Fundamentals of Mechanical Engineering

Clinton Selby

First Edition, 2012

ISBN 978-81-323-2364-8

Published by:
Library Press
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter 1 - Statics and Dynamics (mechanics)

Chapter 2 - Strength of Materials and Solid Mechanics

Chapter 3 - Instrumentation and Measurement

Chapter 4 - Thermodynamics

Chapter 5 - Heat Transfer

Chapter 6 - Energy Transformation

Chapter 7 - HVAC

Chapter 8 - Hydraulics and Pneumatics

Chapter 9 - Fluid Mechanics

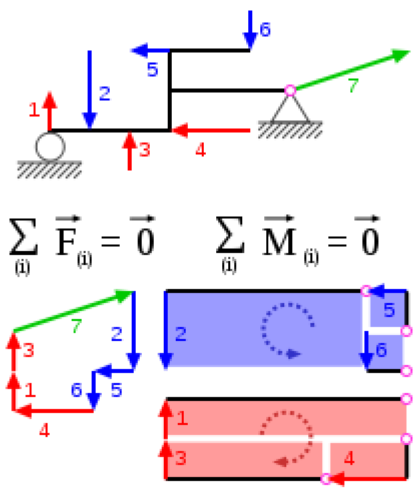
Chapter 10 - Fluid Dynamics

Chapter 11 - Mechatronics and Control Theory

Chapter- 1

Statics and Dynamics (mechanics)

Statics



Example of a beam in static equilibrium. The sum of force and moment is zero.

Statics is the branch of mechanics concerned with the analysis of loads (force, torque/moment) on physical systems in static equilibrium, that is, in a state where the relative positions of subsystems do not vary over time, or where components and structures are at a constant velocity. When in static equilibrium, the system is either at rest, or its center of mass moves at constant velocity.

By Newton's first law, this situation implies that the net force and net torque (also known as moment of force) on every body in the system is zero. From this constraint, such quantities as stress or pressure can be derived. The net forces equaling zero is known as the *first condition for equilibrium,* and the net torque equaling zero is known as the *second condition for equilibrium.*

Vectors

A scalar is a quantity like mass or temperature which only has a magnitude. A vector is a quantity that has both a magnitude and a direction, which is denoted by a bold faced character, an underlined character, or a character with an arrow on it. Vectors can be added using the parallelogram law. Vectors contain components in orthogonal bases. Unit vectors i, j, and k are along the x, y, and z directions.

Equilibrium Equations

The static equilibrium of a particle is an important concept in Statics. A particle is in equilibrium only if the resultant of all forces acting on the particle is equal to zero. In a rectangular coordinate system the equilibrium equations can be represented by three scalar equations, where the sum of forces in all three directions are equal to zero. An engineering application of this concept is determining the tensions of up to three cables under load, for example the forces exerted on each cable of a hoist lifting an object or of guy wires restraining a hot air balloon to the ground.

Moment of a Force

The magnitude of the moment of a force at a point O, is equal to the perpendicular distance from O to the line of action of F, multiplied by the magnitude of the force. The direction of the moment is given by the right hand rule, where counter clockwise (CCW) is out of the page, and clockwise (CW) is into the page. Moments can be added together as vectors.

Trusses

A truss is a structure made of two force members all pin connected to each other. There are two methods to analyzing the forces of members of trusses. The first method is the method of joints. This method uses the free-body-diagram of joints in the structure to determine the forces in each member. This method uses the equilibrium equations at each joint, and the members as forces to calculate the forces in each member. The method of sections uses free-body-diagrams of sections of the truss to obtain unknown forces. This method puts certain sections of the truss in equilibrium so the equations can be used for a limited amount of members at a time.

Moment of Inertia

In classical mechanics, moment of inertia, also called mass moment of inertia, rotational inertia, polar moment of inertia of mass, or the angular mass, (SI units $kg \cdot m^2$) is a measure of an object's resistance to changes to its rotation. It is the inertia of a rotating body with respect to its rotation. The moment of inertia plays much the same role in rotational dynamics as mass does in linear dynamics, describing the relationship between angular momentum and angular velocity, torque and angular acceleration, and several other quantities. The symbol I and sometimes J are usually used to refer to the moment of inertia or polar moment of inertia.

While a simple scalar treatment of the moment of inertia suffices for many situations, a more advanced tensor treatment allows the analysis of such complicated systems as spinning tops and gyroscopic motion.

The concept was introduced by Leonhard Euler in his book Theoria motus corporum solidorum seu rigidorum in 1765. In this book, he discussed the moment of inertia and many related concepts, such as the principal axis of inertia.

Solids

Statics is used in the analysis of structures, for instance in architectural and structural engineering. Strength of materials is a related field of mechanics that relies heavily on the application of static equilibrium. A key concept is the center of gravity of a body at rest: it represents an imaginary point at which all the mass of a body resides. The position of the point relative to the foundations on which a body lies determines its stability towards small movements. If the center of gravity exists outside the foundations, then the body is unstable because there is a torque acting: any small disturbance will cause the body to fall or topple. If the center of gravity exists within the foundations, the body is stable since no net torque acts on the body. If the center of gravity coincides with the foundations, then the body is said to be metastable.

Fluids

Hydrostatics, also known as fluid statics, is the study of fluids at rest. This analyzes bodies of fluid in static equilibrium. The characteristic of any fluid at rest is that the force exerted on any particle of the fluid is the same at all points at the same depth (or altitude) within the fluid. If the net force is greater than zero the fluid will move in the direction of the resulting force. This concept was first formulated in a slightly extended form by the French mathematician and philosopher Blaise Pascal in 1647 and would be later known as Pascal's Law. This law has many important applications in hydraulics. Archimedes, Abū Rayhān al-Bīrūnī, Al-Khazini and Galileo Galilei were also major figures in the development of hydrostatics.

Dynamics

In the field of physics, the study of the causes of motion and changes in motion is **dynamics**. In other words the study of forces and why objects are in motion. *Dynamics* includes the study of the effect of torques on motion. These are in contrast to Kinematics, the branch of classical mechanics that describes the motion of objects without consideration of the causes leading to the motion.

Generally speaking, researchers involved in dynamics study how a physical system might develop or alter over time and study the causes of those changes. In addition, Isaac Newton established the undergirding physical laws which govern dynamics in physics. By studying his system of mechanics, dynamics can be understood. In particular dynamics is mostly related to Newton's second law of motion. However, all three laws of motion are taken into consideration, because these are interrelated in any given observation or experiment.

For classical electromagnetism, it is Maxwell's equations that describe the dynamics. And the dynamics of classical systems involving both mechanics and electromagnetism are described by the combination of Newton's laws, Maxwell's equations, and the Lorentz force.

Force

From Newton, force can be defined as an exertion or pressure which can cause an object to move. The concept of force is used to describe an influence which causes a free body (object) to accelerate. It can be a push or a pull, which causes an object to change direction, to speed or have new velocity, or to deform temporarily or permanently. Generally speaking, force causes an object's state of motion to change.

Newton's laws

Newton described force as the ability to cause a mass to accelerate.

- Newton's first law states that an object in motion will stay in motion unless a force is applied. This law deals with inertia, which is a property of matter that resists acceleration and depends only on mass.
- Newton's second law states that force quantity is equal to mass multiplied by the acceleration ($F = ma$).
- Newton's third law states that for every action, there is an equal but opposite reaction.

Chapter- 2

Strength of Materials and Solid Mechanics

Strength of materials

In materials science, the **strength** of a material is its ability to withstand an applied stress without failure. The applied stress may be tensile, compressive, or shear. It is a subject which deals with loads, elastic and forces acting on the material. For example, an external load applied to an elastic material or internal forces acting on the material. Deformation (e.g. bending) of the material is called **strain**, while the intensity of the internal resisting force is called **stress**. The strength of any material relies on three different type of analytical method: strength, stiffness and stability, where strength means load carrying capacity, stiffness means deformation or elongation, and stability means ability to maintain its initial configuration. Yield strength refers to the point on the engineering stress-strain curve (as opposed to true stress-strain curve) beyond which the material begins deformation that cannot be reversed upon removal of the loading. Ultimate strength refers to the point on the engineering stress-strain curve corresponding to the maximum stress.

A material's strength is dependent on its microstructure. The engineering processes to which a material is subjected can alter this microstructure. The variety of strengthening mechanisms that alter the strength of a material includes work hardening, solid solution strengthening, precipitation hardening and grain boundary strengthening and can be quantified and qualitatively explained. However, strengthening mechanisms are accompanied by the caveat that some mechanical properties of the material may degenerate in an attempt to make the material stronger. For example, in grain boundary strengthening, although yield strength is maximized with decreasing grain size, ultimately, very small grain sizes make the material brittle. In general, the yield strength of a material is an adequate indicator of the material's mechanical strength. Considered in tandem with the fact that the yield strength is the parameter that predicts plastic deformation in the material, one can make informed decisions on how to increase the strength of a material depending its microstructural properties and the desired end effect. Strength is considered in terms of compressive strength, tensile strength, and shear strength, namely the limit states of compressive stress, tensile stress and shear stress, respectively. The effects of dynamic loading are probably the most important practical part of the strength of materials, especially the problem of fatigue. Repeated loading often initiates brittle cracks, which grow slowly until failure occurs.

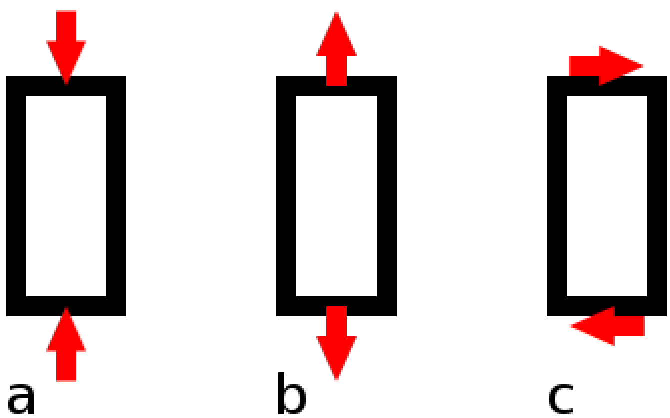
However, the term **strength of materials** most often refers to various methods of calculating stresses in structural members, such as beams, columns and shafts. The methods that can be employed to predict the response of a structure under loading and its susceptibility to various failure modes may take into account various properties of the materials other than material (yield or ultimate) strength. For example failure in buckling is dependent on material stiffness (Young's Modulus).

Types of loadings

- Transverse loading - Forces applied perpendicularly to the longitudinal axis of a member. Transverse loading causes the member to bend and deflect from its original position, with internal tensile and compressive strains accompanying change in curvature.
- Axial loading - The applied forces are collinear with the longitudinal axes of the members. The forces cause the member to either stretch or shorten.
- Torsional loading - Twisting action caused by a pair of externally applied equal and oppositely directed couples acting in a parallel planes or by a single external couple applied to a member that has one end fixed against rotation.

Definitions

Stress terms



A material being loaded in a) compression, b) tension, c) shear

Uniaxial stress is expressed by

$$\sigma = \frac{F}{A},$$

where F is the force [N] acting on an area A [m^2]. The area can be the undeformed area or the deformed area, depending on whether engineering stress or true stress is used.

- *Compressive stress* (or compression) is the stress state caused by an applied load that acts to reduce the length of the material (compression member) in the axis of the applied load, in other words the stress state caused by squeezing the material. A simple case of compression is the uniaxial compression induced by the action of opposite, pushing forces. Compressive strength for materials is generally higher than that of tensile stress. However, structures loaded in compression are subject to additional failure modes dependent on geometry, such as Euler buckling.

- *Tensile stress* is the stress state caused by an applied load that tends to elongate the material in the axis of the applied load, in other words the stress caused by *pulling* the material. The strength of structures of equal cross sectional area loaded in tension is independent of cross section geometry. Materials loaded in tension are susceptible to stress concentrations such as material defects or abrupt changes in geometry. However, materials exhibiting ductile behavior(metals for example) can tolerate some defects while brittle materials (such as ceramics) can fail well below their ultimate stress.

- *Shear stress* is the stress state caused by a pair of opposing forces acting along parallel lines of action through the material, in other words the stress caused by *sliding* faces of the material relative to one another. An example is cutting paper with scissors.

Strength terms

- *Yield strength* is the lowest stress that gives permanent deformation in a material. In some materials, like aluminium alloys, the point of yielding is hard to define, thus it is usually given as the stress required to cause 0.2% plastic strain. This is called a 0.2% proof stress.

- *Compressive strength* is a limit state of compressive stress that leads to compressive failure in the manner of ductile failure (infinite theoretical yield) or in the manner of brittle failure (rupture as the result of crack propagation, or sliding along a weak plane).

- *Tensile strength* or *ultimate tensile strength* is a limit state of tensile stress that leads to tensile failure in the manner of ductile failure (yield as the first stage of failure, some hardening in the second stage and break after a possible "neck" formation) or in the manner of brittle failure (sudden breaking in two or more pieces with a low stress state). Tensile strength can be given as either true stress or engineering stress.

- *Fatigue strength* is a measure of the strength of a material or a component under cyclic loading, and is usually more difficult to assess than the static strength measures. Fatigue strength is given as stress amplitude or stress range ($\Delta\sigma = \sigma_{max} - \sigma_{min}$), usually at zero mean stress, along with the number of cycles to failure.

- *Impact strength*, it is the capability of the material in withstanding by the suddenly applied loads in terms of energy. Often measured with the Izod impact strength test or Charpy impact test, both of which measure the impact energy required to fracture a sample. Volume, modulus of elasticity, distribution of forces, and yield strength effect the impact strength of a material. In order for a material or object to have a higher impact strength the stresses must be distributed evenly throughout the object. It also must have a large volume with a low modulus of elasticity and a high yield strength.

Strain (deformation) terms

- *Deformation* of the material is the change in geometry when stress is applied (in the form of force loading, gravitational field, acceleration, thermal expansion, etc.). Deformation is expressed by the displacement field of the material.
- *Strain* or *reduced deformation* is a mathematical term to express the trend of the deformation change among the material field. Strain is the deformation per unit length. For uniaxial loading - displacements of a specimen (for example a bar element) it is expressed as the quotient of the displacement and the length of the specimen. For 3D displacement fields it is expressed as derivatives of displacement functions in terms of a second order tensor (with 6 independent elements).
- *Deflection* is a term to describe the magnitude to which a structural element bends under a load.

Stress-strain relations

- *Elasticity* is the ability of a material to return to its previous shape after stress is released. In many materials, the relation between applied stress and the resulting strain is directly proportional (up to a certain limit), and a graph representing those two quantities is a straight line.

The slope of this line is known as Young's Modulus, or the "Modulus of Elasticity." The Modulus of Elasticity can be used to determine stress-strain relationships in the linear-elastic portion of the stress-strain curve. The linear-elastic region is either below the yield point, or if a yield point is not well defined for the material, taken to be between 0 and 0.2% strain, and is defined as the region of strain in which no yielding (permanent deformation) occurs.

- *Plasticity* or plastic deformation is the opposite of elastic deformation and is accepted as unrecoverable strain. Plastic deformation is retained even after the relaxation of the applied stress. Most materials in the linear-elastic category are usually capable of plastic deformation. Brittle materials, like ceramics, do not experience any plastic deformation and will fracture under relatively low stress. Materials such as metals usually experience a small amount of plastic deformation before failure while soft or ductile polymers will plasticly deform much more.

Consider the difference between a carrot and chewed bubble gum. The carrot will stretch very little before breaking, but nevertheless will still stretch. The chewed bubble gum, on the other hand, will plastically deform enormously before finally breaking.

Design terms

Ultimate strength is an attribute directly related to a material, rather than just specific specimen of the material, and as such is quoted force per unit of cross section area (N/m^2). The ultimate strength is the maximum stress that a material can withstand before it breaks or weakens. For example, the ultimate tensile strength (UTS) of AISI 1018 Steel is 440 MN/m^2. In general, the SI unit of stress is the pascal, where 1 Pa = 1 N/m^2. In Imperial units, the unit of stress is given as lbf/in^2 or pounds-force per square inch. This unit is often abbreviated as **psi**. One thousand psi is abbreviated **ksi**.

Factor of safety is a design constraint that an engineered component or structure must achieve. $FS = UTS / R$, where FS: the factor of safety, R: The applied stress, and UTS: ultimate stress (psi or N/m^2)

Margin of Safety is also sometimes used to as design constraint. It is defined MS = Factor of safety - 1

For example to achieve a factor of safety of 4, the allowable stress in an AISI 1018 steel component can be worked out as $R = UTS / FS = 440/4 = 110$ MPa, or $R = 110 \times 10^6$ N/m^2.

Design for cyclic loading working stress that have been determined from the ultimate or yield point values of the materials divided by a factor of safety give safe and reliable results only for static loading. Many machine parts are subjected to a non steady and continuously varying stress. Failures in such machine parts are usually caused by repeated loadings and at stresses that are below the yield point. When the part is below fatigue or endurance limit value, it will last indefinitely. However, the fatigue failures due to bending are the most common. These failures usually take place across the crystals. Failure is by fracture or separation without visible yielding or distortion. A purely reversing or cyclic stress means when the stress alternates between equal positive and negative peak stresses during each cycle of operation. Cyclic stress over time can be represented b a sinusoidal curve. In pure cyclic stress, the average stress is zero. When a part is subjected to cyclic stress, also known as range stress (Sr), it has been observed that the failure of the part occurs after a number of stress reversals (N) even if the magnitude of the range stress is below the material's yield strength. Generally, higher the range stress, lesser number of reversals is needed for failure.

Failure theories

We are interested in learning how static mechanical stress can cause failure in machine parts. Static stress means that the stress has been applied slowly and is maintained at a steady level. Failure from cyclic (or dynamic) stress and impact stress will be treated later. Here, we should also keep in mind that, there are many other factors such as, surface wear damage from friction, overheating, chemical corrosion, metallurgical fault or a combination of these and other factors may also cause failure. We will study four important failure theories, namely (1) maximum shear stress theory, (2) maximum normal stress theory, (3) maximum strain energy theory, and (4) maximum distortion energy theory. Out of these four theories of failure, the maximum normal stress theory is only applicable for brittle materials, and the remaining three theories are applicable for ductile materials.

- Maximum Shear stress Theory- This theory postulates that failure will occur in a machine part if the magnitude of the maximum shear stress (tmax) in the part exceeds the shear strength (typ) of the material determined from uniaxial testing.

This theory postulates, that failure will occur when, tmax = typ or max of [|S1-S2|/2, |S2-S3|/2 and |S3-S1|/2] = Syp/2 Dividing both side by 2, max of [|S1-S2|, |S2-S3| and |S3-S1|] = Syp Using a design factor of safety Nfs, the theory formulates the design equation as, max of [|S1-S2|, |S2-S3| and |S3-S1|] should be less than or equal to Syp/Nfs

- Maximum normal stress theory- this theory postulates, that failure will occur in machine part if the maximum normal stress in the part exceeds the ultimate tensile stress of the material as determined from uniaxial testing. This theory deals with brittle materials only. The maximum tensile stress should be less than or equal to ultimate tensile stress divided by factor of safety. The magnitude of the maximum compressive stress should be less than ultimate compressive stress divided by factor of safety.

As the three principal stresses at a point in the part S1, S2, or S3 may be both tensile and compressive stresses, when this theory is applied, we need to check for failures both from tension and compression. The method of application of this theory is to find the maximum tensile stress, and the maximum compressive stress from the given values of S1, S2, and S3. The largest positive value among S1, S2, and S3 is the maximum tensile stress and the smallest negative value is the maximum compressive stress. For example if S1 = 80 MPa, S2 = -100 MPa, and S3 = -150 MPa, then the maximum tensile stress = 80 MPa, and the maximum compressive stress = -150 MPa (smallest negative value!).Thus according to this theory, the safe design condition for brittle material can be given by: The maximum tensile stress should be less than or equal to Sut/Nfs and The magnitude of the maximum compressive stress should less than Suc/Nfs

- Maximum strain energy theory-this theory postulates that failure will occur when the strain energy per unit volume due to the applied stresses in a part equals the strain energy per unit volume at the yield point in uniaxial testing.

Strain energy is the energy stored in a material due elastic deformation, which is, work done during elastic deformation. Work done per unit volume = strain x average stress. During tensile test, stress increases from zero to Syp, that is average stress = Syp/2.Elastic strain at yield point = Syp/E, where E is the elastic modulus of elasticity.Strain energy per unit volume during uniaxial tension = average stress x strain = Syp2/2E

- Maximum distortion energy theory- this theory is also known as shear energy theory or von Mises-Hencky theory. This theory postulates that failure will occur when the distortion energy per unit volume due to the applied stresses in a part equals the distortion energy per unit volume at the yield point in uniaxial testing. The total elastic energy due to strain can be divided into two parts. One part causes change in volume, and the other part causes change in shape. Distortion energy is the amount of energy that is needed to change the shape.

Application of failure theory

Out of the four theories, only the maximum normal stress theory predicts failure for brittle materials. The rest of the three theories are applicable for ductile materials. Out of these three, the distortion energy theory provides most accurate results in majority of the stress conditions.The strain energy theory needs the value of Poisson's ratio of the part material, which is often not readily available. The maximum shear stress theory is conservative. For simple unidirectional normal stresses all theories are equivalent, which means all theories will give the same result.

Solid mechanics

Solid mechanics is the branch of mechanics, physics, and mathematics that concerns the behavior of solid matter under external actions (e.g., external forces, temperature changes, applied displacements, etc.). It is part of a broader study known as continuum mechanics. One of the most common practical applications of solid mechanics is the Euler-Bernoulli beam equation. Solid mechanics extensively uses tensors to describe stresses, strains, and the relationship between them.

Relationship to continuum mechanics

As shown in the following table, solid mechanics inhabits a central place within continuum mechanics. The field of rheology presents an overlap between solid and fluid mechanics.

<table>
<tr><td rowspan="2">Continuum mechanics
The study of the physics of continuous materials</td><td rowspan="2">Solid mechanics
The study of the physics of continuous materials with a defined rest shape.</td><td colspan="2">Elasticity
Describes materials that return to their rest shape after an applied stress.</td></tr>
<tr><td>Plasticity
Describes materials that permanently deform after a</td><td>Rheology
The study of materials with both solid and fluid</td></tr>
</table>

		sufficient applied stress.	characteristics.
	Fluid mechanics The study of the physics of continuous materials which take the shape of their container.	Non-Newtonian fluids	
		Newtonian fluids	

Response models

A material has a rest shape and its shape departs away from the rest shape due to stress. The amount of departure from rest shape is called deformation, the proportion of deformation to original size is called strain. If the applied stress is sufficiently low (or the imposed strain is small enough), almost all solid materials behave in such a way that the strain is directly proportional to the stress; the coefficient of the proportion is called the modulus of elasticity or Young's modulus. This region of deformation is known as the linearly elastic region.

It is most common for analysts in solid mechanics to use linear material models, due to ease of computation. However, real materials often exhibit non-linear behavior. As new materials are used and old ones are pushed to their limits, non-linear material models are becoming more common.

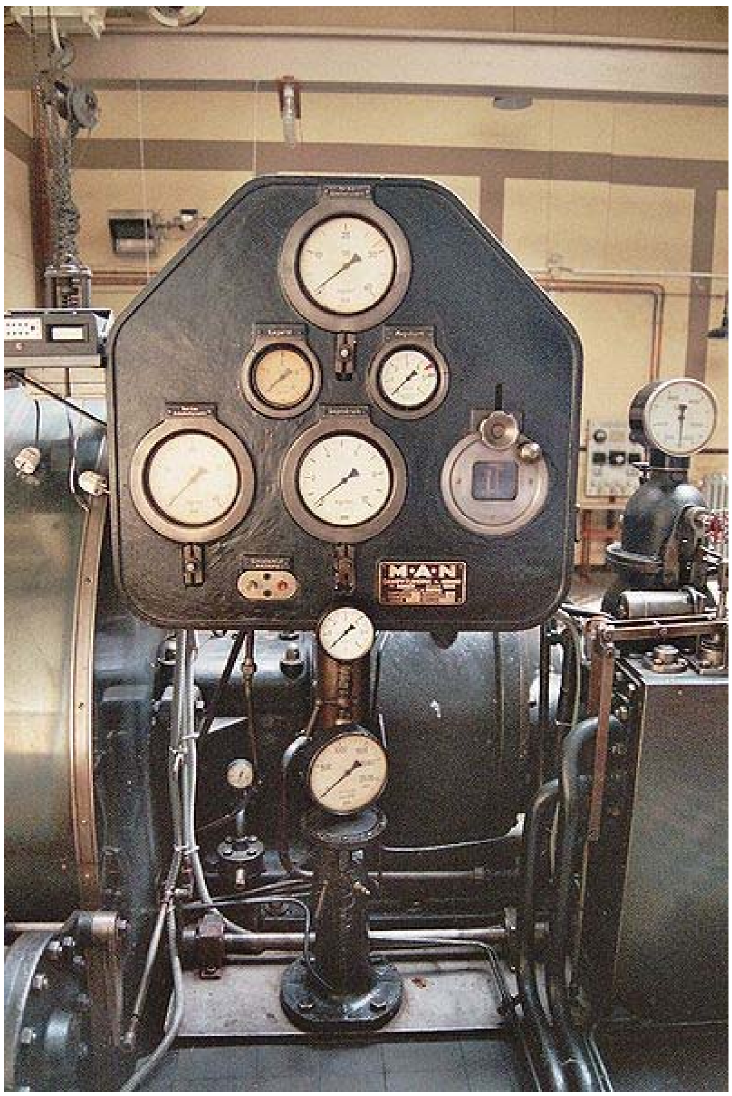
There are three models that describe how a solid responds to an applied stress:

1. Elastically – When an applied stress is removed, the material returns to its undeformed state. Linearly elastic materials, those that deform proportionally to the applied load, can be described by the linear elasticity equations such as Hooke's law.
2. Viscoelastically – These are materials that behave elastically, but also have damping: when the stress is applied and removed, work has to be done against the damping effects and is converted in heat within the material resulting in a hysteresis loop in the stress–strain curve. This implies that the material response has time-dependence.
3. Plastically – Materials that behave elastically generally do so when the applied stress is less than a yield value. When the stress is greater than the yield stress, the material behaves plastically and does not return to its previous state. That is, deformation that occurs after yield is permanent.

