Vacuum Engineering

Mercy Madrigal

First Edition, 2012

ISBN 978-81-323-4258-8

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

Table of Contents

Chapter 1

Ultra-High Vacuum

Ultra-high vacuum (UHV) is the vacuum regime characterised by pressures lower than about 10^{-7} pascal or 100 nanopascals (10^{-9} mbar, ~10^{-9} torr). UHV requires the use of unusual materials in construction and by heating the entire system to 180°C for several hours ("baking") to remove water and other trace gases which adsorb on the surfaces of the chamber. At these low pressures the mean free path of a gas molecule is approximately 40 km, so gas molecules will collide with the chamber walls many times before colliding with each other. Almost all interactions therefore take place on various surfaces in the chamber.

Concepts involved

- Sorption of gases
- Kinetic theory of gases
- Gas transport and pumping
- Vacuum pumps and systems
- Vapour pressure

Material limitations

Materials which are not allowed due to high vapour pressure:

- majority of organic compounds cannot be used:
    - plastics other than teflon and PEEK: gaskets are made of copper, and are single-use; plastics in other uses are replaced with ceramics or metals
    - glues: special glues for high vacuum must be used
- common steel: due to oxidizing, which greatly increases adsorption area, only stainless steel is used
- lead: soldering is performed using lead-free solder
- indium: Indium is commonly used as a deformable gasket material for vacuum seals, especially in cryogenic apparatus, but its low melting point prevents use in baked systems.
- zinc, cadmium: High vapor pressures during system bake-out.

Technical limitations:

- screws: threads have a high surface area and tend to "trap" gases, therefore are avoided
- welding: standard welding cannot be used due to high surface area and introduction of gas chambers, which would collect gas at atmospheric pressure, and release it slowly during evacuation (removal of gas).

Typical uses for ultra-high vacuum

Ultra-high vacuum is necessary for many surface analytic techniques such as:

- X-ray photoelectron spectroscopy (XPS)
- Auger electron spectroscopy (AES)
- Secondary ion mass spectrometry (SIMS)
- Thermal desorption spectroscopy (TPD)
- Thin film growth and preparation techniques with stringent requirements for purity, such as molecular beam epitaxy (MBE), UHV chemical vapor deposition (CVD) and UHV pulsed laser deposition (PLD)
- Angle resolved photoemission spectroscopy (ARPES)

UHV is necessary for these applications to reduce surface contamination, by reducing the number of molecules reaching the sample over a given time period. At 0.1 mPa (10^{-6} Torr), it only takes 1 second to cover a surface with a contaminant, so much lower pressures are needed for long experiments.

UHV is also required for:

- Particle accelerators
- Gravitational wave detectors such as LIGO, VIRGO, GEO 600, and TAMA 300.
- Atomic physics experiments which use cold atoms, such as ion trapping or making Bose-Einstein condensates

and, while not compulsory, can prove beneficial in applications such as:

- Atomic force microscopy. High vacuum enables high Q factors on the cantilever oscillation.
- Scanning tunneling microscopy. High vacuum reduces oxidation and contamination, hence enables imaging and the achievement of atomic resolution on clean metal and semiconductor surfaces, e.g. imaging the surface reconstruction of the unoxidized silicon surface.

Achieving ultra-high vacuum

Extraordinary steps are required to reach UHV, including the following:

- High pumping speed — possibly multiple vacuum pumps in series and/or parallel
- Minimize surface area in the chamber
- High conductance tubing to pumps — short and fat, without obstruction
- Use low-outgassing materials such as certain stainless steels
- Avoid creating pits of trapped gas behind bolts, welding voids, etc.
- Electropolish all metal parts after machining or welding
- Use low vapor pressure materials (ceramics, glass, metals, teflon if unbaked)
- Bake the system to remove water or hydrocarbons adsorbed to the walls
- Chill chamber walls to cryogenic temperatures during use
- Avoid all traces of hydrocarbons, including skin oils in a fingerprint — always use gloves

Outgassing is a significant problem for UHV systems. Outgassing can occur from two sources: surfaces and bulk materials. Outgassing from bulk materials is minimized by careful selection of materials with low vapor pressures (such as glass, stainless steel, and ceramics) for everything inside the system. Even materials which are not generally considered absorbent can outgas, including most plastics and some metals. For example, vessels lined with a highly gas-permeable material such as palladium (which is a high-capacity hydrogen sponge) create special outgassing problems.

