
3.5: Flow Sensors
drift noise
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles
Dynamic Bayesian Network (DBN)
13.5: Bayesian Network Theory

E

eigenvalues
10.3: Eigenvalues and Eigenvectors
eigenvectors
10.3: Eigenvalues and Eigenvectors
endothermic reactors
11.7: Common control architectures and model for reactors
exothermic reactors
11.7: Common control architectures and model for reactors

F

factorial design
14.2: Design of experiments via factorial designs
Fisher's exact
13.9: Discrete Distributions - Hypergeometric, Binomial, and Poisson
fixed points
10.1: Finding fixed points in ODEs and Boolean models
flow
3.5: Flow Sensors
flow sensors
3.5: Flow Sensors
frequency based noise
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles

G

Gaussian noise
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles

H

heat exchanger
6.6: ODE and Excel model of a Heat Exchanger
hidden Markov models
13.7: Occasionally Dishonest Casino? Markov Chains and Hidden Markov Models
humidity sensors
3.8: Miscellaneous Sensors
hypergeometric distribution
13.9: Discrete Distributions - Hypergeometric, Binomial, and Poisson

I

incidence graphs
2.3: Incidence Graphs -Interpretations, Consistency, and Inconsistency
independence
13.4: Bayes Rule, Conditional Probability and Independence
inherent flow characteristic
3.10: Valves - Modeling Dynamics

J

joint probability
13.4: Bayes Rule, Conditional Probability and Independence

L

level sensors
3.4: Level Sensors
limit cycle
10.5: Phase Plane Analysis - Attractors, Spirals, and Limit cycles
linear optimization
8.2: Linear Optimization
linearizing nonlinear ordinary differential equations
10.2: Linearizing ODEs
logical equivalence
5.1: Boolean Models - Truth Tables and State Transition Diagrams
logical inversion
5.1: Boolean Models - Truth Tables and State Transition Diagrams

M

MIMO
12.5: Understanding MIMO Control Through Two Tanks Interaction
MIMO (multiple input, multiple output)
12.1: Determining if a system can be decoupled
12.2: MIMO control using RGA
MISO (multiple input, single output)
12.1: Determining if a system can be decoupled
Model Predictive Control
12.3: MIMO using Model Predictive Control
Monod equation
6.3: Bacterial Chemostat
monotonicity
2.3: Incidence Graphs -Interpretations, Consistency, and Inconsistency
Moser equation
6.3: Bacterial Chemostat
multinomial distributions
13.10: Multinomial Distributions
Multiple Input Multiple Output
12.5: Understanding MIMO Control Through Two Tanks Interaction

N

negative feedback
11.1: Feedback Control
neural network
12.4: Neural Networks for automatic model construction
Niederlinski index
12.2: MIMO control using RGA
noise
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles
nullcline
10.5: Phase Plane Analysis - Attractors, Spirals, and Limit cycles

O

ODE modeling
10.1: Finding fixed points in ODEs and Boolean models
orifice meter
3.5: Flow Sensors
orthogonal arrays
14.1: Design of Experiments via Taguchi Methods - Orthogonal Arrays
overspecification
4.4: Piping and Instrumentation Diagrams - Standard Pitfalls

P

PageRank
13.7: Occasionally Dishonest Casino? Markov Chains and Hidden Markov Models
pH
3.7: pH and Viscosity Sensors
photometric analysis
3.6: Composition Sensors
PID
9.6: PID Downsides and Solutions
PID controller
9.2: P, I, D, PI, PD, and PID control
pink noise
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles
Piping and Instrumentation Diagrams (P&ID)
4.2: Piping and Instrumentation Diagram Standard Notation
4.3: Piping and Instrumentation Diagrams - Location of Controls and Standard Control Structures
4.4: Piping and Instrumentation Diagrams - Standard Pitfalls
plinko
10.3: Eigenvalues and Eigenvectors

plug flow reactors
11.7: Common control architectures and model for reactors
6.7: ODE and Excel Model of an Adiabatic PFR
Poisson distribution
13.9: Discrete Distributions - Hypergeometric, Binomial, and Poisson
pops (noise)
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles
positive feedback
11.1: Feedback Control
probability density function
13.10: Multinomial Distributions
process gain
9.2: P, I, D, PI, PD, and PID control
proportional control
9.6: PID Downsides and Solutions
Proton Exchange Membrane (PEM) fuel cell
5.1: Boolean Models - Truth Tables and State Transition Diagrams
purple noise
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles