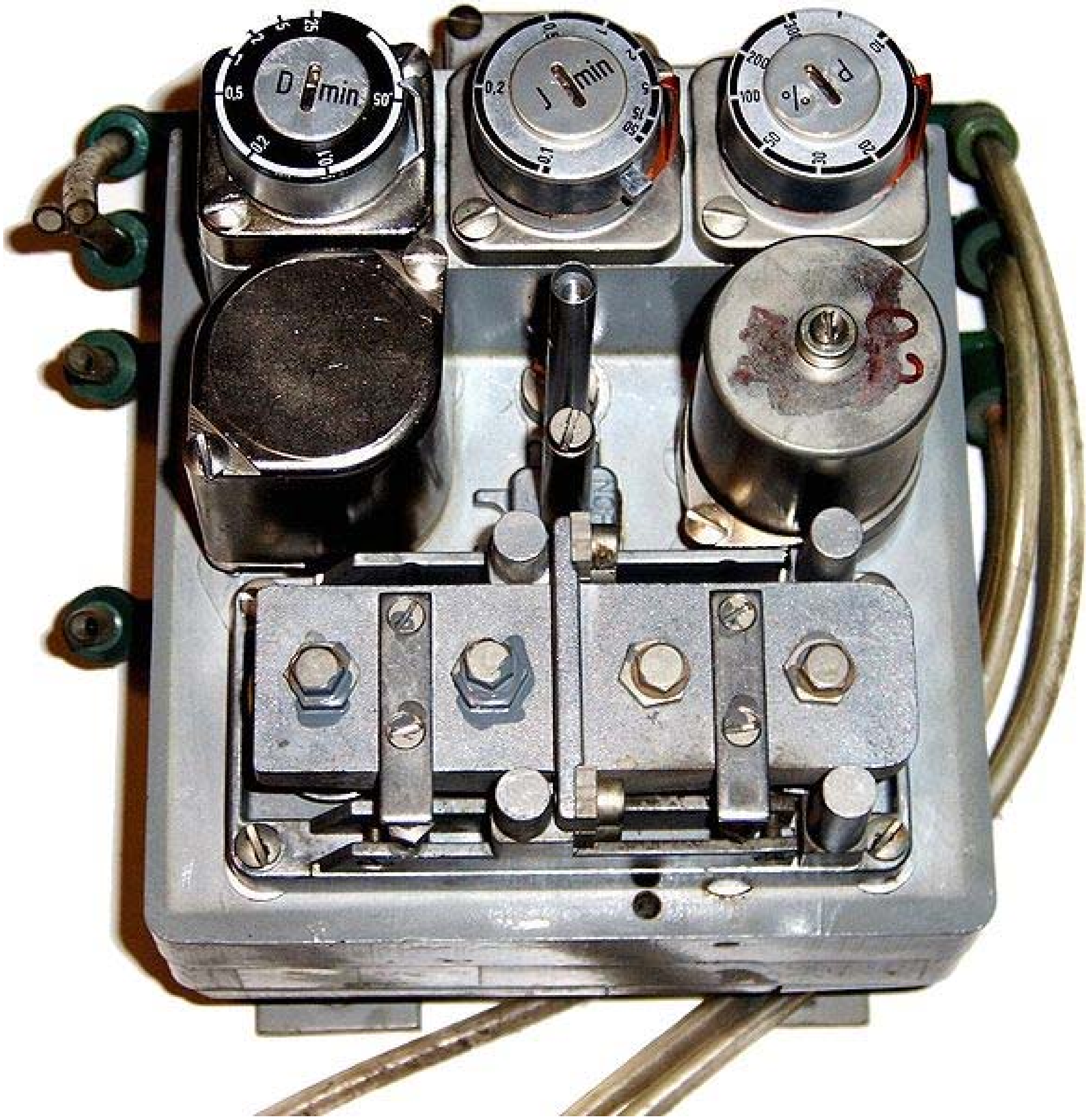
Chapter- 3

Instrumentation and Measurement

Instrumentation



A control post of a steam turbine



Pneumatic PID controller

Instrumentation is defined as the art and science of measurement and control.

An instrument is a device that measures and/or regulates process variables such as flow, temperature, level, or pressure. Instruments include many varied contrivances which can be as simple as valves and transmitters, and as complex as analyzers. Instruments often comprise control systems of varied processes such as refineries, factories, and vehicles. The control of processes is one of the main branches of applied instrumentation. Instrumentation can also refer to handheld devices that measure some desired variable. Diverse handheld instrumentation is common in laboratories, but can be found in the household as well. For example, a smoke detector is a common instrument found in most western homes.

Output instrumentation includes devices such as solenoids, valves, regulators, circuit breakers, and relays. These devices control a desired output variable, and provide either remote or automated control capabilities. These are often referred to as final control elements when controlled remotely or by a control system.

Transmitters are devices which produce an output signal, often in the form of a 4–20 mA electrical current signal, although many other options using voltage, frequency, pressure, or ethernet are possible. This signal can be used for informational purposes, or it can be sent to a PLC, DCS, SCADA system, or other type of computerized controller, where it can be interpreted into readable values and used to control other devices and processes in the system.

Control Instrumentation plays a significant role in both gathering information from the field and changing the field parameters, and as such are a key part of control loops.

History

In the early years of process control, process indicators and control elements such as valves were monitored by an operator that walked around the unit adjusting the valves to obtain the desired temperatures, pressures, and flows. As technology evolved pneumatic controllers were invented and mounted in the field that monitored the process and controlled the valves. This reduced the amount of time process operators were needed to monitor the process. Later years the actual controllers were moved to a central room and signals were sent into the control room to monitor the process and outputs signals were sent to the final control element such as a valve to adjust the process as needed. These controllers and indicators were mounted on a wall called a control board. The operators stood in front of this board walking back and forth monitoring the process indicators. This again reduced the number and amount of time process operators were needed to walk around the units. The basic air signal used during these years was 3-15 psig.

In the 1970s electronic instrumentation began to be manufactured by the instrument companies. Each instrument company came out with their own standard signal for their instrumentation, 10-50ma, 0.25-1.25Volts, 0-10Volts, 1-5volts, and 4-20ma, causing only confusion until the 4-20ma was universally used as a standard electronic instrument signal for transmitters and valves. The transformation of instrumentation from mechanical pneumatic transmitters, controllers, and valves to electronic instruments reduced maintenance costs as electronic instruments were more dependable than mechanical instruments. This also increased efficiency and production due to their increase in accuracy.

The next evolution of instrumentation came with the production of Distributed Control Systems (DCS). The pneumatic and electronic control rooms allowed control from a centralized room, DCS systems allowed control from more than one room or control stations. These stations could be next to each other or miles away. Now a process operator could sit in front of a screen and monitor thousands of points throughout a large unit or complex.

Measurement

Instrumentation is used to measure many parameters (physical values). These parameters include:

- Pressure, either differential or static
- Flow
- Temperature
- Levels of liquids etc.
- Density
- Viscosity
- Other mechanical properties of materials
- Properties of ionising radiation
- Frequency
- Current
- Voltage
- Inductance
- Capacitance
- Resistivity
- Chemical composition
- Chemical properties
- Properties of light
- Vibration
- Weight

Control

Control valve

In addition to measuring field parameters, instrumentation is also responsible for providing the ability to modify some field parameters.

Instrumentation engineering

Instrumentation engineering is the engineering specialization focused on the principle and operation of measuring instruments which are used in design and configuration of automated systems in electrical, pneumatic domains etc. They typically work for industries with automated processes, such as chemical or manufacturing plants, with the goal of improving system productivity, reliability, safety, optimization, and stability. To control the parameters in a process or in a particular system, devices such as microprocessors, microcontrollers or PLCs are used, but their ultimate aim is to control the parameters of a system.

Instrumentation technologists and mechanics

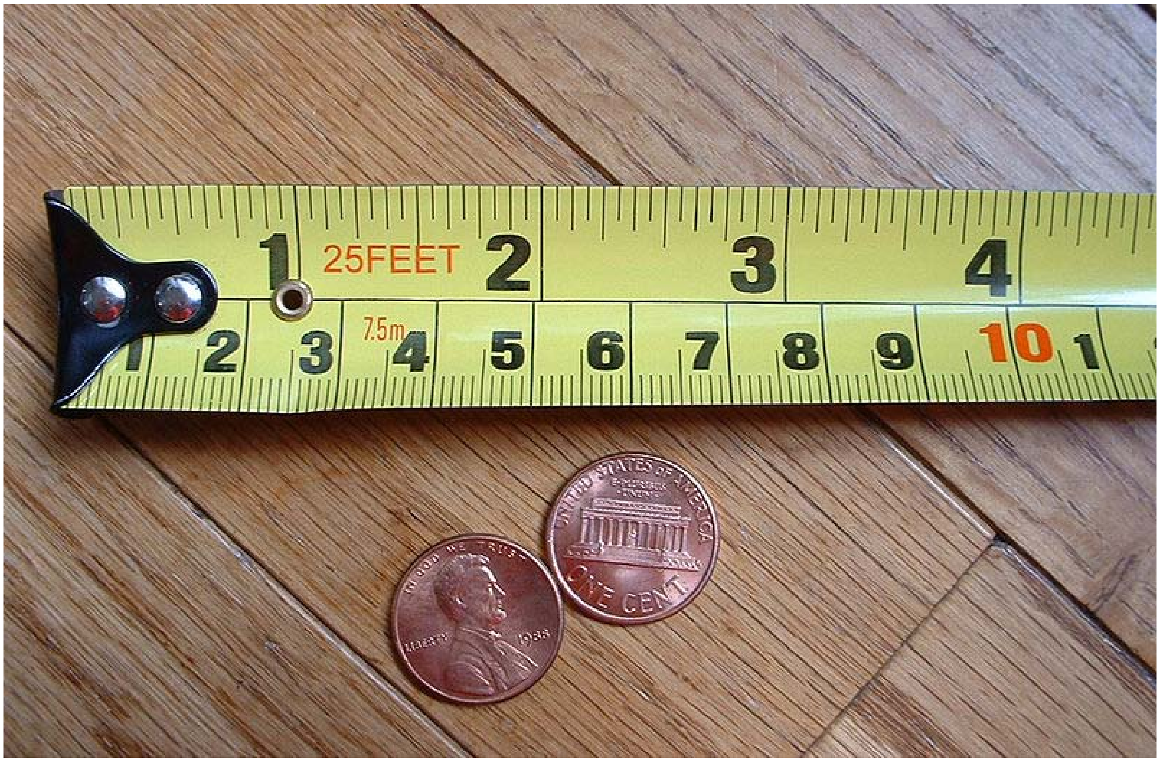
Instrumentation technologists, technicians and mechanics specialize in troubleshooting and repairing and maintenance of instruments and instrumentation systems. This trade is so intertwined with electricians, pipefitters, power engineers, and engineering companies, that one can find him/herself in extremely diverse working situations.

Measurement

Measurement is the process or the result of determining the magnitude of a quantity, such as length or mass, relative to a unit of measurement, such as a meter or a kilogram.

The word *measurement* stems, via the Middle French term *mesure*, from Latin *mēnsūra*, and the verb *metiri*.

Metrology is the science of measurement.



A typical tape measure with both metric and US units and two US pennies for comparison

Standards

With the exception of a few seemingly fundamental quantum constants, units of measurement are essentially arbitrary; in other words, people make them up and then agree to use them. Nothing inherent in nature dictates that an inch has to be a certain length, or that a mile is a better measure of distance than a kilometre. Over the course of human history, however, first for convenience and then for necessity, standards of measurement evolved so that communities would have certain common benchmarks. Laws regulating measurement were originally developed to prevent fraud in commerce. Today, units of measurement are generally defined on a scientific basis, overseen by governmental or supra-governmental agencies, and established in international treaties. The metre, for example, was redefined in 1983 as the distance traveled by light in free space in 1⁄299,792,458 of a second. In the United States, the National Institute of Standards and Technology (NIST), a division of the United States Department of Commerce, regulates commercial measurements. In the United Kingdom, the role is performed by the National Physical Laboratory (NPL).

Units and systems

Imperial system

Before SI units were widely adopted around the world, the British systems of English units and later imperial units were used in Britain, the Commonwealth and the United States. The system came to be known as U.S. customary units in the United States and is still in use there and in a few Caribbean countries. These various systems of measurement have at times been called *foot-pound-second* systems after the Imperial units for distance, weight and time even though the tons, hundredweights, gallons, and nautical miles, for example, are different for the U.S. units. Many Imperial units remain in use in Britain despite the fact that it has officially switched to the SI system. Road signs are still in miles, yards, miles per hour, and so on, people tend to measure their own height in feet and inches and milk is sold in pints, to give just a few examples. Imperial units are used in many other places, for example, in many Commonwealth countries that are considered metricated, land area is measured in acres and floor space in square feet, particularly for commercial transactions (rather than government statistics). Similarly, gasoline is sold by the gallon in many countries that are considered metricated.

Metric system

The metric system is a decimal systems of measurement based on its units for length, the metre and for mass, the kilogram. It exists in several variations, with different choices of base units, though these do not affect its day-to-day use. Since the 1960s, the International System of Units (SI) is the internationally recognized metric system. Metric units of mass, length, and electricity are widely used around the world for both everyday and scientific purposes.

The metric system features a single base unit for many physical quantities. Other quantities are derived from the standard SI units. Multiples and fractions of the units are expressed as powers of ten of each unit. Unit conversions are always simple because they are in the ratio of ten, one hundred, one thousand, etc., so that convenient magnitudes for measurements are achieved by simply moving the decimal place: 1.234 metres is 1234 millimetres or 0.001234 kilometres. The use of fractions, such as 2/5 of a metre, is not prohibited, but uncommon. All lengths and distances, for example, are measured in metres, or thousandths of a metre (millimetres), or thousands of metres (kilometres). There is no profusion of different units with different conversion factors as in the Imperial system which uses, for example, inches, feet, yards, fathoms, rods.

International System of Units

The International System of Units (abbreviated as SI from the French language name *Système International d'Unités*) is the modern revision of the metric system. It is the world's most widely used system of units, both in everyday commerce and in science. The SI was developed in 1960 from the metre-kilogram-second (MKS) system, rather than the centimetre-gram-second (CGS) system, which, in turn, had many variants. At its development the SI also introduced several newly named units that were previously not a part of the metric system. The SI units for the four basic physical quantities: length, time, mass, and temperature are:

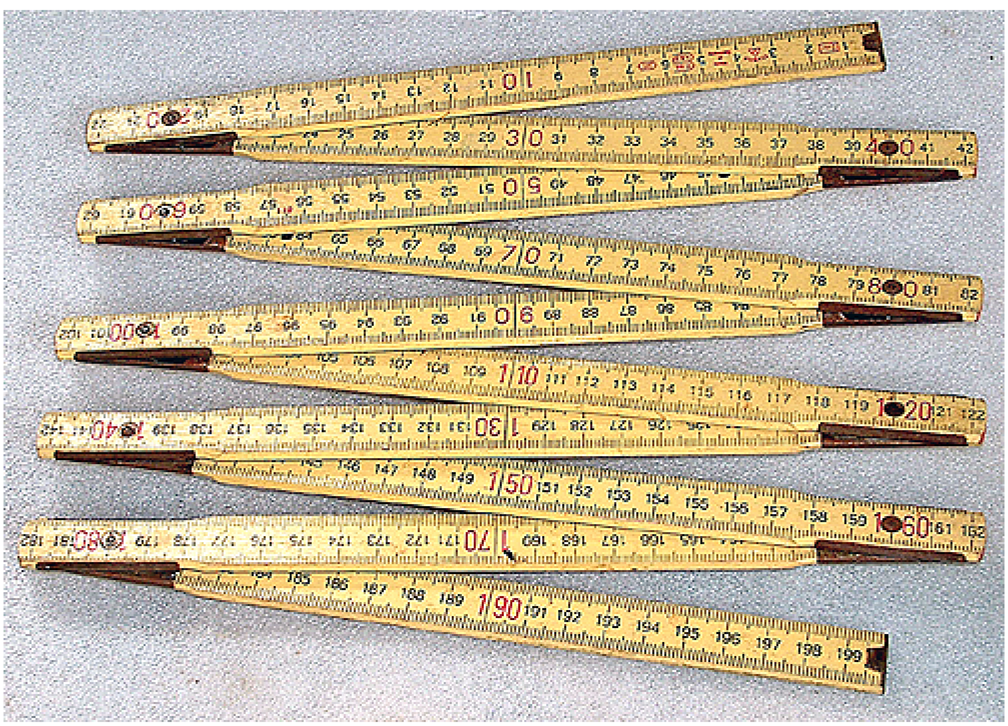
- metre (m): SI unit of length
- second (s): SI unit of time
- kilogram (kg: SI unit of mass
- kelvin (K): SI unit of temperature

There are two types of SI units, base units and derived units. Base units are the simple measurements for time, length, mass, temperature, amount of substance, electric current and light intensity. Derived units are constructed from the base units, for example, the watt, i.e. the unit for power, is defined from the base units as $m^2 \cdot kg \cdot s^{-3}$. Other physical properties may be measured in compound units, such as material density, measured in kg/m^3.

Converting prefixes

The SI allows easy multiplication when switching among units having the same base but different prefixes. To convert from metres to centimetres it is only necessary to multiply the number of metres by 100, since there are 100 centimetres in a metre. Inversely, to switch from centimetres to metres one multiplies the number of centimetres by 0.01 or divide centimetres by 100.

Distance



A 2-metre carpenter's ruler

A ruler or rule is a tool used in, for example, geometry, technical drawing, engineering, and carpentry, to measure distances or to draw straight lines. Strictly speaking, the *ruler* is the instrument used to **rule** straight lines and the calibrated instrument used for determining length is called a *measure*, however common usage calls both instruments *rulers* and the special name *straightedge* is used for an unmarked rule. The use of the word *measure*, in the sense of a measuring instrument, only survives in the phrase *tape measure*, an instrument that can be used to measure but cannot be used to draw straight lines. As can be seen in the photographs on this page, a two-metre carpenter's rule can be folded down to a length of only 20 centimetres, to easily fit in a pocket, and a five-metre long tape measure easily retracts to fit within a small housing.

Some special names

We also use some special names for some multiples of some units.

- 100 kilograms = 1 quintal;1000 kilogram = 1 metric tonne;
- 10 years = 1 decade; 100 years = 1 century; 1000 years = 1 millennium

Building trades

The Australian building trades adopted the metric system in 1966 and the units used for measurement of length are metres (m) and millimetres (mm). Centimetres (cm) are avoided as they cause confusion when reading plans. For example, the length two and a half metres is usually recorded as 2500 mm or 2.5 m; it would be considered non-standard to record this length as 250 cm.

Mass

Mass refers to the intrinsic property of all material objects to resist changes in their momentum. *Weight*, on the other hand, refers to the downward force produced when a mass is in a gravitational field. In free fall, (no net gravitational forces) objects lack weight but retain their mass. The Imperial units of mass include the ounce, pound, and ton. The metric units gram and kilogram are units of mass.

One device for measuring weight or mass is called a weighing scale or, often, simply a *scale*. A spring scale measures force but not mass, a balance compares masses, but requires a gravitational field to operate. Some of the most accurate instruments for measuring weight or mass are based on load cells with a digital read-out, but require a gravitational field to function and would not work in free fall.

Economics

The measures used in economics are physical measures, nominal price value measures and fixed price value measures. These measures differ from one another by the variables they measure and by the variables excluded from measurements. The measurable variables in economics are quantity, quality and distribution. By excluding variables from measurement makes it possible to better focus the measurement on a given variable, yet, this means a narrower approach.

Difficulties

Since accurate measurement is essential in many fields, and since all measurements are necessarily approximations, a great deal of effort must be taken to make measurements as accurate as possible. For example, consider the problem of measuring the time it takes an object to fall a distance of one metre (39 in). Using physics, it can be shown that, in the gravitational field of the Earth, it should take any object about 0.45 second to fall one metre. However, the following are just some of the sources of error that arise. First, this computation used for the acceleration of gravity 9.8 metres per second per second (32.2 ft/s^2). But this measurement is not exact, but only precise to two significant digits. Also, the Earth's gravitational field varies slightly depending on height above sea level and other factors. Next, the computation of .45 seconds involved extracting a square root, a mathematical operation that required rounding off to some number of significant digits, in this case two significant digits.

So far, we have only considered scientific sources of error. In actual practice, dropping an object from a height of a metre stick and using a stopwatch to time its fall, we have other sources of error. First, and most common, is simple carelessness. Then there is the problem of determining the exact time at which the object is released and the exact time it hits the ground. There is also the problem that the measurement of the height and the measurement of the time both involve some error. Finally, there is the problem of air resistance.

Scientific experiments must be carried out with great care to eliminate as much error as possible, and to keep error estimates realistic.

Definitions and theories

Classical definition

In the classical definition, which is standard throughout the physical sciences, *measurement* is the determination or estimation of ratios of quantities. Quantity and measurement are mutually defined: quantitative attributes are those possible to measure, at least in principle. The classical concept of quantity can be traced back to John Wallis and Isaac Newton, and was foreshadowed in Euclid's Elements.In Steven N. S. Cheung's definition,

measurement involves an assignment of numbers for the purposes of ranking, and precision in measurement can only be judged by the extent of the agreement among different observers

Representational theory

In the representational theory, *measurement* is defined as "the correlation of numbers with entities that are not numbers". The strongest form of representational theory is also known as additive conjoint measurement. In this form of representational theory, numbers are assigned based on correspondences or similarities between the structure of number systems and the structure of qualitative systems. A property is quantitative if such structural similarities can be established. In weaker forms of representational theory, such as that implicit within the work of Stanley Smith Stevens, numbers need only be assigned according to a rule.

The concept of measurement is often misunderstood as merely the assignment of a value, but it is possible to assign a value in a way that is not a measurement in terms of the requirements of additive conjoint measurement. One may assign a value to a person's height, but unless it can be established that there is a correlation between measurements of height and empirical relations, it is not a measurement according to additive conjoint measurement theory. Likewise, computing and assigning arbitrary values, like the "book value" of an asset in accounting, is not a measurement because it does not satisfy the necessary criteria.

Information theory

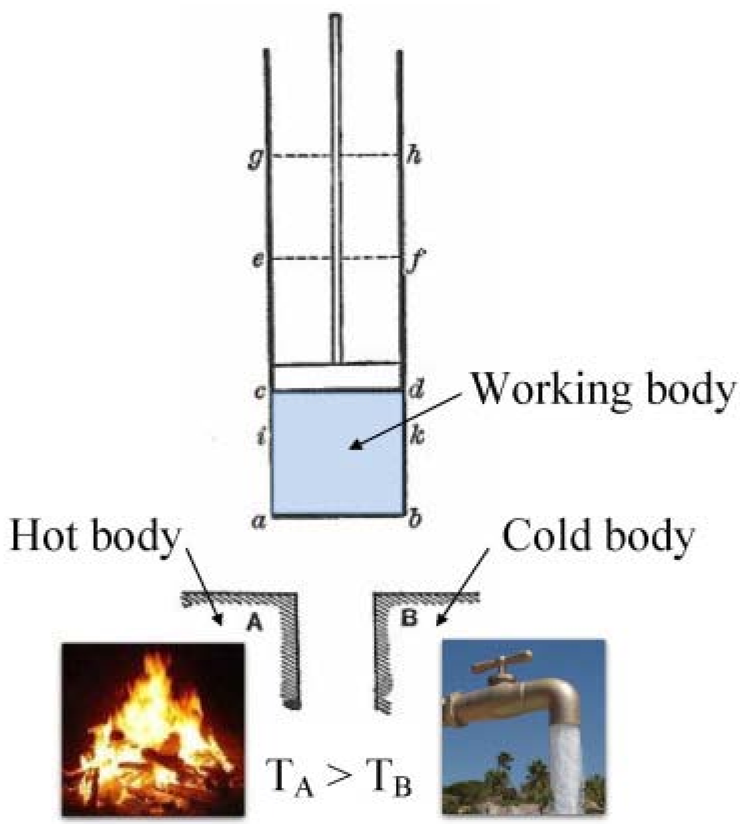
Information theory recognizes that all data are inexact and statistical in nature. Thus the definition of measurement is: "A set of observations that reduce uncertainty where the result is expressed as a quantity." This definition is implied in what scientists actually do when they measure something and report both the mean and statistics of the measurements. In practical terms, one begins with an initial guess as to the value of a quantity, and then, using various methods and instruments, reduces the uncertainty in the value. Note that in this view, unlike the positivist representational theory, all measurements are uncertain, so instead of assigning one value, a range of values is assigned to a measurement. This also implies that there is a continuum between estimation and measurement.