Outgassing from surfaces is a subtler problem. At extremely low pressures, more gas molecules are adsorbed on the walls than are floating in the chamber, so the total surface area inside a chamber is more important than its volume for reaching UHV. Water is a significant source of outgassing because a thin layer of water vapor rapidly adsorbs to everything whenever the chamber is opened to air. Water evaporates from surfaces too slowly to be fully removed at room temperature, but just fast enough to present a continuous level of background contamination. Removal of water and similar gases generally requires baking the UHV system at 200 to 400 °C while vacuum pumps are running. During chamber use, the walls of the chamber may be chilled using liquid nitrogen to reduce outgassing further.

Hydrogen and helium are the most common background gases in a well-designed, well-baked UHV system. Hydrogen diffuses out from the grain boundaries in stainless steel. Helium can diffuse through the steel and glass from the outside air.

There is no single vacuum pump that can operate all the way from atmospheric pressure to ultra-high vacuum. Instead, a series of different pumps is used, according to the appropriate pressure range for each pump. Pumps commonly used to achieve UHV include:

- Turbomolecular pumps (especially compound and/or magnetic bearing types)
- Ion pumps

- Titanium sublimation pumps
- Non-evaporable getter (NEG) pumps
- Cryopumps

UHV pressures are measured with an ion gauge, either a hot filament or an inverted magnetron type.

Finally, special seals and gaskets must be used between components in a UHV system to prevent even trace leakage. Nearly all such seals are all metal, with knife edges on both sides cutting into a soft, copper gasket. This all-metal seal can maintain pressures down to 100 pPa ($\sim 10^{-12}$ Torr).

Measuring high vacuum

Measurement of high vacuum is done using a *nonabsolute gauge* that measures a pressure-related property of the vacuum, for example, its thermal conductivity. See, for example, Pacey. These gauges must be calibrated. The gauges capable of the measuring the lowest pressures are magnetic gauges based upon the pressure dependence of the current in a spontaneous gas discharge in intersecting electric and magnetic fields.

UHV manipulator

A **UHV manipulator** allows an object which is inside a vacuum chamber and under vacuum to be mechanically positioned. It may provide rotary motion, linear motion, or a combination of both. The most complex devices give motion in three axes and rotations around two of those axes. To generate the mechanical movement inside the chamber, two basic mechanisms are commonly employed: a mechanical coupling through the vacuum wall (using a vacuum-tight seal around the coupling), or a magnetic coupling that transfers motion from air-side to vacuum-side. Various forms of motion control are available for manipulators, such as knobs, handwheels, motors, stepping motors, piezoelectric motors, and pneumatics.

The manipulator or sample holder may include features which allow additional control and testing of a sample, such as the ability to apply heat, cooling, voltage, or a magnetic field. Sample heating can be accomplished by electron bombardment or thermal radiation. For electron bombardment, the sample holder is equipped with a filament which emits electrons when biased at a high negative potential. The impact of the electrons bombarding the sample at high energy causes it to heat. For thermal radiation, a filament is mounted close to the sample and resistively heated to high temperature. The infrared energy from the filament heats the sample.

Chapter 2

Vacuum Flange

A KF-25 tee, o-ring, and clamp.

A **vacuum flange** is a flange at the end of a tube used to connect vacuum chambers, tubing and vacuum pumps to each other.

Vacuum flange types

A CF (conflat) pipe and flange with copper gasket.

Several vacuum flange standards exist, and the same flange types are called by different names by different manufacturers and standards organizations.

KF/QF

The ISO standard quick release flange is known by the names Quick Flange (QF), Klein Flange (KF) or NW, sometimes also as DN. The KF designation has been adopted by ISO, DIN, and Pneurop. KF flanges are made with a chamferred back surface that attached with a circular clamp and an elastomeric o-ring that is mounted in a metal centering ring. Standard sizes are indicated by the nominal inner diameter in millimeters for flanges 16 through 50 mm in diameter.

- DN16KF
- DN25KF
- DN40KF
- DN50KF

ISO

The ISO large flange standard is known as LF, LFB, MF or sometimes just ISO flange. As in KF-flanges, the flanges are joined by a centering ring and an elastomeric o-ring. An extra spring-loaded circular clamp is often used around the large diameter o-rings to prevent them from rolling off from the centering ring during mounting.