Q

queuing theory
13.7: Occasionally Dishonest Casino? Markov Chains and Hidden Markov Models

R

random design
14.3: Design of Experiments via Random Design
randomized block design (RBD)
14.3: Design of Experiments via Random Design
ratio control
11.4: Ratio Control
ratio control architecture
11.4: Ratio Control
reboiler
6.5: ODE and Excel model of a Simple Distillation Column
relative gain array (RGA)
12.2: MIMO control using RGA
resistance temperature detector (RTD)
3.2: Temperature Sensors
root locus plots
10.6: Root Locus Plots - Effect of Tuning
rotameter
3.5: Flow Sensors
Routh array
10.7: Routh Stability - Ranges of Parameter Values that are Stable
Routh test
10.7: Routh Stability - Ranges of Parameter Values that are Stable
Routh's theorem
10.4: Using Eigenvalues and Eigenvectors to Find Stability and Solve ODEs

S

separatrix
10.5: Phase Plane Analysis - Attractors, Spirals, and Limit cycles
Shewhart chart
13.3: Six Sigma- What is it and what does it mean?
shot noise
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles
SIMO (single input, multiple output)
12.1: Determining if a system can be decoupled
Single Input Single Output
12.5: Understanding MIMO Control Through Two Tanks Interaction
Single variable Input or Single variable Output (SISO)
12.2: MIMO control using RGA
SISO
12.5: Understanding MIMO Control Through Two Tanks Interaction
SISO (single input, single output)
12.1: Determining if a system can be decoupled
snap (noise)
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles
sonic nozzle
3.5: Flow Sensors
spirals
10.5: Phase Plane Analysis - Attractors, Spirals, and Limit cycles
stationary Gaussian process
13.14: Random sampling from a stationary Gaussian process
step function
7.1: Dirac delta (impulse) function
surge tank
6.1: Surge Tank Model
6.2: Heated Surge Tank

T

Taguchi loss function
14.1: Design of Experiments via Taguchi Methods - Orthogonal Arrays
Taguchi method
14.1: Design of Experiments via Taguchi Methods - Orthogonal Arrays
Taylor series
7.4: Taylor Series
temperature control loops
11.6: Common Control Loops and Model for Temperature Control
Tessier equation
6.3: Bacterial Chemostat
thermometer
3.2: Temperature Sensors
Transmitter Symbols (P&IDs)
4.2: Piping and Instrumentation Diagram Standard Notation
truth tables
5.1: Boolean Models - Truth Tables and State Transition Diagrams

U

ultimate gain
9.3: PID Tuning via Classical Methods

V

valve coefficient
3.10: Valves - Modeling Dynamics
valve rangeability
3.10: Valves - Modeling Dynamics
viscosity
3.7: pH and Viscosity Sensors

W

white noise
2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles

Y

Yates algorithm
14.2: Design of experiments via factorial designs

Z

zero buoyancy force level
3.4: Level Sensors

This page titled Index is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by LibreTexts.

GLOSSARY

S

Sample Word 1 | Sample Definition 1

This page titled Glossary is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by LibreTexts.

DETAILED LICENSING

Overview

Title: Chemical Process Dynamics and Controls (Woolf)

Webpages: 116

All licenses found:

- CC BY 3.0: 94% (109 pages)
- CC BY 4.0: 3.4% (4 pages)
- Undeclared: 2.6% (3 pages)