Quantum mechanics

In quantum mechanics, a measurement is the *collapse of the wavefunction*. The unambiguous meaning of the measurement problem is an unresolved fundamental problem in quantum mechanics.

Chapter- 4

Thermodynamics



Annotated color version of the original 1824 Carnot heat engine showing the hot body (boiler), working body (system, steam), and cold body (water), the letters labeled according to the stopping points in Carnot cycle.

Thermodynamics is the science of energy conversion involving heat and other forms of energy, most notably mechanical work. It studies and interrelates the macroscopic variables, such as temperature, volume and pressure, which describe physical, thermodynamic systems.

Historically, thermodynamics developed out of a desire to increase the efficiency of early steam engines, particularly through the work of French physicist Nicolas Léonard Sadi

Carnot (1824) who believed that engine efficiency was the key that could help France win the Napoleonic Wars. Scottish physicist Lord Kelvin was the first to formulate a concise definition of thermodynamics when he stated in 1854:

Thermo-dynamics is the subject of the relation of heat to forces acting between contiguous parts of bodies, and the relation of heat to electrical agency.

Two fields of thermodynamics emerged in the following decades. Statistical thermodynamics, or statistical mechanics, (1860) concerned itself with statistical predictions of the collective motion of particles from their microscopic behavior, while chemical thermodynamics (1873) studies the nature of the role of entropy in the process of chemical reaction.

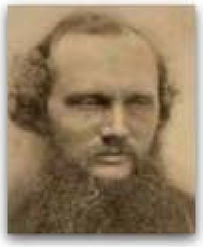
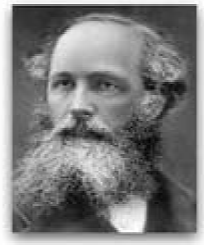
Introduction

The starting point for most thermodynamic considerations are the laws of thermodynamics, which postulate that energy can be exchanged between physical systems as heat or work. They also postulate the existence of a quantity named entropy, which can be defined for any isolated system that is in thermodynamic equilibrium.

In thermodynamics, interactions between large ensembles of objects are studied and categorized. Central to this are the concepts of *system* and *surroundings*. A system is composed of particles, whose average motions define its properties, which in turn are related to one another through equations of state. Properties can be combined to express internal energy and thermodynamic potentials, which are useful for determining conditions for equilibrium and spontaneous processes.

With these tools, thermodynamics can be used to describe how systems respond to changes in their environment. This can be applied to a wide variety of topics in science and engineering, such as engines, phase transitions, chemical reactions, transport phenomena, and even black holes. The results of thermodynamics are essential for other fields of physics and for chemistry, chemical engineering, aerospace engineering, mechanical engineering, cell biology, biomedical engineering, materials science, and economics, to name a few.

The present article is focused mainly on classical thermodynamics, which is concerned with systems in thermodynamic equilibrium. It is wise to distinguish classical thermodynamics from non-equilibrium thermodynamics, which is concerned with systems that are not in thermodynamic equilibrium.

École Polytechnique	Glasgow school	Berlin school	Edinburgh school
 Sadi Carnot (1796-1832)	 William Thomson (1824-1907)	 Rudolf Clausius (1822-1888)	 James Maxwell (1831-1879)
Vienna school	**Gibbsian school**	**Dresden school**	**Dutch school**
 Ludwig Boltzmann (1844-1906)	 Willard Gibbs (1839-1903)	 Gustav Zeuner (1828-1907)	 Johannes der Waals (1837-1923)

The thermodynamicists representative of the original eight founding schools of thermodynamics. The schools with the most-lasting effect in founding the modern versions of thermodynamics are the Berlin school, particularly as established in Rudolf Clausius's 1865 textbook *The Mechanical Theory of Heat*, the Vienna school, with the statistical mechanics of Ludwig Boltzmann, and the Gibbsian school at Yale University, American engineer Willard Gibbs' 1876 *On the Equilibrium of Heterogeneous Substances* launching chemical thermodynamics.

History

The history of thermodynamics as a scientific discipline generally begins with Otto von Guericke who, in 1650, built and designed the world's first vacuum pump and demonstrated a vacuum using his Magdeburg hemispheres. Guericke was driven to make a vacuum in order to disprove Aristotle's long-held supposition that 'nature abhors a vacuum'. Shortly after Guericke, the English physicist and chemist Robert Boyle had learned of Guericke's designs and, in 1656, in coordination with English scientist Robert Hooke, built an air pump. Using this pump, Boyle and Hooke noticed a correlation between pressure, temperature, and volume. In time, Boyle's Law was formulated, which states that pressure and volume are inversely proportional. Then, in 1679, based on these concepts, an associate of Boyle's named Denis Papin built a steam digester, which was a closed vessel with a tightly fitting lid that confined steam until a high pressure was generated.

Later designs implemented a steam release valve that kept the machine from exploding. By watching the valve rhythmically move up and down, Papin conceived of the idea of a piston and a cylinder engine. He did not, however, follow through with his design. Nevertheless, in 1697, based on Papin's designs, engineer Thomas Savery built the first engine, followed by Thomas Newcomen in 1712. Although these early engines were crude and inefficient, they attracted the attention of the leading scientists of the time.

The fundamental concepts of heat capacity and latent heat, which were necessary for the development of thermodynamics, were developed by Professor Joseph Black at the University of Glasgow, where James Watt was employed as an instrument maker. Black and Watt performed experiments together, but it was Watt who conceived the idea of the external condenser which resulted in a large increase in steam engine efficiency. Drawing on all the previous work led Sadi Carnot, the "father of thermodynamics", to publish *Reflections on the Motive Power of Fire* (1824), a discourse on heat, power, energy and engine efficiency. The paper outlined the basic energetic relations between the Carnot engine, the Carnot cycle, and motive power. It marked the start of thermodynamics as a modern science.

The first thermodynamic textbook was written in 1859 by William Rankine, originally trained as a physicist and a civil and mechanical engineering professor at the University of Glasgow. The first and second laws of thermodynamics emerged simultaneously in the 1850s, primarily out of the works of William Rankine, Rudolf Clausius, and William Thomson (Lord Kelvin).

The foundations of statistical thermodynamics were set out by physicists such as James Clerk Maxwell, Ludwig Boltzmann, Max Planck, Rudolf Clausius and J. Willard Gibbs.

During the years 1873-76 the American mathematical physicist Josiah Willard Gibbs published a series of three papers, the most famous being *On the Equilibrium of Heterogeneous Substances*, in which he showed how thermodynamic processes could be graphically analyzed, by studying the energy, entropy, volume, temperature and pressure of the thermodynamic system in such a manner, one can determine if a process would occur spontaneously. During the early 20th century, chemists such as Gilbert N. Lewis, Merle Randall, and E. A. Guggenheim began to apply the mathematical methods of Gibbs to the analysis of chemical processes.

Etymology

The etymology of *thermodynamics* has an intricate history.It was first spelled in a hyphenated form as an adjective (*thermo-dynamic*) and from 1854 to 1868 as the noun *thermo-dynamics* to represent the science of generalized heat engines.

American biophysisist Donald Haynie claims that *thermodynamics* was coined in 1840 from the Greek root θέρμη *therme,* meaning heat and δύναμις,*dynamis,* meaning power. However, this etymology has been cited as unlikely.

Pierre Perrot claims that the term *thermodynamics* was coined by James Joule in 1858 to designate the science of relations between heat and power., however, Joule never used that term, but used instead the term *perfect thermo-dynamic engine* in reference to Thomson’s 1849 phraseology.

By 1858, *thermo-dynamics*, as a functional term, was used in William Thomson's paper *An Account of Carnot's Theory of the Motive Power of Heat.*

Interpretations

Thermodynamics has developed into several related branches of science, each with a different focus.

Classical thermodynamics

Classical thermodynamics is the description of the states of thermodynamical systems at near-equilibrium, using macroscopic, empirical properties directly measurable in the laboratory. It is used to model exchanges of energy, work and heat based on the laws of thermodynamics. The qualifier *classical* reflects the fact that it represents the level of knowledge in the early 19th century. An atomic interpretation of these principles was provided later by the development of statistical mechanics.

Statistical mechanics

Statistical mechanics (or statistical thermodynamics) emerged only with the development of atomic and molecular theories in the late 19th century and early 20th century, giving thermodynamics a molecular interpretation. This field relates the microscopic properties of individual atoms and molecules to the macroscopic or bulk properties of materials that can be observed in everyday life, thereby explaining thermodynamics as a natural result of statistics and mechanics (classical and quantum) at the microscopic level. This statistical approach is in contrast to classical thermodynamics, which is a more phenomenological approach.

Chemical thermodynamics

Chemical thermodynamics is the study of the interrelation of energy with chemical reactions or with a physical change of state within the confines of the laws of thermodynamics.

Treatment of equilibrium

Equilibrium thermodynamics is the systematic study of transformations of matter and energy in systems as they approach equilibrium. The word equilibrium implies a state of balance. In an equilibrium state there are no unbalanced potentials, or driving forces, within the system. A central aim in equilibrium thermodynamics is: given a system in a well-defined initial state, subject to accurately specified constraints, to calculate what the state of the system will be once it has reached equilibrium.

Non-equilibrium thermodynamics is a branch of thermodynamics that deals with systems that are not in thermodynamic equilibrium. Most systems found in nature are not in thermodynamic equilibrium because they are not in stationary states, and are continuously and discontinuously subject to flux of matter and energy to and from other systems. The thermodynamic study of non-equilibrium systems requires more general concepts than are dealt with by equilibrium thermodynamics. Many natural systems still today remain beyond the scope of currently known macroscopic thermodynamic methods.

Laws of thermodynamics

Thermodynamics defines four laws which do not depend on the details of the systems under study or how they interact. Hence these laws are generally valid and can be applied to systems about which one knows nothing other than the balance of energy and matter transfer. Examples of such systems include Einstein's prediction of spontaneous emission, and ongoing research into the thermodynamics of black holes.

These four laws are:

- Zeroth law of thermodynamics: *If two systems are in thermal equilibrium with a third, they are also in thermal equilibrium with each other.*

This statement implies that thermal equilibrium is an equivalence relation on the set of thermodynamic systems under consideration. Systems are said to be in equilibrium if the small, random exchanges between them (eg. Brownian motion) do not lead to a net change in energy. This law is tacitly assumed in every measurement of temperature. Thus, if one seeks to decide if two bodies are at the same temperature, it is not necessary to bring them into contact and measure any changes of their observable properties in time. The law provides a fundamental definition of temperature and justification for the construction of practical thermometers.

It is interesting to note that the zeroth law was not initially recognized as a law. The need to for the zeroth law was not initially realized, so the first, second, and third laws were explicitly stated and found common acceptance in the physics community first. Once the importance of the zeroth law was realized, it was impracticable to renumber the other laws, hence the *zeroth.*

- First law of thermodynamics: *The internal energy of an isolated system is constant.*

The first law of thermodynamics is an expression of the principle of conservation of energy. It states that energy can be transformed (changed from one form to another), but cannot be created or destroyed.

The first law is usually formulated by saying that the change in the internal energy of a closed thermodynamic system is equal to the difference between the of heat supplied to the system and the amount of work done by the system on its surroundings. It is important to note that internal energy is a state of the system whereas heat and work modify the state of the system. In other words, a specific internal energy of a system may be achieved by any combination of heat and work; the manner by which a system achieves a specific internal energy is path independent.

- Second law of thermodynamics: *Heat cannot spontaneously flow from a colder location to a hotter location.*

The second law of thermodynamics is an expression of the universal principle of decay observable in nature. The second law is an observation of the fact that over time, differences in temperature, pressure, and chemical potential tend to even out in a physical system that is isolated from the outside world. Entropy is a measure of how much this evening-out process has progressed. The entropy of an isolated system which is not in equilibrium will tend to increase over time, approaching a maximum value at equilibrium.

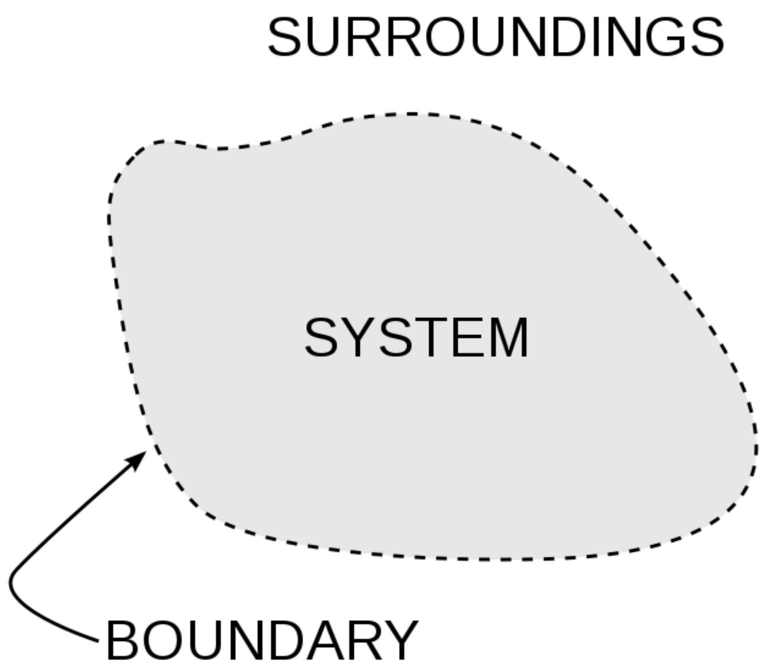
In classical thermodynamics, the second law is a basic postulate applicable to any system involving heat energy transfer; in statistical thermodynamics, the second law is a consequence of the assumed randomness of molecular chaos. There are many versions of the second law, but they all have the same effect, which is to explain the phenomenon of irreversibility in nature.

- Third law of thermodynamics: *As a system approaches absolute zero, all processes cease and the entropy of the system approaches a minimum value.*

The third law of thermodynamics is a statistical law of nature regarding entropy and the impossibility of reaching absolute zero of temperature. This law provides an absolute reference point for the determination of entropy. The entropy determined relative to this point is the absolute entropy. Alternate definitions are, "the entropy of all systems and of all states of a system is smallest at absolute zero," or equivalently "it is impossible to reach the absolute zero of temperature by any finite number of processes".

Absolute zero, at which all activity would stop if it were possible to happen, is −273.15 °C (degrees Celsius), or −459.67 °F (degrees Fahrenheit) or 0 K (kelvin).

System models



A diagram of a generic thermodynamic system

An important concept in thermodynamics is the thermodynamic system, a precisely defined region of the universe under study. Everything in the universe except the system is known as the *surroundings*. A system is separated from the remainder of the universe by a *boundary* which may be notional or not, but which by convention delimits a finite volume. Exchanges of work, heat, or matter between the system and the surroundings take place across this boundary.

In practice, the boundary is simply an imaginary dotted line drawn around a volume when there is going to be a change in the internal energy of that volume. Anything that passes across the boundary that effects a change in the internal energy needs to be accounted for in the energy balance equation. The volume can be the region surrounding a single atom resonating energy, such as Max Planck defined in 1900; it can be a body of steam or air in a steam engine, such as Sadi Carnot defined in 1824; it can be the body of a tropical cyclone, such as Kerry Emanuel theorized in 1986 in the field of atmospheric thermodynamics; it could also be just one nuclide (i.e. a system of quarks) as hypothesized in quantum thermodynamics.

Boundaries are of four types: fixed, moveable, real, and imaginary. For example, in an engine, a fixed boundary means the piston is locked at its position; as such, a constant volume process occurs. In that same engine, a moveable boundary allows the piston to move in and out. For closed systems, boundaries are real while for open system boundaries are often imaginary.

Generally, thermodynamics distinguishes three classes of systems, defined in terms of what is allowed to cross their boundaries:

Interactions of thermodynamic systems

Type of system	Mass flow	Work	Heat
Open	✓	✓	✓
Closed	✗	✓	✓
Isolated	✗	✗	✗

As time passes in an isolated system, internal differences in the system tend to even out and pressures and temperatures tend to equalize, as do density differences. A system in which all equalizing processes have gone to completion is considered to be in a state of thermodynamic equilibrium.

In thermodynamic equilibrium, a system's properties are, by definition, unchanging in time. Systems in equilibrium are much simpler and easier to understand than systems which are not in equilibrium. Often, when analysing a thermodynamic process, it can be assumed that each intermediate state in the process is at equilibrium. This will also considerably simplify the situation. Thermodynamic processes which develop so slowly as to allow each intermediate step to be an equilibrium state are said to be reversible processes.

States and processes

When a system is at equilibrium under a given set of conditions, it is said to be in a definite thermodynamic state. The state of the system can be described by a number of intensive variables and extensive variables. The properties of the system can be described by an equation of state which specifies the relationship between these variables. State may be thought of as the instantaneous quantitative description of a system with a set number of variables held constant.

A thermodynamic process may be defined as the energetic evolution of a thermodynamic system proceeding from an initial state to a final state. Typically, each thermodynamic process is distinguished from other processes in energetic character according to what parameters, such as temperature, pressure, or volume, etc., are held fixed. Furthermore, it is useful to group these processes into pairs, in which each variable held constant is one member of a conjugate pair.

Several commonly studied thermodynamic processes are:

- Isobaric process: occurs at constant pressure
- Isochoric process: occurs at constant volume (also called isometric/isovolumetric)

- Isothermal process: occurs at a constant temperature
- Adiabatic process: occurs without loss or gain of energy by heat
- Isentropic process: a reversible adiabatic process, occurs at a constant entropy
- Isenthalpic process: occurs at a constant enthalpy
- Steady state process: occurs without a change in the internal energy

Instrumentation

There are two types of thermodynamic instruments, the **meter** and the **reservoir**. A thermodynamic meter is any device which measures any parameter of a thermodynamic system. In some cases, the thermodynamic parameter is actually defined in terms of an idealized measuring instrument. For example, the zeroth law states that if two bodies are in thermal equilibrium with a third body, they are also in thermal equilibrium with each other. This principle, as noted by James Maxwell in 1872, asserts that it is possible to measure temperature. An idealized thermometer is a sample of an ideal gas at constant pressure. From the ideal gas law $pV=nRT$, the volume of such a sample can be used as an indicator of temperature; in this manner it defines temperature. Although pressure is defined mechanically, a pressure-measuring device, called a barometer may also be constructed from a sample of an ideal gas held at a constant temperature. A calorimeter is a device which is used to measure and define the internal energy of a system.

A thermodynamic reservoir is a system which is so large that it does not appreciably alter its state parameters when brought into contact with the test system. It is used to impose a particular value of a state parameter upon the system. For example, a pressure reservoir is a system at a particular pressure, which imposes that pressure upon any test system that it is mechanically connected to. The Earth's atmosphere is often used as a pressure reservoir.

Conjugate variables

The central concept of thermodynamics is that of energy, the ability to do work. By the First Law, the total energy of a system and its surroundings is conserved. Energy may be transferred into a system by heating, compression, or addition of matter, and extracted from a system by cooling, expansion, or extraction of matter. In mechanics, for example, energy transfer equals the product of the force applied to a body and the resulting displacement.

Conjugate variables are pairs of thermodynamic concepts, with the first being akin to a "force" applied to some thermodynamic system, the second being akin to the resulting "displacement," and the product of the two equalling the amount of energy transferred. The common conjugate variables are:

- Pressure-volume (the mechanical parameters);
- Temperature-entropy (thermal parameters);
- Chemical potential-particle number (material parameters).

Potentials

Thermodynamic potentials are different quantitative measures of the stored energy in a system. Potentials are used to measure energy changes in systems as they evolve from an initial state to a final state. The potential used depends on the constraints of the system, such as constant temperature or pressure. For example, the Helmholtz and Gibbs energies are the energies available in a system to do useful work when the temperature and volume or the pressure and temperature are fixed, respectively.

The five most well known potentials are:

Name	Symbol	Formula	Natural variables
Internal energy	U	$\int (TdS - pdV + \sum_i \mu_i dN_i)$	$S,V,\{N_i\}$
Helmholtz free energy	F, A	$U - TS$	$T,V,\{N_i\}$
Enthalpy	H	$U + pV$	$S,p,\{N_i\}$
Gibbs free energy	G	$U + pV - TS$	$T,p,\{N_i\}$
Landau Potential (Grand potential)	Ω, Φ_G	$U - TS - \sum_i \mu_i N_i$	$T,V,\{\mu_i\}$

where T is the temperature, S the entropy, p the pressure, V the volume, μ the chemical potential, N the number of particles in the system, and i is the count of particles types in the system.

Thermodynamic potentials can be derived from the energy balance equation applied to a thermodynamic system. Other thermodynamic potentials can also be obtained through Legendre transformation.

Applied fields

- Atmospheric thermodynamics
- Biological thermodynamics
- Black hole thermodynamics
- Chemical thermodynamics
- Classical thermodynamics
- Equilibrium thermodynamics
- Industrial ecology (re: Exergy)
- Maximum entropy thermodynamics
- Psychrometrics
- Quantum thermodynamics
- Statistical thermodynamics
- Thermoeconomics

Chapter- 5

Heat Transfer

Heat transfer is a discipline of thermal engineering that concerns the transfer of thermal energy from one physical system to another. Heat transfer is classified into various mechanisms, such as heat conduction, convection, thermal radiation, and phase-change transfer. Engineers also consider the transfer of mass of differing chemical species, either cold or hot, to achieve heat transfer.

Conduction, also called diffusion, is the direct microscopic exchange of kinetic energy of particles through the boundary between two systems. When an object is at a different temperature from another body or its surroundings, heat flows so that the body and the surroundings reach the same temperature at thermal equilibrium. Such spontaneous heat transfer always occurs from a region of high temperature to another region of lower temperature, as required by the second law of thermodynamics.

Transfer by thermal radiation is the transfer of energy by transmission of electromagnetic radiation described by black body theory.

Overview

Heat is defined in physics as the transfer of thermal energy across a well-defined boundary around a thermodynamic system. It is a characteristics of a process and is not statically contained in matter. In engineering contexts, however, the term *heat transfer* has acquired a specific usage, despite its literal redundancy of the characterization of transfer. In these contexts, *heat* is taken as synonymous to thermal energy. This usage has its origin in the historical interpretation of heat as a fluid (*caloric*) that can be transferred by various causes, and that is also common in the language of laymen and everyday life.

Fundamental methods of heat transfer in engineering include conduction, convection, and radiation. Physical laws describe the behavior and characteristics of each of these methods. Real systems often exhibit a complicated combination of them. Heat transfer methods are used in numerous disciplines, such as automotive engineering, thermal management of electronic devices and systems, climate control, Insulation, materials processing, and power plant engineering.

Various mathematical methods have been developed to solve or approximate the results of heat transfer in systems. Heat transfer is a path function (or process quantity), as opposed to a state quantity; therefore, the amount of heat transferred in a thermodynamic process that changes the state of a system depends on how that process occurs, not only the net difference between the initial and final states of the process. Heat flux is a quantitative, vectorial representation of the heat flow through a surface.

Heat transfer is typically studied as part of a general chemical engineering or mechanical engineering curriculum. Typically, thermodynamics is a prerequisite for heat transfer courses, as the laws of thermodynamics are essential to the mechanism of heat transfer. Other courses related to heat transfer include energy conversion, thermofluids, and mass transfer.

The transport equations for thermal energy (Fourier's law), mechanical momentum (Newton's law for fluids), and mass transfer (Fick's laws of diffusion) are similar and analogies among these three transport processes have been developed to facilitate prediction of conversion from any one to the others.

Mechanisms

The fundamental modes of heat transfer are:

Conduction or diffusion
: The transfer of energy between objects that are in physical contact

Convection
: The transfer of energy between an object and its environment, due to fluid motion

Radiation
: The transfer of energy to or from a body by means of the emission or absorption of electromagnetic radiation

Mass transfer
: The transfer of energy from one location to another as a side effect of physically moving an object containing that energy

Conduction

On a microscopic scale, heat conduction occurs as hot, rapidly moving or vibrating atoms and molecules interact with neighboring atoms and molecules, transferring some of their energy (heat) to these neighboring particles. In other words, heat is transferred by conduction when adjacent atoms vibrate against one another, or as electrons move from one atom to another. Conduction is the most significant means of heat transfer within a solid or between solid objects in thermal contact. Fluids—especially gases—are less conductive. Thermal contact conductance is the study of heat conduction between solid bodies in contact.