The ISO large flanges come in two varieties. The ISO-K (or ISO LF) flanges are joined with double claw clamps which clamp to a circular groove on the tubing side of the flange. The ISO-F (or ISO LFB) flanges have holes for attaching the two flanges with bolts. Two tubes with ISO-K and ISO-F flanges can be joined together by clamping the ISO-K side with single claw clamps which are then bolted to the holes on the ISO-F side.

ISO large flanges are available in sizes from 63 to 500 mm nominal tube diameter).

- DN63LF (63.5 mm)
- DN100LF (102 mm)
- DN160LF (160 mm)
- DN200LF (200 mm)
- DN250LF (254 mm)
- DN320LF (316 mm)
- DN400LF (400 mm)
- DN500LF (500 mm)

CF (Conflat)

CF (ConFlat) flanges use a copper gasket and knife-edge flange to achieve an ultrahigh vacuum seal. The term "ConFlat" is a registered trademark of Varian, Inc., so "CF" is commonly used by other flange manufacturers. Each face of the two mating CF flanges has a knife edge which cuts into the softer metal gasket, providing an extremely leak-tight, metal-to-metal seal. Deformation of the metal gasket fills small defects in the flange, allowing Conflat flanges operate down to 10^{-13} torr (10^{-11} Pa) pressure. The gasket is partially recessed in a groove in each flange. The groove helps hold the gasket in place, which aligns the two flanges and also reduces gasket expansion during bake-out. For stainless steel conflat flanges baking temperatures of 450°C can be achieved; the temperature is limited by the choice of gasket material. CF flanges are sexless and interchangeable. North American flange sizes are given by flange outer diameter in inches while in Europe and Asia, sizes are given by tube inner diameter in millimeters:

European, Asian size	North American size [inches]
DN10	1
DN16	1⅓ ("mini")
DN25	2⅛
DN40 (or: DN35)	2¾
DN50	3⅜

DN63	4½
DN75	4⅝
DN100	6
DN125	6¾
DN160 (or: DN150)	8
DN200	10
DN250	12
	13¼
	14
	16½

A 60 kV high voltage electrical feedthrough on a 4-1/2 inch (or DN63) conflat flange

ConFlat gaskets were originally invented by William Wheeler and other engineers at Varian in an attempt to build a flange that would not leak after baking.

Wheeler

A Wheeler flange is a large wire seal flange often used on large vacuum chambers.

ASA

ANSI has a flange standard called ASA. These flanges are elastomeric o-ring seal and can be used for both vacuum and pressure applications. Flange sizes are indicated by tube nominal inner diameter or by flange outer diameter (in inches): 1 (4.25 O.D.), 1.5 (5.00 O.D.), 2 (6.00 O.D.), 3 (7.50 O.D.), 4 (9.00 O.D.), 6 (11.00 O.D.), 8 (13.5 O.D.), 10 (16.00 O.D.).

Vacuum gaskets

To achieve a vacuum seal, a gasket is required. An elastomeric o-ring gasket can be made of Buna rubber, viton fluoropolymer, silicon rubber or teflon. O-rings can be placed in a groove or may be used in combination with a centering ring or as a "captured" o-ring that is held in place by separate metal rings. Metal gaskets are used in ultra-high vacuum systems where the outgassing of the elastomer could be a significant gas load. A copper ring gasket is used with conflat flanges. Metal wire gaskets made of copper, gold or indium can be used.

Vacuum feedthrough

A vacuum feedthrough is a flange that contains a vacuum-tight electrical, physical or mechanical connection to the vacuum chamber. An electrical feedthrough allows voltages to be applied to components under vacuum, for example a filament or heater. An example of a physical feedthrough is a vacuum tight connection for cooling water. A mechanical feedthrough is used for rotation and translation of components under vacuum. A wobble stick is a mechanical feedthrough device that can be used to pick up, move and otherwise manipulate objects in the vacuum chamber.

Chapter 3

Vacuum Packing and Vacuum Evaporation

Vacuum packing

Vacuum packing is a method of storing food and presenting it for sale. Appropriate types of food are stored in an airless environment, usually in an air-tight pack or bottle to prevent the growth of microorganisms. The vacuum environment removes atmospheric oxygen, protecting the food from spoiling by limiting the growth of aerobic bacteria or fungi, and preventing the evaporation of volatile components. Vacuum packing is commonly used for long-term storage of dry foods such as cereals, nuts, cured meats, cheese, smoked fish, coffee, and potato chips (crisps). It is also for storage of fresh foods such as vegetables, meats, and liquids such as soups in a shorter term because vacuum condition cannot stop bacteria from getting water which can promote their growth. Vacuum packaging food can extend its life by up to 3-5 times.