By Page

- Chemical Process Dynamics and Controls (Woolf) - *CC BY 3.0*
  - Front Matter - *CC BY 3.0*
    - TitlePage - *CC BY 4.0*
    - InfoPage - *CC BY 3.0*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Overview - *CC BY 3.0*
    - 1.1: Introduction to Controls- Background and design methodology - *CC BY 3.0*
    - 1.2: Introduction to DCS- Control system hardware - *CC BY 3.0*
    - 1.3: Current Significance- Process controls and you - *CC BY 3.0*
    - 1.4: Failures in Process Control- Bhopal, Three Mile Island - *CC BY 3.0*
    - 1.5: Process Controls in Everyday Life- Applying process control thinking to everyday situations - *CC BY 3.0*
  - 2: Modeling Basics - *CC BY 3.0*
    - 2.1: Verbal Modeling- process description, control specifications, and connections - *CC BY 3.0*
    - 2.2: Degrees of Freedom- importance, calculation procedure, and examples - *CC BY 3.0*
    - 2.3: Incidence Graphs -Interpretations, Consistency, and Inconsistency - *CC BY 3.0*
    - 2.4: Excel Modeling - logical models, optimization with solver for nonlinear regression, sampling random numbers - *CC BY 3.0*
    - 2.5: Noise Modeling - White, Pink, and Brown Noise, Pops and Crackles - *CC BY 3.0*
    - 2.6: Numerical ODE solving in Excel- Euler's method, Runge Kutta, Dead time in ODE solving - *CC BY 3.0*
    - 2.7: Solving ODEs with Mathematica- How to find numerical and analytical solutions to ODEs with Mathematica - *CC BY 3.0*
    - 2.8: Fitting ODE parameters to data using Excel- Using regression to fit complex models in Excel - *CC BY 3.0*
    - 2.9: Helpful Mathematica Syntax- Hints on how to use Mathematica to model chemical processes - *CC BY 3.0*
  - 3: Sensors and Actuators - *CC BY 3.0*
    - 3.1: Control Systems - Measurement Devices - *CC BY 3.0*
    - 3.2: Temperature Sensors - *CC BY 3.0*
    - 3.3: Pressure Sensors - *CC BY 3.0*
    - 3.4: Level Sensors - *CC BY 3.0*
    - 3.5: Flow Sensors - *CC BY 3.0*
    - 3.6: Composition Sensors - *CC BY 3.0*
    - 3.7: pH and Viscosity Sensors - *CC BY 3.0*
    - 3.8: Miscellaneous Sensors - *CC BY 4.0*
    - 3.9: Valves - Types, Kinds, and Selection - *CC BY 3.0*
    - 3.10: Valves - Modeling Dynamics - *CC BY 3.0*
  - 4: Piping and Instrumentation Diagrams - *CC BY 3.0*
    - 4.1: PandID General Information - *CC BY 3.0*
    - 4.2: Piping and Instrumentation Diagram Standard Notation - *CC BY 3.0*
    - 4.3: Piping and Instrumentation Diagrams - Location of Controls and Standard Control Structures - *CC BY 3.0*
    - 4.4: Piping and Instrumentation Diagrams - Standard Pitfalls - *CC BY 3.0*
    - 4.5: Safety Features in Piping and Instrumentation Diagrams - *CC BY 3.0*
    - 4.6: Regulatory Agencies and Compliance - *CC BY 3.0*
  - 5: Logical Modeling - *CC BY 3.0*
    - 5.1: Boolean Models - Truth Tables and State Transition Diagrams - *CC BY 3.0*
    - 5.2: Logical Control Programs - IF... THEN... WHILE... - *CC BY 3.0*
  - 6: Modeling Case Studies - *CC BY 3.0*

- 6.1: Surge Tank Model - *CC BY 3.0*
- 6.2: Heated Surge Tank - *CC BY 3.0*
- 6.3: Bacterial Chemostat - *CC BY 3.0*
- 6.4: ODE and Excel CSTR model with heat exchange - *CC BY 3.0*
- 6.5: ODE and Excel model of a Simple Distillation Column - *CC BY 3.0*
- 6.6: ODE and Excel model of a Heat Exchanger - *CC BY 3.0*
- 6.7: ODE and Excel Model of an Adiabatic PFR - *CC BY 3.0*
- 6.8: Modeling and PID Controller Example - Cruise Control for an Electric Vehicle - *CC BY 3.0*
- 6.9: Blood Glucose Control in Diabetic Patients - *CC BY 3.0*

◦ 7: Mathematics for Control Systems - *CC BY 3.0*
- 7.1: Dirac delta (impulse) function - *CC BY 3.0*
- 7.2: First-order Differential Equations - *CC BY 3.0*
- 7.3: Second-order Differential Equations - *CC BY 3.0*
- 7.4: Taylor Series - *CC BY 3.0*
- 7.5: Laplace Transforms - *CC BY 3.0*

◦ 8: Optimization - *CC BY 3.0*
- 8.1: Introduction to Optimization - *CC BY 3.0*
- 8.2: Linear Optimization - *CC BY 3.0*
- 8.3: Non-linear Optimization - *CC BY 3.0*