Steady state conduction is a form of conduction that happens when the temperature difference driving the conduction is constant, so that after an equilibration time, the spatial distribution of temperatures in the conducting object does not change any further.

In steady state conduction, the amount of heat entering a section is equal to amount of heat coming out.

Transient conduction occurs when the temperature within an object changes as a function of time. Analysis of transient systems is more complex and often calls for the application of approximation theories or numerical analysis by computer.

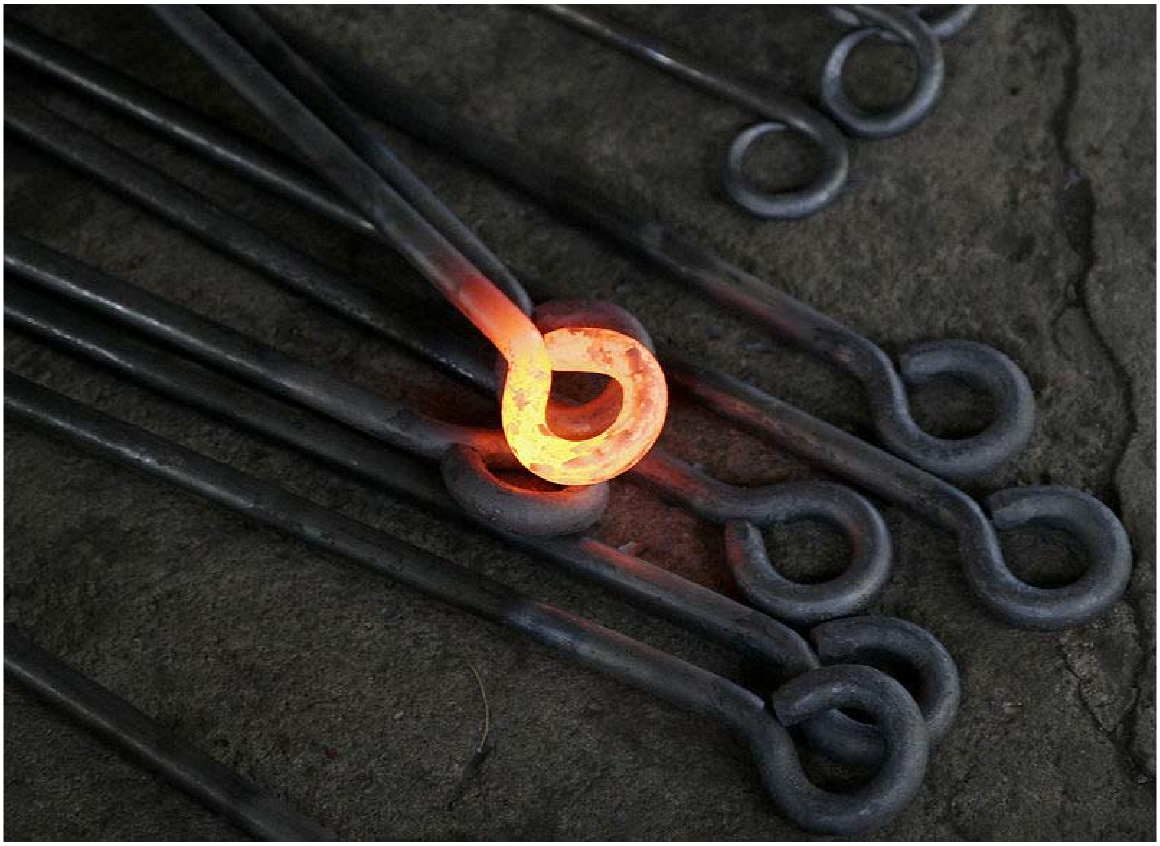
Convection

Convective heat transfer, or convection, is the transfer of heat from one place to another by the movement of fluids. (In physics, the term *fluid* means any substance that deforms under shear stress; it includes liquids, gases, plasmas, and some plastic solids.) Bulk motion of the fluid enhances the heat transfer between the solid surface and the fluid. Convection is usually the dominant form of heat transfer in liquids and gases. Although often discussed as a third method of heat transfer, convection actually describes the combined effects of conduction and fluid flow.

Free, or natural, convection occurs when the fluid motion is caused by buoyancy forces that result from density variations due to variations of temperature in the fluid. *Forced* convection is when the fluid is forced to flow over the surface by external means—such as fans, stirrers, and pumps—creating an artificially induced convection current.

Convection is described by Newton's law of cooling: "The rate of heat loss of a body is proportional to the difference in temperatures between the body and its surroundings."

Radiation



A red-hot iron object, transferring heat to the surrounding environment primarily through thermal radiation.

Thermal radiation is energy emitted by matter as electromagnetic waves due to the pool of thermal energy that all matter possesses that has a temperature above absolute zero. Thermal radiation propagates without the presence of matter through the vacuum of space.

Thermal radiation is a direct result of the random movements of atoms and molecules in matter. Since these atoms and molecules are composed of charged particles (protons and electrons), their movement results in the emission of electromagnetic radiation, which carries energy away from the surface.

Unlike conductive and convective forms of heat transfer, thermal radiation can be concentrated in a small spot by using reflecting mirrors, which is exploited in concentrating solar power generation. For example, the sunlight reflected from mirrors heats the PS10 solar power tower and during the day it can heat water to 285 °C (545 °F).

Mass Transfer

In mass transfer, energy—including thermal energy—is moved by the physical transfer of a hot or cold object from one place to another. This can be as simple as placing hot water in a bottle and heating a bed, or the movement of an iceberg in changing ocean currents. A practical example is thermal hydraulics.

Convection vs. conduction

In a body of fluid that is heated from underneath its container, conduction and convection can be considered to compete for dominance. If heat conduction is too great, fluid moving down by convection is heated by conduction so fast that its downward movement will be stopped due to its buoyancy, while fluid moving up by convection is cooled by conduction so fast that its driving buoyancy will diminish. On the other hand, if heat conduction is very low, a large temperature gradient may be formed and convection might be very strong.

The Rayleigh number (*Ra*) is a measure determining the result of this competition.

$$Ra = \frac{g\Delta\rho L^3}{\mu\alpha} = \frac{g\beta\Delta T L^3}{\nu\alpha}$$

where

- g is acceleration due to gravity
- ρ is the density with $\Delta\rho$ being the density difference between the lower and upper ends
- μ is the dynamic viscosity
- α is the Thermal diffusivity
- β is the volume thermal expansivity (sometimes denoted α elsewhere)
- T is the temperature and
- ν is the kinematic viscosity.

The Rayleigh number can be understood as the ratio between the rate of heat transfer by convection to the rate of heat transfer by conduction; or, equivalently, the ratio between the corresponding timescales (i.e. conduction timescale divided by convection timescale), up to a numerical factor. This can be seen as follows, where all calculations are up to numerical factors depending on the geometry of the system.

The buoyancy force driving the convection is roughly $g\Delta\rho L^3$, so the corresponding pressure is roughly $g\Delta\rho L$. In steady state, this is canceled by the shear stress due to viscosity, and therefore roughly equals $\mu V / L = \mu / T_{conv}$, where V is the typical fluid velocity due to convection and T_{conv} the order of its timescale. The conduction timescale, on the other hand, is of the order of $T_{cond} = L^2 / \alpha$.

Convection occurs when the Rayleigh number is above 1,000–2,000. For example, the Earth's mantle, exhibiting non-stable convection, has Rayleigh number of the order of 1,000, and T_{conv} as calculated above is around 100 million years.

Phase changes

Transfer of heat through a phase transition in the medium—such as water-to-ice, water-to-steam, steam-to-water, or ice-to-water—involves significant energy and is exploited in many ways: steam engines, refrigerators, etc. For example, the Mason equation is an approximate analytical expression for the growth of a water droplet based on the effects of heat transport on evaporation and condensation.

Boiling

Heat transfer in boiling fluids is complex, but of considerable technical importance. It is characterized by an *S*-shaped curve relating heat flux to surface temperature difference.

At low driving temperatures, no boiling occurs and the heat transfer rate is controlled by the usual single-phase mechanisms. As the surface temperature is increased, local boiling occurs and vapor bubbles nucleate, grow into the surrounding cooler fluid, and collapse. This is *sub-cooled nucleate boiling*, and is a very efficient heat transfer mechanism. At high bubble generation rates, the bubbles begin to interfere and the heat flux no longer increases rapidly with surface temperature (this is the departure from nucleate boiling, or DNB). At higher temperatures still, a maximum in the heat flux is reached (the critical heat flux, or CHF). The regime of falling heat transfer that follows is not easy to study, but is believed to be characterized by alternate periods of nucleate and film boiling. Nucleate boiling slows the heat transfer due to gas bubbles on the heater's surface; as mentioned, gas-phase thermal conductivity is much lower than liquid-phase thermal conductivity, so the outcome is a kind of "gas thermal barrier".

At higher temperatures still, the hydrodynamically-quieter regime of film boiling is reached. Heat fluxes across the stable vapor layers are low, but rise slowly with temperature. Any contact between fluid and the surface that may be seen probably leads to the extremely rapid nucleation of a fresh vapor layer ("spontaneous nucleation").

Condensation

Condensation occurs when a vapor is cooled and changes its phase to a liquid. Condensation heat transfer, like boiling, is of great significance in industry. During condensation, the latent heat of vaporization must be released. The amount of the heat is the same as that absorbed during vaporization at the same fluid pressure.

There are several types of condensation:

- Homogeneous condensation, as during a formation of fog.
- Condensation in direct contact with subcooled liquid.

- Condensation on direct contact with a cooling wall of a heat exchanger: This is the most common mode used in industry:
    - Filmwise condensation is when a liquid film is formed on the subcooled surface, and usually occurs when the liquid wets the surface.
    - Dropwise condensation is when liquid drops are formed on the subcooled surface, and usually occurs when the liquid does not wet the surface.

  Dropwise condensation is difficult to sustain reliably; therefore, industrial equipment is normally designed to operate in filmwise condensation mode.

Modeling approaches

Complex heat transfer phenomena can be modeled in different ways.

Heat equation

The heat equation is an important partial differential equation that describes the distribution of heat (or variation in temperature) in a given region over time. In some cases, exact solutions of the equation are available; in other cases the equation must be solved numerically using computational methods. For example, simplified climate models may use Newtonian cooling, instead of a full (and computationally expensive) radiation code, to maintain atmospheric temperatures.

Lumped system analysis

System analysis by the lumped capacitance model is a common approximation in transient conduction that may be used whenever heat conduction within an object is much faster than heat conduction across the boundary of the object.

This is a method of approximation that reduces one aspect of the transient conduction system—that within the object—to an equivalent steady state system. That is, the method assumes that the temperature within the object is completely uniform, although its value may be changing in time.

In this method, the ratio of the conductive heat resistance within the object to the convective heat transfer resistance across the object's boundary, known as the *Biot number*, is calculated. For small Biot numbers, the approximation of *spatially uniform temperature within the object* can be used: it can be presumed that heat transferred into the object has time to uniformly distribute itself, due to the lower resistance to doing so, as compared with the resistance to heat entering the object.

Applications and techniques

Heat transfer has broad application to the functioning of numerous devices and systems. Heat-transfer principles may be used to preserve, increase, or decrease temperature in a wide variety of circumstances.

Insulation and radiant barriers

Heat exposure as part of a fire test for firestop products

Thermal insulators are materials specifically designed to reduce the flow of heat by limiting conduction, convection, or both. Radiant barriers are materials that reflect radiation, and therefore reduce the flow of heat from radiation sources. Good insulators are not necessarily good radiant barriers, and vice versa. Metal, for instance, is an excellent reflector and a poor insulator.

The effectiveness of an insulator is indicated by its **R-value**, or resistance value. The R-value of a material is the inverse of the conduction coefficient (*k*) multiplied by the thickness (*d*) of the insulator. In most of the world, R-values are measured in SI units: square-meter kelvins per watt ($m^2 \cdot K/W$). In the United States, R-values are customarily given in units of British thermal units per hour per square-foot degrees Fahrenheit ($Btu/h \cdot ft^2 \cdot °F$).

$$R = \frac{d}{k}$$

$$C = \frac{Q}{m\Delta T}$$

Rigid fiberglass, a common insulation material, has an R-value of four per inch, while poured concrete, a poor insulator, has an R-value of 0.08 per inch.

The tog is a measure of thermal resistance, commonly used in the textile industry, and often seen quoted on, for example, duvets and carpet underlay.

The effectiveness of a radiant barrier is indicated by its **reflectivity**, which is the fraction of radiation reflected. A material with a high reflectivity (at a given wavelength) has a low emissivity (at that same wavelength), and vice versa. At any specific wavelength, reflectivity = 1 - emissivity. An ideal radiant barrier would have a reflectivity of 1, and would therefore reflect 100 percent of incoming radiation. Vacuum flasks, or Dewars, are silvered to approach this ideal. In the vacuum of space, satellites use multi-layer insulation, which consists of many layers of aluminized (shiny) Mylar to greatly reduce radiation heat transfer and control satellite temperature.

Critical insulation thickness

Low thermal conductivity (*k*) materials reduce heat fluxes. The smaller the *k* value, the larger the corresponding thermal resistance (*R*) value. Thermal conductivity is measured in watts-per-meter per kelvin ($W \cdot m^{-1} \cdot K^{-1}$), represented as *k*. As the thickness of insulating material increases, the thermal resistance—or R-value—also increases.

However, adding layers of insulation has the potential of increasing the surface area, and hence the thermal convection area.

For example, as thicker insulation is added to a cylindrical pipe, the outer radius of the pipe-and-insulation system increases, and therefore surface area increases. The point where the added resistance of increasing insulation thickness becomes overshadowed by the effect of increased surface area is called the critical insulation thickness. In simple cylindrical pipes, this is calculated as a radius:

$$R_{critical} = \frac{k}{h}$$

Heat exchangers

A heat exchanger is a tool built for efficient heat transfer from one fluid to another, whether the fluids are separated by a solid wall so that they never mix, or the fluids are in direct contact. Heat exchangers are widely used in refrigeration, air conditioning, space heating, power generation, and chemical processing. One common example of a heat exchanger is a car's radiator, in which the hot coolant fluid is cooled by the flow of air over the radiator's surface.

Common types of heat exchanger flows include parallel flow, counter flow, and cross flow. In parallel flow, both fluids move in the same direction while transferring heat; in counter flow, the fluids move in opposite directions; and in cross flow, the fluids move at right angles to each other. Common constructions for heat exchanger include shell and tube, double pipe, extruded finned pipe, spiral fin pipe, u-tube, and stacked plate.

When engineers calculate the theoretical heat transfer in a heat exchanger, they must contend with the fact that the driving temperature difference between the two fluids varies with position. To account for this in simple systems, the log mean temperature difference (LMTD) is often used as an "average" temperature. In more complex systems, direct knowledge of the LMTD is not available, and the number of transfer units (NTU) method can be used instead.

Heat dissipation

A heat sink is a component that transfers heat generated within a solid material to a fluid medium, such as air or a liquid. Examples of heat sinks are the heat exchangers used in refrigeration and air conditioning systems, and the radiator in a car (which is also a heat exchanger). Heat sinks also help to cool electronic and optoelectronic devices such as CPUs, higher-power lasers, and light-emitting diodes (LEDs). A heat sink uses its extended surfaces to increase the surface area in contact with the cooling fluid.

Buildings

In cold climates, houses with their heating systems form dissipative systems. In spite of efforts to insulate houses to reduce heat losses via their exteriors, considerable heat is lost, which can make their interiors uncomfortably cool or cold. For the comfort of the inhabitants, the interiors must be maintained out of thermal equilibrium with the external surroundings. In effect, these domestic residences are oases of warmth in a sea of cold, and the thermal gradient between the inside and outside is often quite steep. This can lead to problems such as condensation and uncomfortable air currents, which—if left unaddressed—can cause cosmetic or structural damage to the property. Such issues can be prevented by use of insulation techniques for reducing heat loss.

Thermal transmittance is the rate of transfer of heat through a structure divided by the difference in temperature across the structure. It is expressed in watts per square meter per kelvin, or W/m^2K. Well-insulated parts of a building have a low thermal transmittance, whereas poorly-insulated parts of a building have a high thermal transmittance.

A thermostat is a device capable of starting the heating system when the house's interior falls below a set temperature, and of stopping that same system when another (higher) set temperature has been achieved. Thus, the thermostat controls the flow of energy into the house, that energy eventually being dissipated to the exterior.

Thermal energy storage

Thermal energy storage refers to technologies that store energy in a thermal reservoir for later use. They can be employed to balance energy demand between daytime and nighttime. The thermal reservoir may be maintained at a temperature above (hotter) or below (colder) than that of the ambient environment. Applications include later use in space heating, domestic or process hot water, or to generate electricity. Most practical active solar heating systems have storage for a few hours to a day's worth of heat collected.

Evaporative cooling

Evaporative cooling is a physical phenomenon in which evaporation of a liquid, typically into surrounding air, cools an object or a liquid in contact with it. Latent heat describes the amount of heat that is needed to evaporate the liquid; this heat comes from the liquid itself and the surrounding gas and surfaces. The greater the difference between the two temperatures, the greater the evaporative cooling effect. When the temperatures are the same, no net evaporation of water in air occurs; thus, there is no cooling effect. A simple example of natural evaporative cooling is perspiration, or sweat, which the body secretes in order to cool itself. An evaporative cooler is a device that cools air through the simple evaporation of water.

Radiative cooling

Radiative cooling is the process by which a body loses heat by radiation. It is an important effect in the Earth's atmosphere. In the case of the Earth-atmosphere system, it refers to the process by which long-wave (infrared) radiation is emitted to balance the absorption of short-wave (visible) energy from the Sun. Convective transport of heat and evaporative transport of latent heat both remove heat from the surface and redistribute it in the atmosphere, making it available for radiative transport at higher altitudes.

Laser cooling

Laser cooling refers to techniques in which atomic and molecular samples are cooled through the interaction with one or more laser light fields. The most common method of laser cooling is Doppler cooling. In Doppler cooling, the frequency of the laser light is tuned slightly below an electronic transition in the atom. Thus, the atoms would absorb more photons if they moved towards the light source, due to the Doppler effect. If an excited atom then emits a photon spontaneously, it will be accelerated. The result of the absorption and emission process is to reduce the speed of the atom. Eventually the mean velocity, and therefore the kinetic energy of the atoms, will be reduced. Since the temperature of an ensemble of atoms is a measure of the random internal kinetic energy, this is equivalent to cooling the atoms.

Sympathetic cooling is a process in which particles of one type cool particles of another type. Typically, atomic ions that can be directly laser-cooled are used to cool nearby ions or atoms. This technique allows cooling of ions and atoms that cannot be laser cooled directly.

Magnetic cooling

Magnetic evaporative cooling is a technique for lowering the temperature of a group of atoms. The process confines atoms using a magnetic field. Over time, individual atoms will become much more energetic than the others due to random collisions, and will escape—removing energy from the system and reducing the temperature of the remaining group. This process is similar to the familiar process by which standing water becomes water vapor.

Other

A heat pipe is a passive device constructed in such a way that it acts as though it has extremely high thermal conductivity. Heat pipes use latent heat and capillary action to move heat, and can carry many times as much heat as a similar-sized copper rod. Originally invented for use in satellites, they have applications in personal computers.

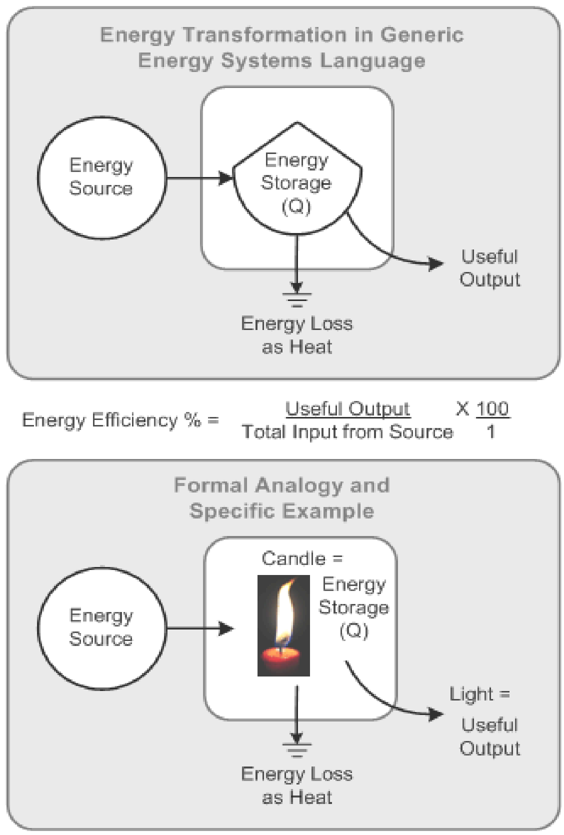
A thermocouple is a junction between two different metals that produces a voltage related to a temperature difference. Thermocouples are a widely used type of temperature sensor for measurement and control, and can also be used to convert heat into electric power.

A thermopile is an electronic device that converts thermal energy into electrical energy. It is composed of thermocouples. Thermopiles do not measure the absolute temperature, but generate an output voltage proportional to a temperature difference. Thermopiles are widely used, e.g., they are the key component of infrared thermometers, such as those used to measure body temperature via the ear.

A thermal diode or thermal rectifier is a device that preferentially passes heat in one direction: a "one-way valve" for heat.

Chapter- 6

Energy Transformation



Energy Transformation in Energy Systems Language

In physics, the term energy describes the capacity to produce changes within a system, without regard to limitations in transformation imposed by entropy. Changes in total energy of systems can only be accomplished by adding or subtracting energy from them, as energy is a quantity which is conserved, according to the first law of thermodynamics. According to special relativity, changes in the energy of systems will also coincide with changes in the system's mass, and the total amount of mass of a system is a measure of its energy.

Energy in a system may be **transformed** so that it resides in a different state. Energy in many states may be used to do many varieties of physical work. Energy may be used in natural processes or machines, or else to provide some service to society (such as heat, light, or motion). For example, an internal combustion engine converts the potential chemical energy in gasoline and oxygen into heat, which is then transformed into the propulsive energy (kinetic energy that moves a vehicle.) A solar cell converts solar radiation into electrical energy that can then be used to light a bulb or power a computer.

The generic name for a device which converts energy from one form to another is a transducer.

In general, most types of energy, save for thermal energy, may be converted to any other kind of energy, with a theoretical efficiency of 100%. Such efficiencies might occur in practice, such as when chemical potential energy is completely converted into kinetic energies, and vice versa, only in isolated systems.

Conversion of other types of energies to heat also may occur with high efficiency but a perfect level would be only possible for isolated systems also.

If there is nothing beyond the frontiers of the universe then the only real isolated system would be the universe itself. Currently we do not have the knowledge or technology to create an isolated system from a portion of the universe.

Exceptions for perfect efficiency (even for isolated systems) occur when energy has already been partly distributed among many available quantum states for a collection of particles, which are freely allowed to explore any state of momentum and position (phase space). In such circumstances, a measure called entropy, or evening-out of energy distribution in such states, dictates that future states of the system must be of at least equal evenness in energy distribution. (There is no way, taking the universe as a whole, to collect energy into fewer states, once it has spread to them).

A consequence of this requirement is that there are limitations to the efficiency with which thermal energy can be converted to other kinds of energy, since thermal energy in equilibrium at a given temperature already represents the maximal evening-out of energy between all possible states. Such energy is sometimes considered "degraded energy," because it is not entirely usable. The second law of thermodynamics is a way of stating that, for this reason, thermal energy in a system may be converted to other kinds of energy with efficiencies approaching 100%, only if the entropy (even-ness or disorder) of the universe is increased by other means, to compensate for the decrease in entropy associated with the disappearance of the thermal energy and its entropy content. Otherwise, only a part of thermal energy may be converted to other kinds of energy (and thus, useful work), since the remainder of the heat must be reserved to be transferred to a thermal reservoir at a lower temperature, in such a way that the increase in entropy for this process more than compensates for the entropy decrease associated with transformation of the rest of the heat into other types of energy.

History of energy transformation from the early universe

Energy transformations in the universe over time are (generally) characterized by various kinds of energy which has been available since the Big Bang, later being "released" (that is, transformed to more active types of energy such as kinetic or radiant energy), when a triggering mechanism is available to do it. A direct transformation of energy occurs when hydrogen produced in the big bang collects into structures such as planets, in a process during which gravitational potential may be converted directly into heat. In Jupiter, Saturn, Uranus, and Neptune, for example, such heat from continued collapse of the planets' large gases atmospheres continues to drive most of the planets' weather systems, with atmospheric bands, winds, and powerful storms.

Familiar examples of other such processes transforming energy from the big bang include nuclear decay, in which energy is released which was originally "stored" in heavy isotopes, such as uranium and thorium. This energy was stored at the time of these elements' nucleosynthesis, a process which ultimately uses the gravitational potential energy released from the gravitational collapse of supernovae, to store energy in the creation of these heavy elements before they were incorporated into the solar system and the Earth. This energy in uranium is triggered for sudden-release in nuclear fission bombs, and similar stored energies in atomic nuclei are released spontaneously, during most types of radioactive decay. In this process, heat from decay of these atoms in the core of the Earth is transformed immediately to heat. This heat in turn may lift mountains, via plate tectonics and orogenesis. This slow lifting of terrain thus represents a kind of gravitational potential energy storage of the heat energy. The stored potential energy may be released to active kinetic energy in landslides, after a triggering event. Earthquakes also release stored elastic potential energy in rocks, a kind of mechanical potential energy which has been produced ultimately from the same radioactive heat sources. Thus, according to present understanding, familiar events such as landslides and earthquakes release energy which has been stored as potential energy in the Earth's gravitational field, or elastic strain (mechanical potential energy) in rocks. Prior to this, the energy represented by these events had been stored in heavy atoms, ever since the time that gravitational potentials transforming energy in the collapse of long-destroyed stars created these atoms, and in doing so, stored the energy within them.

In other similar chain of transformations beginning at the dawn of the universe, nuclear fusion of hydrogen in the Sun releases another store of potential energy which was created at the time of the Big Bang. At that time, according to theory, space expanded and the universe cooled too rapidly for hydrogen to completely fuse into heavier elements. This meant that hydrogen represents a store of potential energy which can be released by nuclear fusion. Such a fusion process is triggered by heat and pressure generated from gravitational collapse of hydrogen clouds when they produce stars, and some of the fusion energy is then transformed into sunlight. Such sunlight may again be stored as gravitational potential energy after it strikes the Earth, as (for example) snow-avalanches, or when water evaporates from oceans and is deposited high above sea level (where, after being released at a hydroelectric dam, it can be used to drive turbine/generators to produce electricity). Sunlight also drives many weather phenomena on Earth. An example of a solar-mediated weather event is a hurricane, which occurs when large unstable areas of warm ocean, heated over months, give up some of their thermal energy suddenly to power a few days of violent air movement. Sunlight is also captured by plants as *chemical potential energy*, when carbon dioxide and water are converted into a combustible combination of carbohydrates, lipids, and oxygen. Release of this energy as heat and light may be triggered suddenly by a spark, in a forest fire; or it may be available more slowly for animal or human metabolism, when these molecules are ingested, and catabolism is triggered by enzyme action.

Through all of these transformation chains, potential energy stored at the time of the Big Bang is later released by intermediate events, sometimes being stored in a number of ways over time between releases, as more active energy. In all these events, one kind of energy is converted to other types of energy, including heat.

Examples of sets of energy conversions in machines

For instance, a coal-fired power plant involves these power transfers:

1. Chemical energy in the coal converted to thermal energy
2. Thermal energy converted to kinetic energy in steam
3. Kinetic energy converted to mechanical energy in the turbine
4. Mechanical energy of the turbine converted to electrical energy, which is the ultimate output

In such a system, the last step is almost perfectly efficient, the first and second steps are fairly efficient, but the third step is relatively inefficient. The most efficient gas-fired electrical power stations can achieve 50% conversion efficiency. Oil and coal fired stations achieve less.

In a conventional automobile, these power transfers are involved:

1. Potential energy in the fuel converted to kinetic energy of expanding gas via combustion
2. Kinetic energy of expanding gas converted to linear piston movement
3. Linear piston movement converted to rotary crankshaft movement
4. Rotary crankshaft movement passed into transmission assembly
5. Rotary movement passed out of transmission assembly
6. Rotary movement passed through differential
7. Rotary movement passed out of differential to drive wheels
8. Rotary movement of drive wheels converted to linear motion of the vehicle.

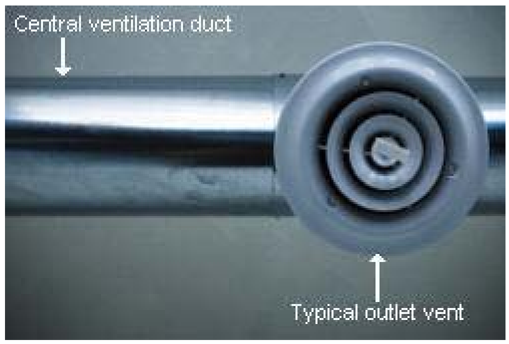
Other energy conversions

There are many different machines and transducers that convert one energy form into another. A short list of examples follows:

- Thermoelectric (Heat → Electricity)
- Geothermal power (Heat→ Electricity)
- Heat engines, such as the internal combustion engine used in cars, or the steam engine (Heat → Mechanical energy)
- Ocean thermal power (Heat → Electricity)
- Hydroelectric dams (Gravitational potential energy → Electricity)
- Electric generator (Kinetic energy or Mechanical work → Electricity)
- Fuel cells (Chemical energy → Electricity)
- Battery (electricity) (Chemical energy → Electricity)
- Fire (Chemical energy → Heat and Light)
- Electric lamp (Electricity → Heat and Light)
- Microphone (Sound → Electricity)
- Wave power (Mechanical energy → Electricity)
- Windmills (Wind energy → Electricity or Mechanical energy)
- Piezoelectrics (Strain → Electricity)
- Acoustoelectrics (Sound → Electricity)
- Friction (Kinetic energy → Heat)

Chapter- 7

HVAC

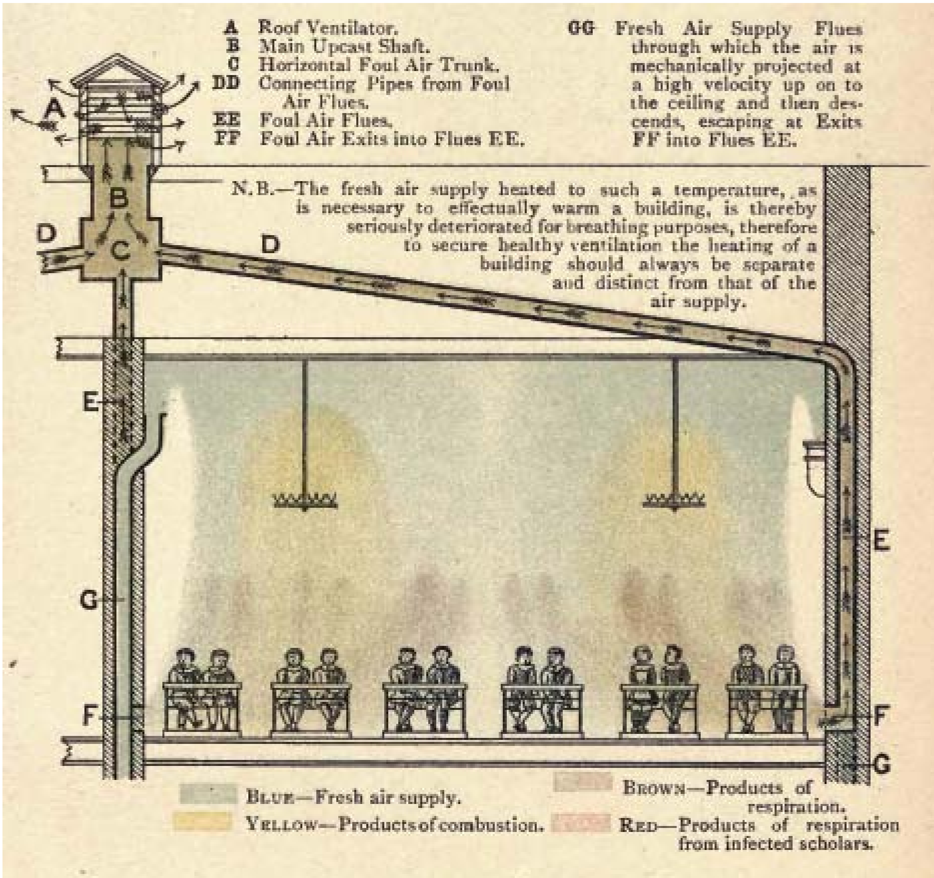


HVAC systems use ventilation air ducts installed throughout a building that supply conditioned air to a room through rectangular or round outlet vents, called diffusers; and ducts that remove air through return-air grilles

HVAC (Heating, Ventilating, and Air Conditioning) refers to technology of indoor or automotive environmental comfort. HVAC system design is a major subdiscipline of mechanical engineering, based on the principles of thermodynamics, fluid mechanics, and heat transfer. Refrigeration is sometimes added to the field's abbreviation as HVAC&R or HVACR, or ventilating is dropped as in HACR (such as the designation of HACR-rated circuit breakers).

HVAC is important in the design of medium to large industrial and office buildings such as skyscrapers and in marine environments such as aquariums, where safe and healthy building conditions are regulated with temperature and humidity, as well as "fresh air" from outdoors.

Background



Ventilation (architecture) on the downdraught system, by impulsion, or the 'plenum' principle, applied to schoolrooms (1899)

Heating, ventilating, and air conditioning is based on inventions and discoveries made by Nikolay Lvov, Michael Faraday, Willis Carrier, Reuben Trane, James Joule, William Rankine, Sadi Carnot, and many others.

The invention of the components of HVAC systems went hand-in-hand with the industrial revolution, and new methods of modernization, higher efficiency, and system control are constantly introduced by companies and inventors all over the world. The three central functions of heating, ventilating, and air-conditioning are interrelated, providing thermal comfort, acceptable indoor air quality, within reasonable installation, operation, and maintenance costs. HVAC systems can provide ventilation, reduce air infiltration, and maintain pressure relationships between spaces. How air is delivered to, and removed from spaces is known as room air distribution.

In modern buildings the design, installation, and control systems of these functions are integrated into one or more HVAC systems. For very small buildings, contractors normally "size" and select HVAC systems and equipment. For larger buildings, building services designers and engineers, such as mechanical, architectural, or building services engineers analyze, design, and specify the HVAC systems, and specialty mechanical contractors build and commission them. Building permits and code-compliance inspections of the installations are normally required for all sizes of buildings.

The HVAC industry is a worldwide enterprise, with career opportunities including operation and maintenance, system design and construction, equipment manufacturing and sales, and in education and research. The HVAC industry had been historically regulated by the manufacturers of HVAC equipment, but Regulating and Standards organizations such as HARDI, ASHRAE, SMACNA, ACCA, Uniform Mechanical Code, International Mechanical Code, and AMCA have been established to support the industry and encourage high standards and achievement.

Design of the HVAC system.

The starting point in carrying out a heat estimate both for cooling and heating will depends on the ambient and inside conditions specified. However before taking up the heat load calculation, it is necessary to work out the fresh air requirement for each area in details, as pressurization is an important requirement.

Heating

Central heating unit

There are many different types of standard heating systems. Central heating is often used in cold climates to heat private houses and public buildings. Such a system contains a boiler, furnace, or heat pump to heat water, steam, or air, all in a central location such as a furnace room in a home or a mechanical room in a large building. The use of water as the heat transfer medium is known as hydronics. The system also contains either ductwork, for forced air systems, or piping to distribute a heated fluid and radiators to transfer this heat to the air. The term *radiator* in this context is misleading since most heat transfer from the heat exchanger is by convection, not radiation. The radiators may be mounted on walls or buried in the floor to give under-floor heat.

In boiler fed or radiant heating systems, all but the simplest systems have a pump to circulate the water and ensure an equal supply of heat to all the radiators. The heated water can also be fed through another (secondary) heat exchanger inside a storage cylinder to provide hot running water.

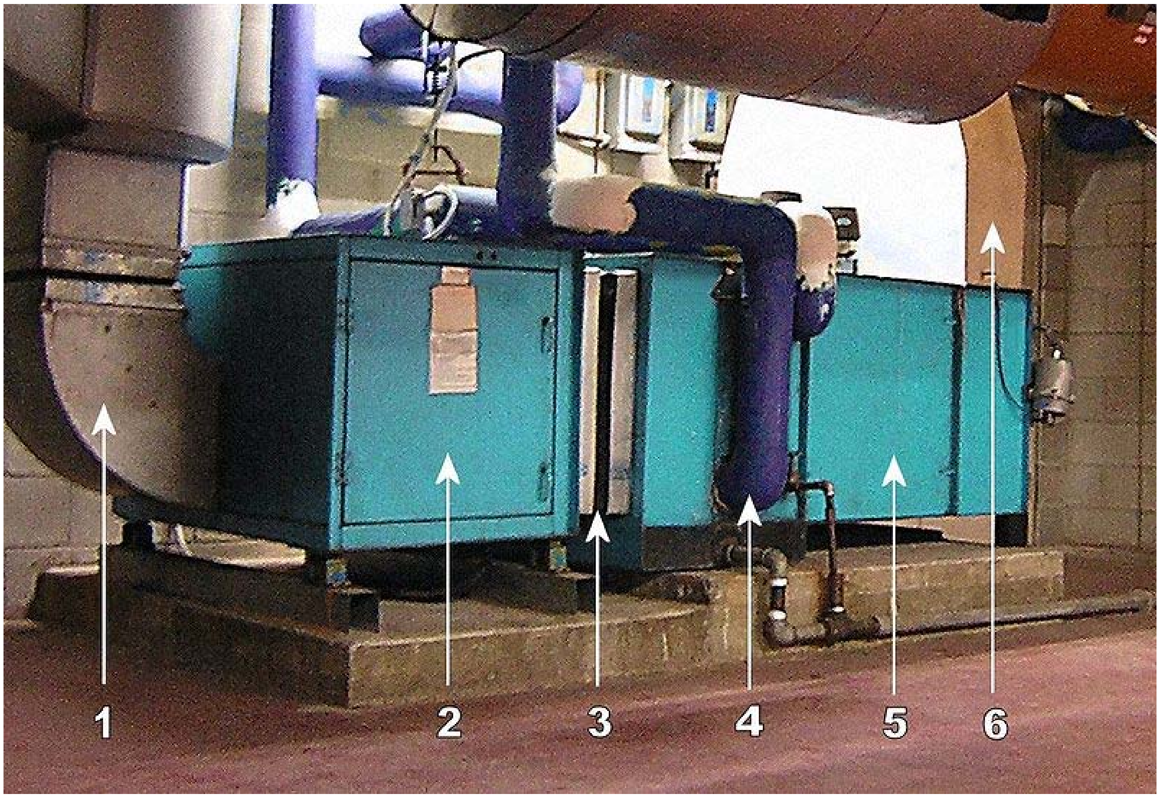
Forced air systems send heated air through ductwork. During warm weather the same ductwork can be used for air conditioning. The forced air can also be filtered or put through air cleaners.

Heating can also be provided from electric, or resistance heating using a filament that becomes hot when electric current is caused to pass through it. This type of heat can be found in electric baseboard heaters, portable electric heaters, and as backup or supplemental heating for heat pump (or reverse heating) system.

The heating elements (radiators or vents) should be located in the coldest part of the room, typically next to the windows to minimize condensation and offset the convective air current formed in the room due to the air next to the window becoming negatively buoyant due to the cold glass. Devices that direct vents away from windows to prevent "wasted" heat defeat this design intent. Cold air drafts can contribute significantly to subjectively feeling colder than the average room temperature. Therefore, it is important to control the air leaks from outside in addition to proper design of the **heating system**.

The invention of central heating is often credited to the ancient Romans, who installed a system of air ducts called a hypocaust in the walls and floors of public baths and private villas.

Ventilating



An air handling unit is used for the heating and cooling of air in a central location.

Ventilating is the process of "changing" or replacing air in any space to control temperature or remove moisture, odors, smoke, heat, dust, airborne bacteria, carbon dioxide, and to replenish oxygen. Ventilation includes both the exchange of air to the outside as well as circulation of air within the building. It is one of the most important factors for maintaining acceptable indoor air quality in buildings. Methods for ventilating a building may be divided into *mechanical/forced* and *natural* types. Ventilation is used to remove unpleasant smells and excessive moisture, introduce outside air, to keep interior building air circulating, and to prevent stagnation of the interior air.

Mechanical or forced ventilation

"Mechanical" or "forced" ventilation is provided by an air handler and used to control indoor air quality. Excess humidity, odors, and contaminants can often be controlled via dilution or replacement with outside air. However, in humid climates much energy is required to remove excess moisture from ventilation air.

Kitchens and bathrooms typically have mechanical exhaust to control odors and sometimes humidity. Factors in the design of such systems include the flow rate (which is a function of the fan speed and exhaust vent size) and noise level. If ducting for the fans traverse unheated space (e.g., an attic), the ducting should be insulated as well to prevent condensation on the ducting. Direct drive fans are available for many applications, and can reduce maintenance needs.

Ceiling fans and table/floor fans circulate air within a room for the purpose of reducing the perceived temperature because of evaporation of perspiration on the skin of the occupants. Because hot air rises, ceiling fans may be used to keep a room warmer in the winter by circulating the warm stratified air from the ceiling to the floor. Ceiling fans do not provide ventilation as defined as the introduction of outside air.

Natural ventilation

Natural ventilation is the ventilation of a building with outside air without the use of a fan or other mechanical system. It can be achieved with openable windows or trickle vents when the spaces to ventilate are small and the architecture permits. In more complex systems warm air in the building can be allowed to rise and flow out upper openings to the outside (stack effect) thus forcing cool outside air to be drawn into the building naturally through openings in the lower areas. These systems use very little energy but care must be taken to ensure the occupants' comfort. In warm or humid months, in many climates, maintaining thermal comfort solely via natural ventilation may not be possible so conventional air conditioning systems are used as backups. Air-side economizers perform the same function as natural ventilation, but use mechanical systems' fans, ducts, dampers, and control systems to introduce and distribute cool outdoor air when appropriate.

Air conditioning

Air conditioning and refrigeration are provided through the removal of heat. The definition of cold is the absence of heat and all air conditioning systems work on this basic principle. Heat can be removed through the process of radiation, convection, and Heat cooling through a process called the refrigeration cycle. The conduction mediums such as water, air, ice, and chemicals are referred to as refrigerants.

An air conditioning system, or a standalone air conditioner, provides cooling, ventilation, and humidity control for all or part of a house or building.

The refrigerant cycle consists of four essential elements to create a cooling effect. The system refrigerant starts its cycle in a gaseous state. The compressor pumps the refrigerant gas up to a high pressure and temperature. From there it enters a heat exchanger (sometimes called a "condensing coil") where it loses energy (heat) to the outside. In the process the refrigerant condenses into a liquid. The liquid refrigerant is returned indoors to another heat exchanger ("evaporating coil"). A metering device allows the liquid to flow in at a low pressure at the proper rate. As the liquid refrigerant evaporates it aborbs energy (heat) from the inside air, returns to the compressor, and the cycle repeats. In the process, heat is absorbed from indoors, and transferred outdoors, resulting in cooling of the building.

Central, 'all-air' air conditioning systems are often installed in modern residences, offices, and public buildings, but are difficult to retrofit (install in a building that was not designed to receive it) because of the bulky air ducts required. A duct system must be carefully maintained to prevent the growth of pathogenic bacteria in the ducts. An alternative to large ducts to carry the needed air to heat or cool an area is the use of remote fan coils or split systems. These systems, although most often seen in residential applications, are gaining popularity in small commercial buildings. The evaporator coil is connected to a remote condenser unit using piping instead of ducts.

Dehumidification in an air conditioning system is provided by the evaporator. Since the evaporator operates at a temperature below dew point, moisture in the air condenses on the evaporator coil tubes. This moisture is collected at the bottom of the evaporator in a condensate pan and is removed by piping it to a central drain or onto the ground outside. A dehumidifier is an air-conditioner-like device that controls the humidity of a room or building. It is often employed in basements which have a higher relative humidity because of their lower temperature (and propensity for damp floors and walls). In food retailing establishments, large open chiller cabinets are highly effective at dehumidifying the internal air. Conversely, a humidifier increases the humidity of a building.

Air-conditioned buildings often have sealed windows, because open windows would disrupt the attempts of the HVAC system to maintain constant indoor air conditions.

All modern air conditioning systems, down to small "window" units, are equipped with internal air filters. These are generally of a light weight gauze-type element, and must be replaced as conditions warrant (some models may be washable). For example, a building in a high-dust environment, or a home with furry pets, will need to have the filters changed more often than buildings without these dirt loads. Failure to replace these filters as needed will contribute to a lower heat-exchange rate, resulting in wasted energy, shortened equipment life, and higher energy bills; also low air flow can result in "iced-up" or "iced-over" evaporator coils, and then there is no air flow at all. Additionally, very dirty or plugged filters can cause overheating during a heating cycle, and can possibly result in damage to the furnace unit or even fire.

It is important to keep in mind that because an air conditioner moves heat from the indoor (evaporator) coil to the outdoor (condenser) coil, the latter must be kept just as clean as the former. This means that, in addition to replacing the air filter at the evaporator coil, it is also necessary to regularly clean the condenser coil. Failure to keep the condenser clean will eventually result in harm to the compressor, because the condenser coil is responsible for discharging both the indoor heat (as picked up by the evaporator) plus the heat generated by the electric motor driving the compressor.

Outside, "fresh" air is generally drawn into the system by a vent into the evaporator section. Adjustment of the percentage of return air made up of fresh air can usually be adjusted by manipulating the opening of this vent.

Energy efficiency

For the last 20 to 30 years, manufacturers of HVAC equipment have been making an effort to make the systems they manufacture more efficient. This was originally driven by rising energy costs, and has more recently been driven by increased awareness of environmental issues. In the USA, the EPA has also imposed tighter restrictions. There are several methods for making HVAC systems more efficient.

Heating energy

Water heating is more efficient for heating buildings and was the standard many years ago. Today forced air systems can double for air conditioning and are more popular.

A couple of benefits of forced air systems, which are now widely applied in churches, schools and high-end residences,are 1) better air conditioned effect 2) up to 15-20% energy saving, and 3) evenly conditioned effect. A drawback is the installation cost, which might be slightly higher than traditional HVAC system.

Energy efficiency can be improved even more in central heating systems by introducing zoned heating. This allows a more granular application of heat, similar to non-central heating systems. Zones are controlled by multiple thermostats. In water heating systems the thermostats control zone valves, and in forced air systems they control zone dampers inside the vents which selectively block the flow of air. In this case, the control system is very critical to maintain a proper temperature.

Geothermal Heat Pump

Geothermal heat pumps are similar to ordinary heat pumps, but instead of using heat found in outside air, they rely on the stable, even heat of the earth to provide heating, air conditioning and, in most cases, hot water. From Montana's −70 °F (−57 °C) temperature, to the highest temperature ever recorded in the U.S.—134 °F (56.7 °C) in Death Valley, California, in 1913—many parts of the country experience seasonal temperature extremes. A few feet below the earth's surface, however, the ground remains at a relatively constant temperature. Although the temperatures vary according to latitude, at 6 feet (1.83 m) underground, temperatures range from 45 to 75 °F (7.2 to 23.9 °C).

While they may be more costly to install initially than regular heat pumps, they can produce markedly lower energy bills—30 percent to 40 percent lower, according to estimates from the U.S. Environmental Protection Agency.

Ventilation energy recovery

Energy recovery systems sometimes utilize heat recovery ventilation or energy recovery ventilation systems that employ heat exchangers or enthalpy wheels to recover sensible or latent heat from exhausted air. This is done by transfer of energy to the incoming outside fresh air.

Air conditioning energy

The performance of vapor compression refrigeration cycles is limited by thermodynamics. These air conditioning and heat pump devices *move* heat rather than convert it from one form to another, so *thermal efficiencies* do not appropriately describe the performance of these devices. The **Coefficient-of-Performance** (COP) measures performance, but this dimensionless measure has not been adopted, but rather the **Energy Efficiency Ratio** (*EER*). EER is the Energy Efficiency Ratio based on a 35 °C (95 °F) outdoor temperature. To more accurately describe the performance of air conditioning equipment over a typical cooling season a modified version of the EER is used, and is the **Seasonal Energy Efficiency Ratio** (*SEER*). SEER ratings are based on seasonal temperature averages instead of a constant 35 °C outdoor temperature. The current industry minimum SEER rating is 13 SEER. The SEER article describes it further, and presents some economic comparisons using this useful performance measure.

Engineers have pointed out some areas where efficiency of the existing hardware could be improved. For example, the fan blades used to move the air are usually stamped from sheet metal, an economical method of manufacture, but as a result they are not aerodynamically efficient. A well-designed blade could reduce electrical power required to move the air by a third.

- Chilled beam
- Circulator pump
- Cooling tower
- Damper (flow)
- Dedicated outdoor air system
- Diffuser
- Displacement Ventilation
- Duct
- Economizer
- Evaporative cooler
- Fan coil unit
- Fan (mechanical)
- Heater
- Heat exchanger, including 'coils'
- Heat Pump
- Heat recovery ventilator
- Humidifier / Dehumidifier
- HVAC control system
- Piping
- Valve
- Variable air volume
- Variable-frequency drive, for fine control of pumps
- Underfloor air distribution

HVAC industry and standards

North America

USA

In the United States, HVAC engineers generally are members of the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE). ASHRAE is an international technical society for all individuals and organizations interested in HVAC. The Society, organized into Regions, Chapters, and Student Branches, allows exchange of HVAC knowledge and experiences for the benefit of the field's practitioners and the public. ASHRAE provides many opportunities to participate in the development of new knowledge via, for example, research and its many Technical Committees. These committees meet typically twice per year at the ASHRAE Annual and Winter Meetings. A popular product show, the AHR Expo, is held in conjunction with each Winter Meeting. The Society has approximately 50,000 members and has headquarters at Atlanta, Georgia, USA.