◦ 9: Proportional-Integral-Derivative (PID) Control - *CC BY 3.0*
- 9.1: Constructing Block Diagrams- Visualizing control measurements - *CC BY 3.0*
- 9.2: P, I, D, PI, PD, and PID control - *CC BY 3.0*
- 9.3: PID Tuning via Classical Methods - *CC BY 3.0*
- 9.4: PID tuning via Frequency Responses with Bode Plots - *CC BY 3.0*
- 9.5: PID tuning via Optimization - *CC BY 3.0*
- 9.6: PID Downsides and Solutions - *CC BY 3.0*

◦ 10: Dynamical Systems Analysis - *CC BY 3.0*
- 10.1: Finding fixed points in ODEs and Boolean models - *CC BY 3.0*
- 10.2: Linearizing ODEs - *CC BY 3.0*
- 10.3: Eigenvalues and Eigenvectors - *CC BY 3.0*
- 10.4: Using Eigenvalues and Eigenvectors to Find Stability and Solve ODEs - *CC BY 3.0*
- 10.5: Phase Plane Analysis - Attractors, Spirals, and Limit cycles - *CC BY 3.0*
- 10.6: Root Locus Plots - Effect of Tuning - *CC BY 3.0*
- 10.7: Routh Stability - Ranges of Parameter Values that are Stable - *CC BY 3.0*

◦ 11: Control Architectures - *CC BY 3.0*
- 11.1: Feedback Control - *CC BY 3.0*
- 11.2: Feed Forward Control - *CC BY 3.0*
- 11.3: Cascade Control - *CC BY 3.0*
- 11.4: Ratio Control - *CC BY 3.0*
- 11.5: Common control loops and model for liquid pressure and liquid level - *CC BY 3.0*
- 11.6: Common Control Loops and Model for Temperature Control - *CC BY 3.0*
- 11.7: Common control architectures and model for reactors - *CC BY 3.0*

◦ 12: Multiple Input, Multiple Output (MIMO) Control - *CC BY 3.0*
- 12.1: Determining if a system can be decoupled - *CC BY 3.0*
- 12.2: MIMO control using RGA - *CC BY 3.0*
- 12.3: MIMO using Model Predictive Control - *CC BY 3.0*
- 12.4: Neural Networks for automatic model construction - *CC BY 3.0*
- 12.5: Understanding MIMO Control Through Two Tanks Interaction - *CC BY 3.0*

◦ 13: Statistics and Probability Background - *CC BY 3.0*
- 13.1: Basic statistics- mean, median, average, standard deviation, z-scores, and p-value - *CC BY 3.0*
- 13.2: SPC- Basic Control Charts- Theory and Construction, Sample Size, X-Bar, R charts, S charts - *CC BY 3.0*
- 13.3: Six Sigma- What is it and what does it mean? - *CC BY 3.0*
- 13.4: Bayes Rule, Conditional Probability and Independence - *CC BY 3.0*
- 13.5: Bayesian Network Theory - *CC BY 3.0*
- 13.6: Learning and analyzing Bayesian networks with Genie - *CC BY 3.0*
- 13.7: Occasionally Dishonest Casino? Markov Chains and Hidden Markov Models - *CC BY 3.0*
- 13.8: Continuous Distributions- normal and exponential - *CC BY 3.0*
- 13.9: Discrete Distributions - Hypergeometric, Binomial, and Poisson - *CC BY 3.0*
- 13.10: Multinomial Distributions - *CC BY 3.0*
- 13.11: Comparisons of two means - *CC BY 3.0*
- 13.12: Factor analysis and ANOVA - *CC BY 3.0*
- 13.13: Correlation and Mutual Information - *CC BY 3.0*
- 13.14: Random sampling from a stationary Gaussian process - *CC BY 3.0*

◦ 14: Design of Experiments - *CC BY 3.0*
- 14.1: Design of Experiments via Taguchi Methods - Orthogonal Arrays - *CC BY 3.0*
- 14.2: Design of experiments via factorial designs - *CC BY 3.0*
- 14.3: Design of Experiments via Random Design - *CC BY 3.0*
- 14.4: Summary- Summary on Control Architectures' philosophies, advantages, and disadvantages. - *CC BY 3.0*

◦ Back Matter - *CC BY 3.0*
- Index - *CC BY 4.0*
- Glossary - *CC BY 4.0*
- Detailed Licensing - *Undeclared*

This page titled Chemical Process Dynamics and Controls (Woolf) is shared under a CC BY 3.0 license and was authored, remixed, and/or curated by Peter Woolf et al. via source content that was edited to the style and standards of the LibreTexts platform.