The most recognized standards for HVAC design is based on ASHRAE data. ASHRAE is the American Society of Heating, Refrigerating and Air-Conditioning Engineers. The ASHRAE Handbook's most general volume, of four, is Fundamentals; it includes heating and cooling calculations. Each volume of the ASHRAE Handbook is updated every four years. The design professional must consult ASHRAE data for the standards of design and care as the typical building codes provides little to no information on HVAC design practices; such codes, such as the UMC and IMC, do include much details on installation requirements, however. Other useful reference materials include items from SMACNA, ACCA, and technical trade journals.

American design standards are legislated in the Uniform Mechanical Code or International Mechanical Code. In certain states, counties, or cities, either of these codes may be adopted and amended via various legislative processes. These codes are updated and published by the International Association of Plumbing and Mechanical Officials (IAPMO) or the International Code Council (ICC) respectively, on a 3-year code development cycle. Typically, local Building Permit Departments are charged with enforcement of these standards on private and certain public properties.

In the United States, as well as throughout the world, HVAC contractors and companies are members of NADCA, the National Air Duct Cleaners Association. NADCA was formed in 1989 as a non-profit association of companies engaged in the cleaning of HVAC systems. Its mission was to promote source removal as the only acceptable method of cleaning and to establish industry standards for the association. NADCA has expanded its mission to include the representation of qualified companies engaged in the assessment, cleaning, and restoration of HVAC systems, and to assist its members in providing high quality service to their customers. The goal of the association is to be the number one source for the HVAC cleaning and restoration services: first time, every time. NADCA has experienced phenomenal membership growth and has been extremely successful with the training and certification of air systems cleaning specialists, mold remediators, and HVAC inspectors. The association has also published important standards and guidelines, educational materials, and other useful information for the consumer and members of NADCA. Their headquarters are located in Washington, D.C.

Europe

United Kingdom

The Chartered Institute of Building Services Engineers is a body that covers the essential Service (systems architecture) that allow buildings to operate. It includes the electrotechnical, heating, ventilating, air conditioning, refrigeration and plumbing industries. To train as a building services engineer, the academic requirement are GCSEs (A-C) / Standard Grades (1-3) in Maths and Science, which are important in measurements, planning and theory. Employers will often want a degree in a branch of engineering, such as building environment engineering, electrical engineering or mechanical engineering. To become a full member of CIBSE, and so also to be registered by the Engineering Council UK as a chartered engineer, one must also attain an Honours Degree and a Masters Degree in a relevent engineering subject.

CIBSE publishes several guides to HVAC design relevent to the UK market, and also the Republic of Ireland, Australia, New Zealand and Hong Kong. These guides include for various recommended design criteria and standards, some of which are cited within the UK building regulations, and therefore form a legislative requirment for major building services works. The main guides are:

- Guide A: Environmental Design
- Guide B: Heating, Ventilating, Air Conditioning and Refrigeration
- Guide C: Reference Data
- Guide D: Transportation systems in Buildings
- Guide E: Fire Safety Engineering
- Guide F: Energy Efficiency in Buildings
- Guide G: Public Health Engineering
- Guide H: Building Control Systems
- Guide J: Weather, Solar and Illuminance Data
- Guide K: Electricity in Buildings
- Guide L: Sustainability
- Guide M: Maintenance Engineering and Management

Within the construction sector, it is the job of the building services engineer to design and oversee the installation and maintenance of the essential services such as gas, electricity, water, heating and lighting, as well as many others. These all help to make buildings comfortable and healthy places to live and work in. Building Services is part of a sector that has over 51,000 businesses and employs over 500,000 people. This sector has an annual turnover of £19.3 billion which represents 2%-3% of the GDP.

Australia

Air Conditioning and Mechanical Contractors Association of Australia (AMCA)
Australian Institute of Refrigeration, Air Conditioning and Heating (AIRAH),
CIBSE

Asia

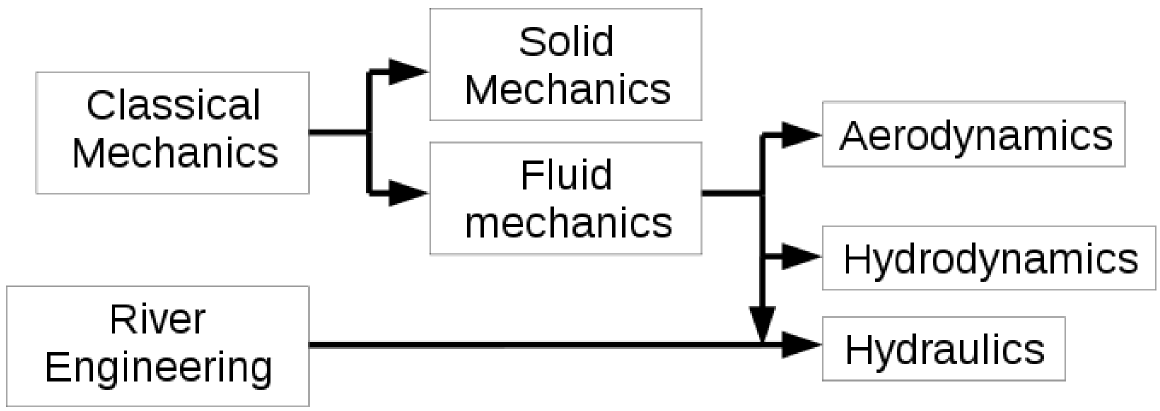
India

The Indian Society of Heating, Refrigerating and Air Conditioning Engineers (ISHRAE) was established to promote the HVAC industry in India. ISHRAE is an associate of ASHRAE. ISHRAE was started at Delhi in 1981 and a chapter was started in Bangalore in 1989. Between 1989 & 1993, ISHRAE chapters were formed in all major cities in India and also in the Middle East.

Chapter- 8

Hydraulics and Pneumatics

Hydraulics



Hydraulics and other studies

Hydraulics is a topic in applied science and engineering dealing with the mechanical properties of liquids. Fluid mechanics provides the theoretical foundation for hydraulics, which focuses on the engineering uses of fluid properties. In fluid power, hydraulics is used for the generation, control, and transmission of power by the use of pressurized liquids. Hydraulic topics range through most science and engineering disciplines, and cover concepts such as pipe flow, dam design, fluidics and fluid control circuitry, pumps, turbines, hydropower, computational fluid dynamics, flow measurement, river channel behavior and erosion.

Free surface hydraulics is the branch of hydraulics dealing with free surface flow, such as occurring in rivers, canals, lakes, estuaries and seas. Its sub-field **open channel flow** studies the flow in open channels.

The word "hydraulics" originates from the Greek word *ὑδραυλικός* (*hydraulikos*) which in turn originates from *ὕδωρ* (*hydor*, Greek for water) and *αὐλός* (*aulos*, meaning pipe).

Ancient and medieval era

Early uses of water power date back to Mesopotamia and ancient Egypt, where irrigation has been used since the 6th millennium BC and water clocks had been used since the early 2nd millennium BC. Other early examples of water power include the Qanat system in ancient Persia and the Turpan water system in ancient China.

Greek / Hellenistic world

Greeks continued and sophisticated the construction of water and hydraulic power systems. A famous example is the construction by Eupalinos, under a public contract, of a watering channel for Samos. An early example of the usage of hydraulic wheel, probably the earliest in Europe, is the Perachora wheel (3rd c. BC).

Notable is the construction of the first hydraulic automata by Ctesibius (flourished c. 270 BC) and Hero of Alexandria (c. 10–80 AD). Hero describes a number of working machines using hydraulic power, such as the force pump, which is known from many Roman sites as having been used for raising water and in fire engines.

China

In ancient China there was Sunshu Ao (6th century BC), Ximen Bao (5th century BC), Du Shi (circa 31 AD), Zhang Heng (78 - 139 AD), and Ma Jun (200 - 265 AD), while medieval China had Su Song (1020 - 1101 AD) and Shen Kuo (1031–1095). Du Shi employed a waterwheel to power the bellows of a blast furnace producing cast iron. Zhang Heng was the first to employ hydraulics to provide motive power in rotating an armillary sphere for astronomical observation.

Sri Lanka

Moat and gardens at Sigirya

In ancient Sri Lanka, hydraulics were widely used in the ancient kingdoms of Anuradhapura and Polonnaruwa. The discovery of the principle of the valve tower, or valve pit, for regulating the escape of water is credited to ingenuity more than 2,000 years ago. By the first century A.D, several large-scale irrigation works had been completed. Macro- and micro-hydraulics to provide for domestic horticultural and agricultural needs, surface drainage and erosion control, ornamental and recreational water courses and retaining structures and also cooling systems were in place in Sigiriya, Sri Lanka. The citadel on the massive rock at the site includes cisterns for collecting water.

Innovations in Ancient Rome

Aqueduct of Segovia

In Ancient Rome many different hydraulic applications were developed, including public water supplies, innumerable aqueducts, power using watermills and hydraulic mining. They were among the first to make use of the siphon to carry water across valleys, and used hushing on a large scale to prospect for and then extract metal ores. They used lead widely in plumbing systems for domestic and public supply, such as feeding thermae.

Hydraulic mining was used in the gold-fields of northern Spain, which was conquered by Augustus in 25 BC. The alluvial gold-mine of Las Medulas was one of the largest of their mines. It was worked by at least 7 long aqueducts, and the water streams were used to erode the soft deposits, and then wash the tailings for the valuable gold content.

Modern era (C. 1600–1870)

Benedetto Castelli

In 1619 Benedetto Castelli (1576 - 1578–1643), a student of Galileo Galilei, published the book *Della Misura dell'Acque Correnti* or "On the Measurement of Running Waters", one of the foundations of modern hydrodynamics. He served as a chief consultant to the Pope on hydraulic projects, i.e., management of rivers in the Papal States, beginning in 1626.

Blaise Pascal

Blaise Pascal (1623–1662-1672) studied fluid hydrodynamics and hydrostatics, centered on the principles of hydraulic fluids. His inventions include the hydraulic press, which multiplied a smaller force acting on a larger area into the application of a larger force totaled over a smaller area, transmitted through the same pressure (or same change of pressure) at both locations. Pascal's law or principle states that for an incompressible fluid at rest, the difference in pressure is proportional to the difference in height and this difference remains the same whether or not the overall pressure of the fluid is changed by applying an external force. This implies that by increasing the pressure at any point in a confined fluid, there is an equal increase at every other point in the container, i.e., any change in pressure applied at any point of the fluid is transmitted undiminished throughout the fluids.

Jean Louis Marie Poiseuille

A French physician, Poiseuille researched the flow of blood through the body and discovered an important law governing the rate of flow with the diameter of the tube in which flow occurred.

Pneumatics

Preserved Porter Locomotive Company No. 3290 of 1923

Pneumatics is a branch of technology, which deals with the study and application of use of pressurized gas to affect mechanical motion.

Pneumatic systems are extensively used in industry, where factories are commonly plumbed with compressed air or other compressed inert gases. This is because a centrally-located and electrically-powered compressor that powers cylinders and other pneumatic devices through solenoid valves is often able to provide motive power in a cheaper, safer, more flexible, and more reliable way than a large number of electric motors and actuators.

Pneumatics also has applications in dentistry, construction, mining, and other areas.

Examples of pneumatic systems and components

- Air brakes on buses and trucks
- Air brakes, on trains
- Air compressors
- Air engines for pneumatically powered vehicles
- Barostat systems used in Neurogastroenterology and for researching electricity
- Cable jetting, a way to install cables in ducts
- Compressed-air engine and compressed-air vehicles

- Gas-operated reloading
- Holman Projector, a pneumatic anti-aircraft weapon
- Lego pneumatics can be used to build pneumatic models

- Pipe organs:
    - Electro-pneumatic action
    - Tubular-pneumatic action

- Pneumatic actuator
- Pneumatic air guns
- Pneumatic cylinder
- Pneumatic Launchers, a type of spud gun
- Pneumatic mail systems
- Pneumatic motor
- Pneumatic tire

- Pneumatic tools:
    - Jackhammer used by road workers
    - Pneumatic nailgun

- Pressure regulator
- Pressure sensor
- Pressure switch

- Vacuum pump

Gases used in pneumatic systems

Pneumatic systems in fixed installations such as factories use compressed air because a sustainable supply can be made by compressing atmospheric air. The air usually has moisture removed and a small quantity of oil added at the compressor, to avoid corrosion of mechanical components and to lubricate them.

Factory-plumbed, pneumatic-power users need not worry about poisonous leakages as the gas is commonly just air. Smaller or stand-alone systems can use other compressed gases which are an asphyxiation hazard, such as nitrogen - often referred to as OFN (oxygen-free nitrogen), when supplied in cylinders.

Any compressed gas other than air is an asphyxiation hazard - including nitrogen, which makes up approximately 80% of air. Compressed oxygen (approx. 20% of air) would not asphyxiate, but it would be an extreme fire hazard, so is never used in pneumatically powered devices.

Portable pneumatic tools and small vehicles such as Robot Wars machines and other hobbyist applications are often powered by compressed carbon dioxide because containers designed to hold it such as soda stream canisters and fire extinguishers are readily available, and the phase change between liquid and gas makes it possible to obtain a larger volume of compressed gas from a lighter container than compressed air would allow. Carbon dioxide is an asphyxiant and can also be a freezing hazard when vented inappropriately.

Comparison to hydraulics

Both pneumatics and hydraulics are applications of fluid power. Pneumatics uses an easily compressible gas such as air or a suitable pure gas, while hydraulics uses relatively incompressible liquid media such as oil. Most industrial pneumatic applications use pressures of about 80 to 100 pounds per square inch (550 to 690 kPa). Hydraulics applications commonly use from 1,000 to 5,000 psi (6.9 to 34 MPa), but specialized applications may exceed 10,000 psi (69 MPa).

Advantages of pneumatics

- **Simplicity of Design And Control**
  - Machines are easily designed using standard cylinders & other components. Control is as easy as it is simple ON - OFF type control.
- **Reliability**
  - Pneumatic systems tend to have long operating lives and require very little maintenance.
  - Because gas is compressible, the equipment is less likely to be damaged by shock. The gas in pneumatics absorbs excessive force, whereas the fluid of hydraulics directly transfers force.
- **Storage**
  - Compressed Gas can be stored, allowing the use of machines when electrical power is lost.
- **Safety**
  - Very low chance of fire (compared to hydraulic oil).
  - Machines can be designed to be overload safe.

Advantages of hydraulics

- Liquid (as a gas is also a 'fluid') does not absorb any of the supplied energy.
- Capable of moving much higher loads and providing much higher forces due to the incompressibility.
- The hydraulic working fluid is basically incompressible, leading to a minimum of spring action. When hydraulic fluid flow is stopped, the slightest motion of the load releases the pressure on the load; there is no need to "bleed off" pressurized air to release the pressure on the load.

Pneumatic logic

Pneumatic logic systems (sometimes called **air logic control**) are often used to control industrial processes, consisting of primary logic units such as:

- And Units
- Or Units
- 'Relay or Booster' Units
- Latching Units
- 'Timer' Units
- Sorteberg relay
- fluidics amplifiers with no moving parts other than the air itself

Pneumatic logic is a reliable and functional control method for industrial processes. In recent years, these systems have largely been replaced by electrical control systems, due to the smaller size and lower cost of electrical components. Pneumatic devices are still used in processes where compressed air is the only energy source available or upgrade cost, safety, and other considerations outweigh the advantage of modern digital control.

Chapter- 9

Fluid Mechanics

Fluid mechanics is the study of fluids and the forces on them. (Fluids include liquids, gases, and plasmas.) Fluid mechanics can be divided into fluid kinematics, the study of fluid motion, and fluid dynamics, the study of the effect of forces on fluid motion, which can further be divided into fluid statics, the study of fluids at rest, and fluid kinetics, the study of fluids in motion. It is a branch of continuum mechanics, a subject which models matter without using the information that it is made out of atoms, that is, it models matter from a macroscopic viewpoint rather than from a microscopic viewpoint. Fluid mechanics, especially fluid dynamics, is an active field of research with many unsolved or partly solved problems. Fluid mechanics can be mathematically complex. Sometimes it can best be solved by numerical methods, typically using computers. A modern discipline, called computational fluid dynamics (CFD), is devoted to this approach to solving fluid mechanics problems. Also taking advantage of the highly visual nature of fluid flow is particle image velocimetry, an experimental method for visualizing and analyzing fluid flow.

Brief history

The study of fluid mechanics goes back at least to the days of ancient Greece, when Archimedes investigated fluid statics and buoyancy and formulated his famous law known now as the Archimedes Principle. Rapid advancement in fluid mechanics began with Leonardo da Vinci (observation and experiment), Evangelista Torricelli (barometer), Isaac Newton (viscosity) and Blaise Pascal (hydrostatics), and was continued by Daniel Bernoulli with the introduction of mathematical fluid dynamics in *Hydrodynamica* (1738). Inviscid flow was further analyzed by various mathematicians (Leonhard Euler, d'Alembert, Lagrange, Laplace, Poisson) and viscous flow was explored by a multitude of engineers including Poiseuille and Gotthilf Heinrich Ludwig Hagen. Further mathematical justification was provided by Claude-Louis Navier and George Gabriel Stokes in the Navier–Stokes equations, and boundary layers were investigated (Ludwig Prandtl), while various scientists (Osborne Reynolds, Andrey Kolmogorov, Geoffrey Ingram Taylor) advanced the understanding of fluid viscosity and turbulence.

Relationship to continuum mechanics

Fluid mechanics is a subdiscipline of continuum mechanics, as illustrated in the following table.

<table>
<tr><td rowspan="4">Continuum mechanics
The study of the physics of continuous materials</td><td rowspan="2">Solid mechanics
The study of the physics of continuous materials with a defined rest shape.</td><td colspan="2">Elasticity
Describes materials that return to their rest shape after an applied stress.</td></tr>
<tr><td>Plasticity
Describes materials that permanently deform after a sufficient applied stress.</td><td rowspan="2">Rheology
The study of materials with both solid and fluid characteristics.</td></tr>
<tr><td rowspan="2">Fluid mechanics
The study of the physics of continuous materials which take the shape of their container.</td><td>Non-Newtonian fluids</td></tr>
<tr><td colspan="2">Newtonian fluids</td></tr>
</table>

In a mechanical view, a fluid is a substance that does not support shear stress; that is why a fluid at rest has the shape of its containing vessel. A fluid at rest has no shear stress.

Assumptions

Like any mathematical model of the real world, fluid mechanics makes some basic assumptions about the materials being studied. These assumptions are turned into equations that must be satisfied if the assumptions are to be held true. For example, consider an incompressible fluid in three dimensions. The assumption that mass is conserved means that for any fixed closed surface (such as a sphere) the rate of mass passing from *outside* to *inside* the surface must be the same as rate of mass passing the other way. (Alternatively, the mass *inside* remains constant, as does the mass *outside*). This can be turned into an integral equation over the surface.

Fluid mechanics assumes that every fluid obeys the following:

- Conservation of mass
- Conservation of energy
- Conservation of momentum
- The *continuum hypothesis*, detailed below.

Further, it is often useful (at subsonic conditions) to assume a fluid is incompressible – that is, the density of the fluid does not change. Liquids can often be modelled as incompressible fluids, whereas gases cannot.

Similarly, it can sometimes be assumed that the viscosity of the fluid is zero (the fluid is *inviscid*). Gases can often be assumed to be inviscid. If a fluid is viscous, and its flow contained in some way (e.g. in a pipe), then the flow at the boundary must have zero velocity. For a viscous fluid, if the boundary is not porous, the shear forces between the fluid and the boundary results also in a zero velocity for the fluid at the boundary. This is called the no-slip condition. For a porous media otherwise, in the frontier of the containing vessel, the slip condition is not zero velocity, and the fluid has a discontinuous velocity field between the free fluid and the fluid in the porous media (this is related to the Beavers and Joseph condition).

The continuum hypothesis

Fluids are composed of molecules that collide with one another and solid objects. The continuum assumption, however, considers fluids to be continuous. That is, properties such as density, pressure, temperature, and velocity are taken to be well-defined at "infinitely" small points, defining a REV (Reference Element of Volume), at the geometric order of the distance between two adjacent molecules of fluid. Properties are assumed to vary continuously from one point to another, and are averaged values in the REV. The fact that the fluid is made up of discrete molecules is ignored.

The continuum hypothesis is basically an approximation, in the same way planets are approximated by point particles when dealing with celestial mechanics, and therefore results in approximate solutions. Consequently, assumption of the continuum hypothesis can lead to results which are not of desired accuracy. That said, under the right circumstances, the continuum hypothesis produces extremely accurate results.

Those problems for which the continuum hypothesis does not allow solutions of desired accuracy are solved using statistical mechanics. To determine whether or not to use conventional fluid dynamics or statistical mechanics, the Knudsen number is evaluated for the problem. The Knudsen number is defined as the ratio of the molecular mean free path length to a certain representative physical length scale. This length scale could be, for example, the radius of a body in a fluid. (More simply, the Knudsen number is how many times its own diameter a particle will travel on average before hitting another particle). Problems with Knudsen numbers at or above unity are best evaluated using statistical mechanics for reliable solutions.

Navier–Stokes equations

The **Navier–Stokes equations** (named after Claude-Louis Navier and George Gabriel Stokes) are the set of equations that describe the motion of fluid substances such as liquids and gases. These equations state that changes in momentum (force) of fluid particles depend only on the external pressure and internal viscous forces (similar to friction) acting on the fluid. Thus, the Navier–Stokes equations describe the balance of forces acting at any given region of the fluid.

The Navier–Stokes equations are differential equations which describe the motion of a fluid. Such equations establish relations among the rates of change of the variables of interest. For example, the Navier–Stokes equations for an ideal fluid with zero viscosity states that acceleration (the rate of change of velocity) is proportional to the derivative of internal pressure.

This means that solutions of the Navier–Stokes equations for a given physical problem must be sought with the help of calculus. In practical terms only the simplest cases can be solved exactly in this way. These cases generally involve non-turbulent, steady flow (flow does not change with time) in which the Reynolds number is small.

For more complex situations, such as global weather systems like El Niño or lift in a wing, solutions of the Navier–Stokes equations can currently only be found with the help of computers. This is a field of sciences by its own called computational fluid dynamics.

General form of the equation

The general form of the Navier–Stokes equations for the conservation of momentum is:

$$\rho\frac{D\mathbf{v}}{Dt} = \nabla \cdot \mathbb{P} + \rho\mathbf{f}$$

where

- ρ is the fluid density,
- $\frac{D}{Dt}$ is the substantive derivative (also called the material derivative),
- $\mathbf{v}$ is the velocity vector,
- $\mathbf{f}$ is the body force vector, and
- $\mathbb{P}$ is a tensor that represents the surface forces applied on a fluid particle (the stress tensor).

Unless the fluid is made up of spinning degrees of freedom like vortices, $\mathbb{P}$ is a symmetric tensor. In general, (in three dimensions) $\mathbb{P}$ has the form:

$$\mathbb{P} = \begin{pmatrix} \sigma_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_{yy} & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_{zz} \end{pmatrix}$$

where

- σ are normal stresses,
- τ are tangential stresses (shear stresses).

The above is actually a set of three equations, one per dimension. By themselves, these aren't sufficient to produce a solution. However, adding conservation of mass and appropriate boundary conditions to the system of equations produces a solvable set of equations.

Newtonian versus non-Newtonian fluids

A **Newtonian fluid** (named after Isaac Newton) is defined to be a fluid whose shear stress is linearly proportional to the velocity gradient in the direction perpendicular to the plane of shear. This definition means regardless of the forces acting on a fluid, it *continues to flow*. For example, water is a Newtonian fluid, because it continues to display fluid properties no matter how much it is stirred or mixed. A slightly less rigorous definition is that the drag of a small object being moved slowly through the fluid is proportional to the force applied to the object. (Compare friction). Important fluids, like water as well as most gases, behave — to good approximation — as a Newtonian fluid under normal conditions on Earth.

By contrast, stirring a non-Newtonian fluid can leave a "hole" behind. This will gradually fill up over time – this behaviour is seen in materials such as pudding, oobleck, or sand (although sand isn't strictly a fluid). Alternatively, stirring a non-Newtonian fluid can cause the viscosity to decrease, so the fluid appears "thinner" (this is seen in non-drip paints). There are many types of non-Newtonian fluids, as they are defined to be something that fails to obey a particular property — for example, most fluids with long molecular chains can react in a non-Newtonian manner.

Equations for a Newtonian fluid

The constant of proportionality between the shear stress and the velocity gradient is known as the viscosity. A simple equation to describe Newtonian fluid behaviour is

$$\tau = -\mu \frac{dv}{dy}$$

where

τ is the shear stress exerted by the fluid ("drag")
μ is the fluid viscosity – a constant of proportionality
$\frac{dv}{dy}$ is the velocity gradient perpendicular to the direction of shear.

For a Newtonian fluid, the viscosity, by definition, depends only on temperature and pressure, not on the forces acting upon it. If the fluid is incompressible and viscosity is constant across the fluid, the equation governing the shear stress (in Cartesian coordinates) is

$$\tau_{ij} = \mu \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right)$$

where

τ_{ij} is the shear stress on the i^{th} face of a fluid element in the j^{th} direction
v_i is the velocity in the i^{th} direction
x_j is the j^{th} direction coordinate.

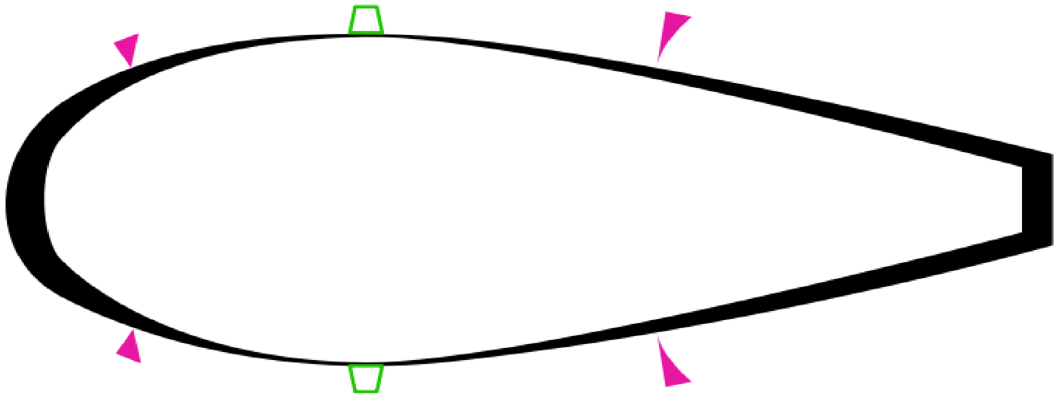
If a fluid does not obey this relation, it is termed a non-Newtonian fluid, of which there are several types.

Among fluids, two rough broad divisions can be made: ideal and non-ideal fluids. An ideal fluid really does not exist, but in some calculations, the assumption is justifiable. An Ideal fluid is non viscous- offers no resistance whatsoever to a shearing force.

One can group real fluids into Newtonian and non-Newtonian. Newtonian fluids agree with Newton's law of viscosity. Non-Newtonian fluids can be either plastic, bingham plastic, pseudoplastic, dilatant, thixotropic, rheopectic, viscoelatic.

Chapter- 10

Fluid Dynamics



Typical aerodynamic teardrop shape, showing the pressure distribution as the thickness of the black line and showing the velocity in the boundary layer as the violet triangles. The green vortex generators prompt the transition to turbulent flow and prevent back-flow also called flow separation from the high pressure region in the back. The surface in front is as smooth as possible or even employs shark like skin, as any turbulence here will reduce the energy of the airflow. The Kammback also prevents back flow from the high pressure region in the back across the spoilers to the convergent part. Putting stuff inside out results in tubes; they also face the problem of flow separation in their divergent parts, so called diffusers. Cutting the shape into halves results in an aerofoil with the low pressure region on top leading to lift (force).

In physics, **fluid dynamics** is a sub-discipline of fluid mechanics that deals with **fluid flow**—the natural science of fluids (liquids and gases) in motion. It has several subdisciplines itself, including aerodynamics (the study of air and other gases in motion) and **hydrodynamics** (the study of liquids in motion). Fluid dynamics has a wide range of applications, including calculating forces and moments on aircraft, determining the mass flow rate of petroleum through pipelines, predicting weather patterns, understanding nebulae in interstellar space and reportedly modeling fission weapon detonation. Some of its principles are even used in traffic engineering, where traffic is treated as a continuous fluid.

Fluid dynamics offers a systematic structure that underlies these practical disciplines, that embraces empirical and semi-empirical laws derived from flow measurement and used to solve practical problems. The solution to a fluid dynamics problem typically involves calculating various properties of the fluid, such as velocity, pressure, density, and temperature, as functions of space and time.

Historically, *hydrodynamics* meant something different than it does today. Before the twentieth century, hydrodynamics was synonymous with fluid dynamics. This is still reflected in names of some fluid dynamics topics, like magnetohydrodynamics and hydrodynamic stability—both also applicable in, as well as being applied to, gases.

Equations of fluid dynamics

The foundational axioms of fluid dynamics are the conservation laws, specifically, conservation of mass, conservation of linear momentum (also known as Newton's Second Law of Motion), and conservation of energy (also known as First Law of Thermodynamics). These are based on classical mechanics and are modified in quantum mechanics and general relativity. They are expressed using the Reynolds Transport Theorem.

In addition to the above, fluids are assumed to obey the *continuum assumption*. Fluids are composed of molecules that collide with one another and solid objects. However, the continuum assumption considers fluids to be continuous, rather than discrete. Consequently, properties such as density, pressure, temperature, and velocity are taken to be well-defined at infinitesimally small points, and are assumed to vary continuously from one point to another. The fact that the fluid is made up of discrete molecules is ignored.

For fluids which are sufficiently dense to be a continuum, do not contain ionized species, and have velocities small in relation to the speed of light, the momentum equations for Newtonian fluids are the Navier-Stokes equations, which is a non-linear set of differential equations that describes the flow of a fluid whose stress depends linearly on velocity gradients and pressure. The unsimplified equations do not have a general closed-form solution, so they are primarily of use in Computational Fluid Dynamics. The equations can be simplified in a number of ways, all of which make them easier to solve. Some of them allow appropriate fluid dynamics problems to be solved in closed form.

In addition to the mass, momentum, and energy conservation equations, a thermodynamical equation of state giving the pressure as a function of other thermodynamic variables for the fluid is required to completely specify the problem. An example of this would be the perfect gas equation of state:

$$p = \frac{\rho R_u T}{M}$$

where p is pressure, ρ is density, R_u is the gas constant, M is the molar mass and T is temperature.

Compressible vs incompressible flow

All fluids are compressible to some extent, that is changes in pressure or temperature will result in changes in density. However, in many situations the changes in pressure and temperature are sufficiently small that the changes in density are negligible. In this case the flow can be modeled as an incompressible flow. Otherwise the more general compressible flow equations must be used.

Mathematically, incompressibility is expressed by saying that the density ρ of a fluid parcel does not change as it moves in the flow field, i.e.,

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = 0\,,$$

where D / Dt is the substantial derivative, which is the sum of local and convective derivatives. This additional constraint simplifies the governing equations, especially in the case when the fluid has a uniform density.

For flow of gases, to determine whether to use compressible or incompressible fluid dynamics, the Mach number of the flow is to be evaluated. As a rough guide, compressible effects can be ignored at Mach numbers below approximately 0.3. For liquids, whether the incompressible assumption is valid depends on the fluid properties (specifically the critical pressure and temperature of the fluid) and the flow conditions (how close to the critical pressure the actual flow pressure becomes). Acoustic problems always require allowing compressibility, since sound waves are compression waves involving changes in pressure and density of the medium through which they propagate.

Viscous vs inviscid flow

Viscous problems are those in which fluid friction has significant effects on the fluid motion.

The Reynolds number, which is a ratio between inertial and viscous forces, can be used to evaluate whether viscous or inviscid equations are appropriate to the problem.

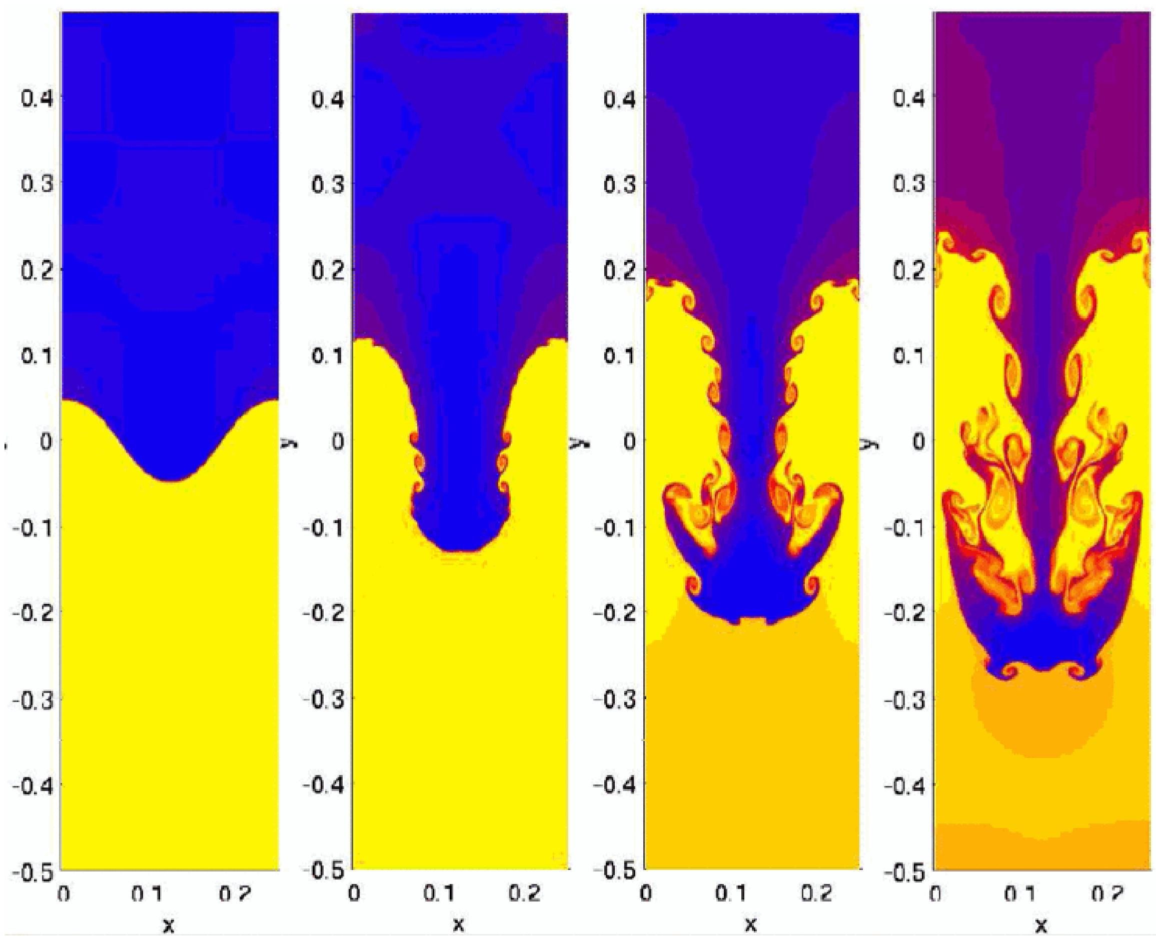
Stokes flow is flow at very low Reynolds numbers, $Re \ll 1$, such that inertial forces can be neglected compared to viscous forces.

On the contrary, high Reynolds numbers indicate that the inertial forces are more significant than the viscous (friction) forces. Therefore, we may assume the flow to be an inviscid flow, an approximation in which we neglect viscosity completely, compared to inertial terms.

This idea can work fairly well when the Reynolds number is high. However, certain problems such as those involving solid boundaries, may require that the viscosity be included. Viscosity often cannot be neglected near solid boundaries because the no-slip condition can generate a thin region of large strain rate (known as Boundary layer) which enhances the effect of even a small amount of viscosity, and thus generating vorticity. Therefore, to calculate net forces on bodies (such as wings) we should use viscous flow equations. As illustrated by d'Alembert's paradox, a body in an inviscid fluid will experience no drag force. The standard equations of inviscid flow are the Euler equations. Another often used model, especially in computational fluid dynamics, is to use the Euler equations away from the body and the boundary layer equations, which incorporates viscosity, in a region close to the body.

The Euler equations can be integrated along a streamline to get Bernoulli's equation. When the flow is everywhere irrotational and inviscid, Bernoulli's equation can be used throughout the flow field. Such flows are called potential flows.

Steady vs unsteady flow



Hydrodynamics simulation of the Rayleigh–Taylor instability

When all the time derivatives of a flow field vanish, the flow is considered to be a **steady flow**. Steady-state flow refers to the condition where the fluid properties at a point in the system do not change over time. Otherwise, flow is called unsteady. Whether a particular flow is steady or unsteady, can depend on the chosen frame of reference. For instance, laminar flow over a sphere is steady in the frame of reference that is stationary with respect to the sphere. In a frame of reference that is stationary with respect to a background flow, the flow is unsteady.

Turbulent flows are unsteady by definition. A turbulent flow can, however, be statistically stationary. According to Pope:

The random field $U(x,t)$ is statistically stationary if all statistics are invariant under a shift in time.

This roughly means that all statistical properties are constant in time. Often, the mean field is the object of interest, and this is constant too in a statistically stationary flow.

Steady flows are often more tractable than otherwise similar unsteady flows. The governing equations of a steady problem have one dimension fewer (time) than the governing equations of the same problem without taking advantage of the steadiness of the flow field.

Laminar vs turbulent flow

Turbulence is flow characterized by recirculation, eddies, and apparent randomness. Flow in which turbulence is not exhibited is called laminar. It should be noted, however, that the presence of eddies or recirculation alone does not necessarily indicate turbulent flow—these phenomena may be present in laminar flow as well. Mathematically, turbulent flow is often represented via a Reynolds decomposition, in which the flow is broken down into the sum of an average component and a perturbation component.

It is believed that turbulent flows can be described well through the use of the Navier–Stokes equations. Direct numerical simulation (DNS), based on the Navier–Stokes equations, makes it possible to simulate turbulent flows at moderate Reynolds numbers. Restrictions depend on the power of the computer used and the efficiency of the solution algorithm. The results of DNS have been found to agree well with experimental data for some flows.

Most flows of interest have Reynolds numbers much too high for DNS to be a viable option, given the state of computational power for the next few decades. Any flight vehicle large enough to carry a human ($L > 3$ m), moving faster than 72 km/h (20 m/s) is well beyond the limit of DNS simulation ($Re = 4$ million). Transport aircraft wings (such as on an Airbus A300 or Boeing 747) have Reynolds numbers of 40 million (based on the wing chord). In order to solve these real-life flow problems, turbulence models will be a necessity for the foreseeable future. Reynolds-averaged Navier–Stokes equations (RANS) combined with turbulence modeling provides a model of the effects of the turbulent flow. Such a modeling mainly provides the additional momentum transfer by the Reynolds stresses, although the turbulence also enhances the heat and mass transfer. Another promising methodology is large eddy simulation (LES), especially in the guise of detached eddy simulation (DES)—which is a combination of RANS turbulence modeling and large eddy simulation.

Newtonian vs non-Newtonian fluids

Sir Isaac Newton showed how stress and the rate of strain are very close to linearly related for many familiar fluids, such as water and air. These Newtonian fluids are modeled by a coefficient called viscosity, which depends on the specific fluid.

However, some of the other materials, such as emulsions and slurries and some visco-elastic materials (e.g. blood, some polymers), have more complicated *non-Newtonian* stress-strain behaviours. These materials include *sticky liquids* such as latex, honey, and lubricants which are studied in the sub-discipline of rheology.

Subsonic vs transonic, supersonic and hypersonic flows

While many terrestrial flows (e.g. flow of water through a pipe) occur at low mach numbers, many flows of practical interest (e.g. in aerodynamics) occur at high fractions of the Mach Number M=1 or in excess of it (supersonic flows). New phenomena occur at these Mach number regimes (e.g. shock waves for supersonic flow, transonic instability in a regime of flows with M nearly equal to 1, non-equilibrium chemical behavior due to ionization in hypersonic flows) and it is necessary to treat each of these flow regimes separately.

Magnetohydrodynamics

Magnetohydrodynamics is the multi-disciplinary study of the flow of electrically conducting fluids in electromagnetic fields. Examples of such fluids include plasmas, liquid metals, and salt water. The fluid flow equations are solved simultaneously with Maxwell's equations of electromagnetism.

Other approximations

There are a large number of other possible approximations to fluid dynamic problems. Some of the more commonly used are listed below.

- The **Boussinesq approximation** neglects variations in density except to calculate buoyancy forces. It is often used in free convection problems where density changes are small.
- **Lubrication theory** and **Hele-Shaw flow** exploits the large aspect ratio of the domain to show that certain terms in the equations are small and so can be neglected.

- **Slender-body theory** is a methodology used in Stokes flow problems to estimate the force on, or flow field around, a long slender object in a viscous fluid.
- The **shallow-water equations** can be used to describe a layer of relatively inviscid fluid with a free surface, in which surface gradients are small.
- The **Boussinesq equations** are applicable to surface waves on thicker layers of fluid and with steeper surface slopes.
- **Darcy's law** is used for flow in porous media, and works with variables averaged over several pore-widths.
- In rotating systems, the **quasi-geostrophic approximation** assumes an almost perfect balance between pressure gradients and the Coriolis force. It is useful in the study of atmospheric dynamics.

Terminology in fluid dynamics

The concept of pressure is central to the study of both fluid statics and fluid dynamics. A pressure can be identified for every point in a body of fluid, regardless of whether the fluid is in motion or not. Pressure can be measured using an aneroid, Bourdon tube, mercury column, or various other methods.

Some of the terminology that is necessary in the study of fluid dynamics is not found in other similar areas of study. In particular, some of the terminology used in fluid dynamics is not used in fluid statics.

Terminology in incompressible fluid dynamics

The concepts of total pressure and dynamic pressure arise from Bernoulli's equation and are significant in the study of all fluid flows. (These two pressures are not pressures in the usual sense—they cannot be measured using an aneroid, Bourdon tube or mercury column.) To avoid potential ambiguity when referring to pressure in fluid dynamics, many authors use the term static pressure to distinguish it from total pressure and dynamic pressure. Static pressure is identical to pressure and can be identified for every point in a fluid flow field.

In *Aerodynamics*, L.J. Clancy writes: *To distinguish it from the total and dynamic pressures, the actual pressure of the fluid, which is associated not with its motion but with its state, is often referred to as the static pressure, but where the term pressure alone is used it refers to this static pressure.*

A point in a fluid flow where the flow has come to rest (i.e. speed is equal to zero adjacent to some solid body immersed in the fluid flow) is of special significance. It is of such importance that it is given a special name—a stagnation point. The static pressure at the stagnation point is of special significance and is given its own name—stagnation pressure. In incompressible flows, the stagnation pressure at a stagnation point is equal to the total pressure throughout the flow field.

Terminology in compressible fluid dynamics

In a compressible fluid, such as air, the temperature and density are essential when determining the state of the fluid. In addition to the concept of total pressure (also known as stagnation pressure), the concepts of total (or stagnation) temperature and total (or stagnation) density are also essential in any study of compressible fluid flows. To avoid potential ambiguity when referring to temperature and density, many authors use the terms static temperature and static density. Static temperature is identical to temperature; and static density is identical to density; and both can be identified for every point in a fluid flow field.

The temperature and density at a stagnation point are called stagnation temperature and stagnation density.

A similar approach is also taken with the thermodynamic properties of compressible fluids. Many authors use the terms total (or stagnation) enthalpy and total (or stagnation) entropy. The terms static enthalpy and static entropy appear to be less common, but where they are used they mean nothing more than enthalpy and entropy respectively, and the prefix "static" is being used to avoid ambiguity with their 'total' or 'stagnation' counterparts. Because the 'total' flow conditions are defined by isentropically bringing the fluid to rest, the total (or stagnation) entropy is by definition always equal to the "static" entropy.

Chapter- 11

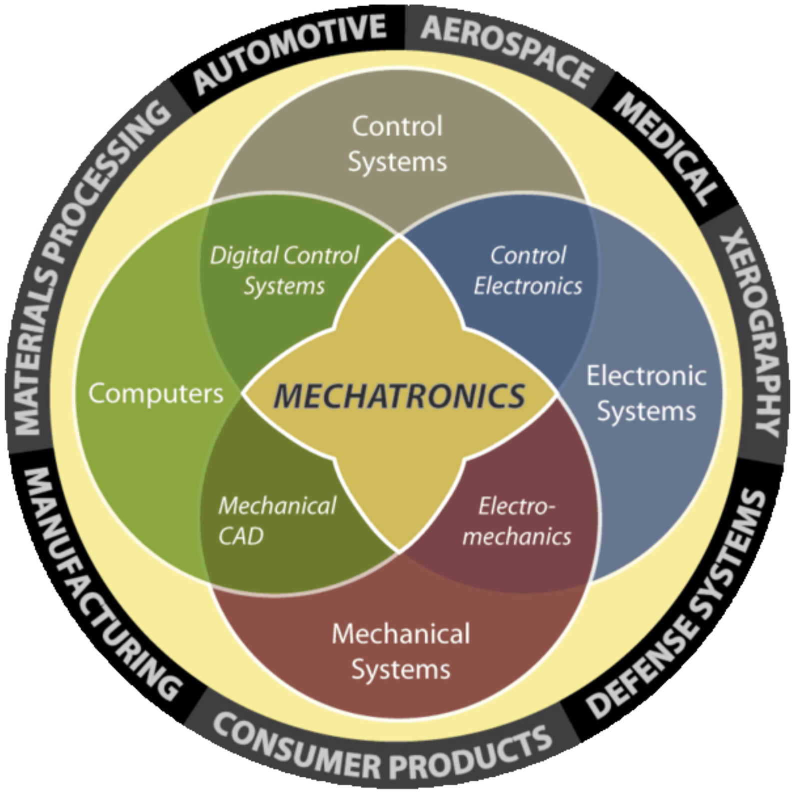
Mechatronics and Control Theory

Mechatronics

Mechatronics is the synergistic combination of Mechanical engineering, Electronic engineering, Computer engineering, Control engineering, and Systems Design engineering in order to design, and manufacture useful products. Mechatronics is a multidisciplinary engineering system design, that is to say it rejects splitting engineering into separate disciplines.

French standard NF E 01-010 gives the following definition: “approach aiming at the synergistic integration of mechanics, electronics, control theory, and computer science within product design and manufacturing, in order to improve and/or optimize its functionality".

Description



Aerial Venn diagram from RPI's website describes the various fields that make up Mechatronics

A mechatronics engineer unites the principles of mechanics, electronics, and computing to generate a simpler, more economical and reliable system. Mechatronics is centered on mechanics, electronics, computing, control engineering, molecular engineering (from nanochemistry and biology), and optical engineering, which, combined, make possible the generation of simpler, more economical, reliable and versatile systems. The portmanteau "mechatronics" was coined by Tetsuro Mori, the senior engineer of the Japanese company Yaskawa in 1969. An industrial robot is a prime example of a mechatronics system; it includes aspects of electronics, mechanics, and computing to do its day-to-day jobs.

Engineering cybernetics deals with the question of control engineering of mechatronic systems. It is used to control or regulate such a system. Through collaboration, the mechatronic modules perform the production goals and inherit flexible and agile manufacturing properties in the production scheme. Modern production equipment consists of mechatronic modules that are integrated according to a control architecture. The most known architectures involve hierarchy, polyarchy, heterarchy, and hybrid. The methods for achieving a technical effect are described by control algorithms, which might or might not utilize formal methods in their design. Hybrid systems important to mechatronics include production systems, synergy drives, planetary exploration rovers, automotive subsystems such as anti-lock braking systems and spin-assist, and every-day equipment such as autofocus cameras, video, hard disks, and CD players.

Science Fiction

Mechatronics has been commonly used in science fiction such as the popular Terminator movies.

Course structure

Mechatronic students take courses from across the various fields listed below:

- Mechanical engineering and materials science subjects
- Electronic engineering subjects
- Computer engineering subjects
- Computer science subjects
- Systems and control engineering subjects
- Optomechanics (optical engineering) subjects
- Robotics subjects

Application

- Machine vision
- Automation and robotics
- Servo-mechanics
- Sensing and control systems
- Automotive engineering, automotive equipment in the design of subsystems such as anti-lock braking systems
- Computer-machine controls, such as computer driven machines like IE CNC milling machines
- Expert systems
- Industrial goods
- Consumer products
- Mechatronics systems
- Medical mechatronics, medical imaging systems
- Structural dynamic systems
- Transportation and vehicular systems
- Mechatronics as the new language of the automobile
- Diagnostic, reliability, and control system techniques
- Computer aided and integrated manufacturing systems
- Computer-aided design
- Engineering and manufacturing systems
- Packaging

Physical implementations

For most mechatronic systems, the main issue is no more how to implement a control system, but how to implement actuators and what is the energy source. Within the mechatronic field, mainly two technologies are used to produce the movement: the piezo-electric actuators and motors, or the electromagnetic actuators and motors. Maybe the most famous mechatronics systems are the well known camera autofocus system or camera anti-shake systems.

Concerning the energy sources, most of the applications use batteries. But a new trend is arriving and is the energy harvesting, allowing transforming into electricity mechanical energy from shock, vibration, or thermal energy from thermal variation, and so on.

Variant of the field

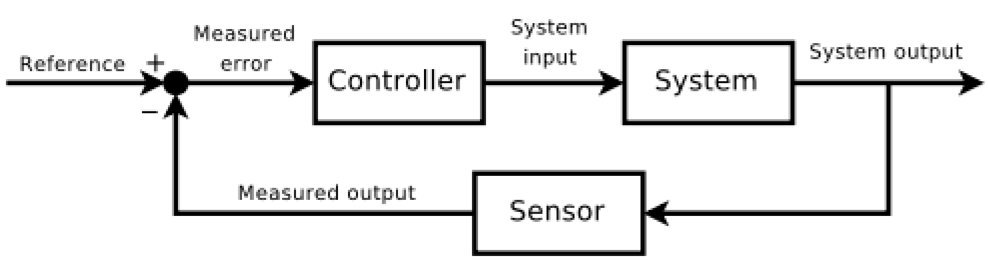
An emerging variant of this field is biomechatronics, whose purpose is to integrate mechanical parts with a human being, usually in the form of removable gadgets such as an exoskeleton. This is the "real-life" version of cyberware.

Another emerging variant is Electronical or electronics design centric ECAD/MCAD co-design. Electronical is where the integration and co-design between the design team and design tools of an electronics centric system and the design team and design tools of that systems physical/mechanical enclosure takes place.

Education

Countries offering education in mechatronics are Japan, Malaysia, France, Germany, United States, UK, Sweden, Canada, Australia and Singapore, among others.

Control theory



The concept of the feedback loop to control the dynamic behavior of the system: this is negative feedback, because the sensed value is subtracted from the desired value to create the error signal which is amplified by the controller.

Control theory is an interdisciplinary branch of engineering and mathematics, that deals with the behavior of dynamical systems. The desired output of a system is called the *reference*. When one or more output variables of a system need to follow a certain reference over time, a controller manipulates the inputs to a system to obtain the desired effect on the output of the system.

Overview

Control theory is

- a theory that deals with influencing the behavior of dynamical systems
- an interdisciplinary subfield of science, which originated in engineering and mathematics, and evolved into use by the social sciences, like psychology, sociology and criminology.

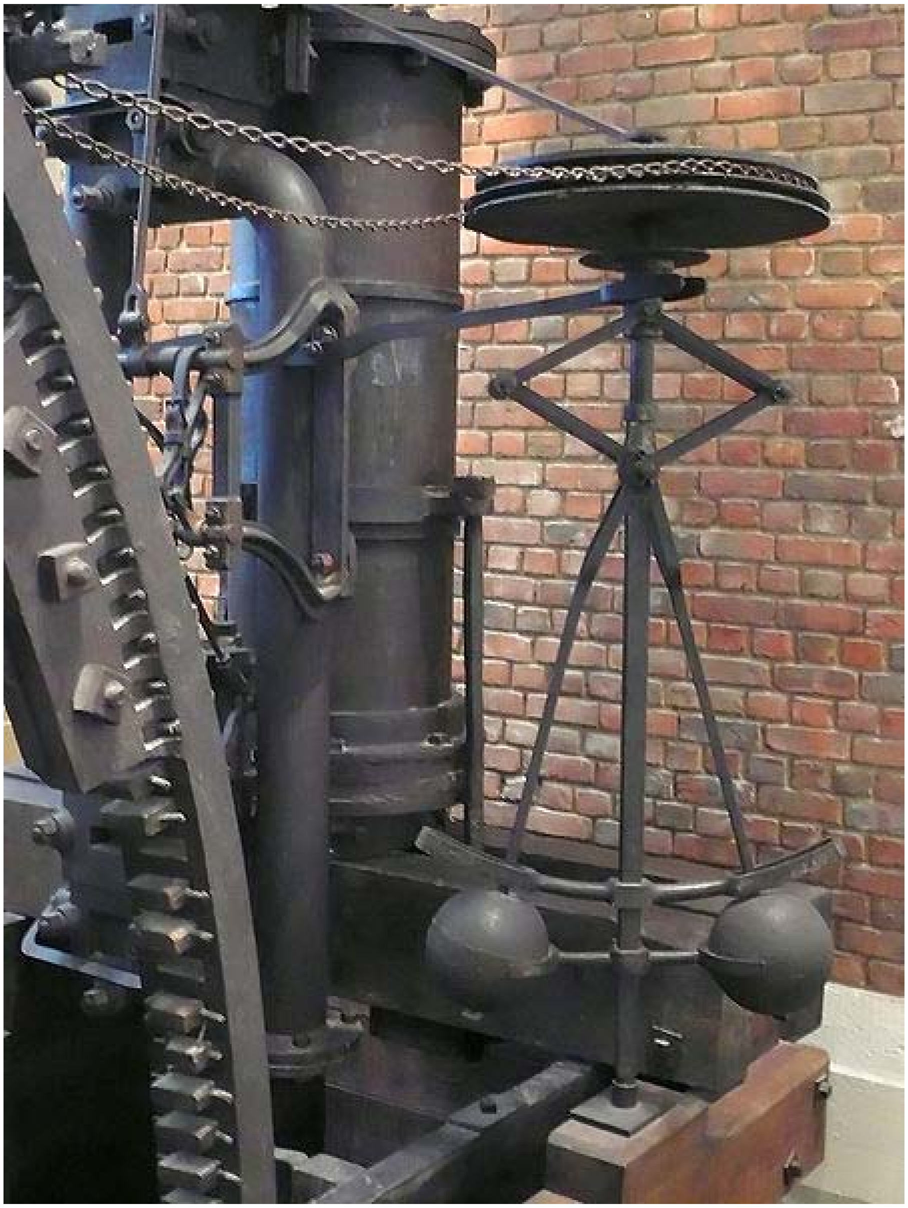
An example

Consider a car's cruise control, which is a device designed to maintain vehicle speed at a constant *desired* or *reference* speed provided by the driver. The *controller* is the cruise control, the *plant* is the car, and the *system* is the car and the cruise control. The system output is the car's speed, and the control itself is the engine's throttle position which determines how much power the engine generates.

A primitive way to implement cruise control is simply to lock the throttle position when the driver engages cruise control. However, if the cruise control is engaged on a stretch of flat road, then the car will travel slower going uphill and faster when going downhill. This type of controller is called an open-loop controller because no measurement of the system output (the car's speed) is used to alter the control (the throttle position.) As a result, the controller can not compensate for changes acting on the car, like a change in the slope of the road.

In a **closed-loop control system**, a sensor monitors the system output (the car's speed) and feeds the data to a controller which adjusts the control (the throttle position) as necessary to maintain the desired system output (match the car's speed to the reference speed.) Now when the car goes uphill the decrease in speed is measured, and the throttle position changed to increase engine power, speeding the vehicle. Feedback from measuring the car's speed has allowed the controller to dynamically compensate for changes to the car's speed. It is from this feedback that the paradigm of the control *loop* arises: the control affects the system output, which in turn is measured and looped back to alter the control.

History



Centrifugal governor in a Boulton & Watt engine of 1788

Although control systems of various types date back to antiquity, a more formal analysis of the field began with a dynamics analysis of the centrifugal governor, conducted by the physicist James Clerk Maxwell in 1868 entitled *On Governors*. This described and analyzed the phenomenon of "hunting", in which lags in the system can lead to overcompensation and unstable behavior. This generated a flurry of interest in the topic, during which Maxwell's classmate Edward John Routh generalized the results of Maxwell for the general class of linear systems. Independently, Adolf Hurwitz analyzed system stability using differential equations in 1877, resulting in what is now known as the Routh-Hurwitz theorem.

A notable application of dynamic control was in the area of manned flight. The Wright Brothers made their first successful test flights on December 17, 1903 and were distinguished by their ability to control their flights for substantial periods (more so than the ability to produce lift from an airfoil, which was known). Control of the airplane was necessary for safe flight.

By World War II, control theory was an important part of fire-control systems, guidance systems and electronics. The Space Race also depended on accurate spacecraft control. However, control theory also saw an increasing use in fields such as economics.

People in systems and control

Many active and historical figures made significant contribution to control theory, including, for example:

- Alexander Lyapunov (1857–1918) in the 1890s marks the beginning of stability theory.
- Harold S. Black (1898–1983), invented the concept of negative feedback amplifiers in 1927. He managed to develop stable negative feedback amplifiers in the 1930s.
- Harry Nyquist (1889–1976), developed the Nyquist stability criterion for feedback systems in the 1930s.
- Richard Bellman (1920–1984), developed dynamic programming since the 1940s.
- Andrey Kolmogorov (1903–1987) co-developed the Wiener-Kolmogorov filter (1941).
- Norbert Wiener (1894–1964) co-developed the Wiener-Kolmogorov filter and coined the term cybernetics in the 1940s.
- John R. Ragazzini (1912–1988) introduced digital control and the z-transform in the 1950s.
- Lev Pontryagin (1908–1988) introduced the maximum principle and the bang-bang principle.

Classical control theory

To avoid the problems of the open-loop controller, control theory introduces feedback. A closed-loop controller uses feedback to control states or outputs of a dynamical system. Its name comes from the information path in the system: process inputs (e.g. voltage applied to an electric motor) have an effect on the process outputs (e.g. velocity or torque of the motor), which is measured with sensors and processed by the controller; the result (the control signal) is used as input to the process, closing the loop.

Closed-loop controllers have the following advantages over open-loop controllers:

- disturbance rejection (such as unmeasured friction in a motor)
- guaranteed performance even with model uncertainties, when the model structure does not match perfectly the real process and the model parameters are not exact

- unstable processes can be stabilized
- reduced sensitivity to parameter variations
- improved reference tracking performance

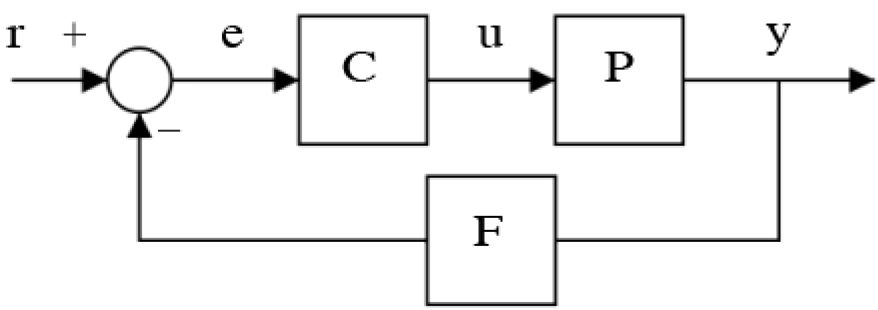
In some systems, closed-loop and open-loop control are used simultaneously. In such systems, the open-loop control is termed feedforward and serves to further improve reference tracking performance.

A common closed-loop controller architecture is the PID controller.

Closed-loop transfer function

The output of the system *y(t)* is fed back through a sensor measurement *F* to the reference value *r(t)*. The controller *C* then takes the error *e* (difference) between the reference and the output to change the inputs *u* to the system under control *P*. This is shown in the figure. This kind of controller is a closed-loop controller or feedback controller.

This is called a single-input-single-output (*SISO*) control system; *MIMO* (i.e. Multi-Input-Multi-Output) systems, with more than one input/output, are common. In such cases variables are represented through vectors instead of simple scalar values. For some distributed parameter systems the vectors may be infinite-dimensional (typically functions).



If we assume the controller *C*, the plant *P*, and the sensor *F* are linear and time-invariant (i.e.: elements of their transfer function *C(s)*, *P(s)*, and *F(s)* do not depend on time), the systems above can be analysed using the Laplace transform on the variables. This gives the following relations:

$$Y(s) = P(s)U(s)$$
$$U(s) = C(s)E(s)$$
$$E(s) = R(s) - F(s)Y(s).$$

Solving for *Y*(*s*) in terms of *R*(*s*) gives:

$$Y(s) = \left(\frac{P(s)C(s)}{1 + F(s)P(s)C(s)}\right) R(s) = H(s)R(s).$$

The expression $H(s) = \frac{P(s)C(s)}{1 + F(s)P(s)C(s)}$ is referred to as the *closed-loop transfer function* of the system. The numerator is the forward (open-loop) gain from *r* to *y*, and the denominator is one plus the gain in going around the feedback loop, the so-called loop gain. If $|P(s)C(s)| \gg 1$, i.e. it has a large norm with each value of *s*, and if $|F(s)| \approx 1$, then *Y(s)* is approximately equal to *R(s)* and the output closely tracks the reference input.

PID controller

The PID controller is probably the most-used feedback control design. *PID* is an acronym for *Proportional-Integral-Differential*, referring to the three terms operating on the error signal to produce a control signal. If *u(t)* is the control signal sent to the system, *y(t)* is the measured output and *r(t)* is the desired output, and tracking error $e(t) = r(t) - y(t)$, a PID controller has the general form

$$u(t) = K_P e(t) + K_I \int e(t) \mathrm{d}t + K_D \frac{\mathrm{d}}{\mathrm{d}t} e(t).$$

The desired closed loop dynamics is obtained by adjusting the three parameters K_P, K_I and K_D, often iteratively by "tuning" and without specific knowledge of a plant model. Stability can often be ensured using only the proportional term. The integral term permits the rejection of a step disturbance (often a striking specification in process control). The derivative term is used to provide damping or shaping of the response. PID controllers are the most well established class of control systems: however, they cannot be used in several more complicated cases, especially if MIMO systems are considered.

Applying Laplace transformation results in the transformed PID controller equation

$$u(s) = K_P e(s) + K_I \frac{1}{s} e(s) + K_D s e(s)$$
$$u(s) = (K_P + K_I \frac{1}{s} + K_D s) e(s)$$

with the PID controller transfer function

$$C(s) = (K_P + K_I \frac{1}{s} + K_D s).$$

Modern control theory

In contrast to the frequency domain analysis of the classical control theory, modern control theory utilizes the time-domain state space representation, a mathematical model of a physical system as a set of input, output and state variables related by first-order differential equations. To abstract from the number of inputs, outputs and states, the variables are expressed as vectors and the differential and algebraic equations are written in matrix form (the latter only being possible when the dynamical system is linear). The state space representation (also known as the "time-domain approach") provides a convenient and compact way to model and analyze systems with multiple inputs and outputs. With inputs and outputs, we would otherwise have to write down Laplace transforms to encode all the information about a system. Unlike the frequency domain approach, the use of the state space representation is not limited to systems with linear components and zero initial conditions. "State space" refers to the space whose axes are the state variables. The state of the system can be represented as a vector within that space.)

Topics in control theory

Stability

The *stability* of a general dynamical system with no input can be described with Lyapunov stability criteria. A linear system that takes an input is called bounded-input bounded-output (BIBO) stable if its output will stay bounded for any bounded input. Stability for nonlinear systems that take an input is input-to-state stability (ISS), which combines Lyapunov stability and a notion similar to BIBO stability. For simplicity, the following descriptions focus on continuous-time and discrete-time linear systems.

Mathematically, this means that for a causal linear system to be stable all of the poles of its transfer function must satisfy some criteria depending on whether a continuous or discrete time analysis is used:

- In continuous time, the Laplace transform is used to obtain the transfer function. A system is stable if the poles of this transfer function lie strictly in the open left half of the complex plane (i.e. the real part of all the poles is less than zero).
- In discrete time the Z-transform is used. A system is stable if the poles of this transfer function lie strictly inside the unit circle. i.e. the magnitude of the poles is less than one).

When the appropriate conditions above are satisfied a system is said to be asymptotically stable: the variables of an asymptotically stable control system always decrease from their initial value and do not show permanent oscillations. Permanent oscillations occur when a pole has a real part exactly equal to zero (in the continuous time case) or a modulus equal to one (in the discrete time case). If a simply stable system response neither decays nor grows over time, and has no oscillations, it is marginally stable: in this case the system transfer function has non-repeated poles at complex plane origin (i.e. their real and complex component is zero in the continuous time case). Oscillations are present when poles with real part equal to zero have an imaginary part not equal to zero.

Differences between the two cases are not a contradiction. The Laplace transform is in Cartesian coordinates and the Z-transform is in circular coordinates, and it can be shown that:

- the negative-real part in the Laplace domain can map onto the interior of the unit circle
- the positive-real part in the Laplace domain can map onto the exterior of the unit circle

If a system in question has an impulse response of

$$x[n] = 0.5^n u[n]$$

then the Z-transform (see this example), is given by

$$X(z) = \frac{1}{1 - 0.5z^{-1}}$$

which has a pole in $z = 0.5$ (zero imaginary part). This system is BIBO (asymptotically) stable since the pole is *inside* the unit circle.

However, if the impulse response was

$$x[n] = 1.5^n u[n]$$

then the Z-transform is

$$X(z) = \frac{1}{1 - 1.5z^{-1}}$$

which has a pole at $z = 1.5$ and is not BIBO stable since the pole has a modulus strictly greater than one.

Numerous tools exist for the analysis of the poles of a system. These include graphical systems like the root locus, Bode plots or the Nyquist plots.

Mechanical changes can make equipment (and control systems) more stable. Sailors add ballast to improve the stability of ships. Cruise ships use antiroll fins that extend transversely from the side of the ship for perhaps 30 feet (10 m) and are continuously rotated about their axes to develop forces that oppose the roll.

Controllability and observability

Controllability and observability are main issues in the analysis of a system before deciding the best control strategy to be applied, or whether it is even possible to control or stabilize the system. Controllability is related to the possibility of forcing the system into a particular state by using an appropriate control signal. If a state is not controllable, then no signal will ever be able to control the state. If a state is not controllable, but its dynamics are stable, then the state is termed Stabilizable. Observability instead is related to the possibility of "observing", through output measurements, the state of a system. If a state is not observable, the controller will never be able to determine the behaviour of an unobservable state and hence cannot use it to stabilize the system. However, similar to the stabilizability condition above, if a state cannot be observed it might still be detectable.

From a geometrical point of view, looking at the states of each variable of the system to be controlled, every "bad" state of these variables must be controllable and observable to ensure a good behaviour in the closed-loop system. That is, if one of the eigenvalues of the system is not both controllable and observable, this part of the dynamics will remain untouched in the closed-loop system. If such an eigenvalue is not stable, the dynamics of this eigenvalue will be present in the closed-loop system which therefore will be unstable. Unobservable poles are not present in the transfer function realization of a state-space representation, which is why sometimes the latter is preferred in dynamical systems analysis.

Solutions to problems of uncontrollable or unobservable system include adding actuators and sensors.

Control specification

Several different control strategies have been devised in the past years. These vary from extremely general ones (PID controller), to others devoted to very particular classes of systems (especially robotics or aircraft cruise control).

A control problem can have several specifications. Stability, of course, is always present: the controller must ensure that the closed-loop system is stable, regardless of the open-loop stability. A poor choice of controller can even worsen the stability of the open-loop system, which must normally be avoided. Sometimes it would be desired to obtain particular dynamics in the closed loop: i.e. that the poles have $Re[\lambda] < -\overline{\lambda}$, where $\overline{\lambda}$ is a fixed value strictly greater than zero, instead of simply asking that $Re[\lambda] < 0$.

Another typical specification is the rejection of a step disturbance; including an integrator in the open-loop chain (i.e. directly before the system under control) easily achieves this. Other classes of disturbances need different types of sub-systems to be included.

Other "classical" control theory specifications regard the time-response of the closed-loop system: these include the rise time (the time needed by the control system to reach the desired value after a perturbation), peak overshoot (the highest value reached by the response before reaching the desired value) and others (settling time, quarter-decay). Frequency domain specifications are usually related to robustness.

Modern performance assessments use some variation of integrated tracking error (IAE,ISA,CQI).

Model identification and robustness

A control system must always have some robustness property. A robust controller is such that its properties do not change much if applied to a system slightly different from the mathematical one used for its synthesis. This specification is important: no real physical system truly behaves like the series of differential equations used to represent it mathematically. Typically a simpler mathematical model is chosen in order to simplify calculations, otherwise the true system dynamics can be so complicated that a complete model is impossible.

System identification

The process of determining the equations that govern the model's dynamics is called system identification. This can be done off-line: for example, executing a series of measures from which to calculate an approximated mathematical model, typically its transfer function or matrix. Such identification from the output, however, cannot take account of unobservable dynamics. Sometimes the model is built directly starting from known physical equations: for example, in the case of a mass-spring-damper system we know that $m\ddot{x}(t) = -Kx(t) - B\dot{x}(t)$. Even assuming that a "complete" model is used in designing the controller, all the parameters included in these equations (called "nominal parameters") are never known with absolute precision; the control system will have to behave correctly even when connected to physical system with true parameter values away from nominal.

Some advanced control techniques include an "on-line" identification process. The parameters of the model are calculated ("identified") while the controller itself is running: in this way, if a drastic variation of the parameters ensues (for example, if the robot's arm releases a weight), the controller will adjust itself consequently in order to ensure the correct performance.

Analysis

Analysis of the robustness of a SISO control system can be performed in the frequency domain, considering the system's transfer function and using Nyquist and Bode diagrams. Topics include gain and phase margin and amplitude margin. For MIMO and, in general, more complicated control systems one must consider the theoretical results devised for each control technique: i.e., if particular robustness qualities are needed, the engineer must shift his attention to a control technique including them in its properties.

Constraints

A particular robustness issue is the requirement for a control system to perform properly in the presence of input and state constraints. In the physical world every signal is limited. It could happen that a controller will send control signals that cannot be followed by the physical system: for example, trying to rotate a valve at excessive speed. This can produce undesired behavior of the closed-loop system, or even break actuators or other subsystems. Specific control techniques are available to solve the problem: model predictive control, and anti-wind up systems. The latter consists of an additional control block that ensures that the control signal never exceeds a given threshold.

System classifications

Linear Systems control

For MIMO systems, pole placement can be performed mathematically using a state space representation of the open-loop system and calculating a feedback matrix assigning poles in the desired positions. In complicated systems this can require computer-assisted calculation capabilities, and cannot always ensure robustness. Furthermore, all system states are not in general measured and so observers must be included and incorporated in pole placement design.

Nonlinear Systems control

Processes in industries like robotics and the aerospace industry typically have strong nonlinear dynamics. In control theory it is sometimes possible to linearize such classes of systems and apply linear techniques, but in many cases it can be necessary to devise from scratch theories permitting control of nonlinear systems. These, e.g., feedback linearization, backstepping, sliding mode control, trajectory linearization control normally take advantage of results based on Lyapunov's theory. Differential geometry has been widely used as a tool for generalizing well-known linear control concepts to the non-linear case, as well as showing the subtleties that make it a more challenging problem.

Decentralized Systems

When the system is controlled by multiple controllers, the problem is one of decentralized control. Decentralization is helpful in many ways, for instance, it helps control systems operate over a larger geographical area. The agents in decentralized control systems can interact using communication channels and coordinate their actions.

Main control strategies

Every control system must guarantee first the stability of the closed-loop behavior. For linear systems, this can be obtained by directly placing the poles. Non-linear control systems use specific theories (normally based on Aleksandr Lyapunov's Theory) to ensure stability without regard to the inner dynamics of the system. The possibility to fulfill different specifications varies from the model considered and the control strategy chosen. Here a summary list of the main control techniques is shown:

Adaptive control
: Adaptive control uses on-line identification of the process parameters, or modification of controller gains, thereby obtaining strong robustness properties. Adaptive controls were applied for the first time in the aerospace industry in the 1950s, and have found particular success in that field.

Hierarchical control
: A Hierarchical control system is a type of Control System in which a set of devices and governing software is arranged in a hierarchical tree. When the links in the tree are implemented by a computer network, then that hierarchical control system is also a form of Networked control system.

Intelligent control
: Intelligent control uses various AI computing approaches like neural networks, Bayesian probability, fuzzy logic, machine learning, evolutionary computation and genetic algorithms to control a dynamic system.

Optimal control
: Optimal control is a particular control technique in which the control signal optimizes a certain "cost index": for example, in the case of a satellite, the jet thrusts needed to bring it to desired trajectory that consume the least amount of fuel. Two optimal control design methods have been widely used in industrial applications, as it has been shown they can guarantee closed-loop stability. These are Model Predictive Control (MPC) and Linear-Quadratic-Gaussian control (LQG). The first can more explicitly take into account constraints on the signals in the system, which is an important feature in many industrial processes. However, the "optimal control" structure in MPC is only a means to achieve such a result, as it does not optimize a true performance index of the closed-loop control system. Together with PID controllers, MPC systems are the most widely used control technique in process control.

Robust control
: Robust control deals explicitly with uncertainty in its approach to controller design. Controllers designed using *robust control* methods tend to be able to cope with small differences between the true system and the nominal model used for design. The early methods of Bode and others were fairly robust; the state-space methods invented in the 1960s and 1970s were sometimes found to lack robustness. A modern example of a robust control technique is H-infinity loop-shaping developed by Duncan McFarlane and Keith Glover of Cambridge University, United Kingdom. Robust methods aim to achieve robust performance and/or stability in the presence of small modeling errors.

Stochastic control
: Stochastic control deals with control design with uncertainty in the model. In typical stochastic control problems, it is assumed that there exist random noise and disturbances in the model and the controller, and the control design must take into account these random deviations.