Vacuum packing is also used to reduce greatly the bulk of non-food items. For example, clothing and bedding can be stored in bags evacuated with a domestic vacuum cleaner or a dedicated vacuum sealer. This technique is sometimes used to compact household waste, for example where a charge is made for each full bag collected. Vacuum packing can be used to reduce bulk of inflatable items as well.

Vacuum packaging products using plastic bags, canisters, bottles, or mason jars are available for home use.

Vacuum packaging delicate food items can be done by using an inert gas kit, typically available on chamber vacuum sealers. After air has been removed, an inert gas (such as nitrogen) is added to maintain the preservation of packaged food while preventing damage. An example of inert gas for packaging delicate foods is potato chips.

External Sealers

External vacuum sealers involve a bag being attached to the vacuum-sealing machine externally. The machine will remove the air and seal the bag, which is all done outside the machine.

Chamber Sealers

Chamber sealers require the entire product to be placed within the machine. Like external sealers, a plastic bag is typically used for packaging. Once the product is placed in the machine, the lid is closed and air is removed. Once the air is removed, the bag is sealed and the atmosphere within the chamber is returned back to normal. The lid is then opened and the product removed. Chamber sealers are typically used for higher-volume packaging.

Manufacturers of chamber type vacuum packing machines include: Cryovac, Multivac, Sammic, VC999, New Diamond Vac , Jaw Feng Machinery ,Sevana and several others.

Preventing Freezer Burn

When foods are frozen without preparation, freezer burn can occur. It happens when the surface of the food is dehydrated, and this leads to a dried and leathery appearance. Freezer burn also ruins the flavor and texture of foods. Vacuum packing prevents freezer burn by preventing the food from exposure to the cold, dry air.

Sous-vide Cooking

Vacuum packaging also allows for a special cooking method, Sous-vide. Sous-vide, meaning "under vacuum" in French, involves poaching food that is vacuum sealed in a plastic bag.

Security

Due to an oxygen-poor environment, anaerobic microorganism can proliferate, so vacuum packing is often used in combination with other treatment.

Vacuum evaporation

Vacuum evaporation is the process of causing the pressure in a liquid-filled container to be reduced below the vapor pressure of the liquid, causing the liquid to evaporate at a lower temperature than normal. Although the process can be applied to any type of liquid at any vapor pressure, it is generally used to describe the boiling of water by lowering the container's internal pressure below standard atmospheric pressure and causing the water to boil at room temperature.

When the process is applied to food and the water is evaporated and removed, the food can be stored for long periods of time without spoiling. It is also used when boiling a substance at normal temperatures would chemically change the consistency of the product, such as egg whites coagulating when attempting to dehydrate the albumen into a powder.

This process was invented by Henri Nestlé in 1866, of Nestlé Chocolate fame, although the Shakers were already using a vacuum pan earlier than that.

This process is used industrially to make such food products as evaporated milk for milk chocolate, and tomato paste for ketchup.

Vacuum evaporation is also a form of physical vapor deposition used in the semiconductor, microelectronics, and optical industries and in this context is a process of depositing thin films of material onto surfaces. Such a technique consists of pumping a vacuum chamber to pressures of less than 10^{-5} torr and heating a material to produce a flux of vapor in order to deposit the material onto a surface. The material to be vaporized is typically heated until its vapor pressure is high enough to produce a flux of several Angstroms per second by using an electrically resistive heater or bombardment by a high voltage beam.

Chapter 4

Vacuum Deposition and Vacuum Chamber

Vacuum deposition

Vacuum deposition is a family of processes used to deposit layers atom-by-atom or molecule-by-molecule at sub-atmospheric pressure (vacuum) on a solid surface. The layers may be as thin as one atom to millimeters thick (freestanding structures). There may be multiple layers of different materials (e.g. optical coatings). A thickness of less than one micrometre is generally called a thin film while a thickness greater than one micrometre is called a coating. The vacuum environment may serve one or more purposes including:

- reducing the particle density so that the mean free path for collision is long
- reducing the particle density of undesirable atoms and molecules (contaminants)
- providing a low pressure plasma environment
- providing a means for controlling gas and vapor composition
- providing a means for mass flow control into the processing chamber.

Condensing particles may come from a variety of sources including:

- thermal evaporation, Evaporation (deposition)
- sputtering
- cathodic arc vaporization
- laser ablation
- decomposition of a chemical vapor precursor, chemical vapor deposition

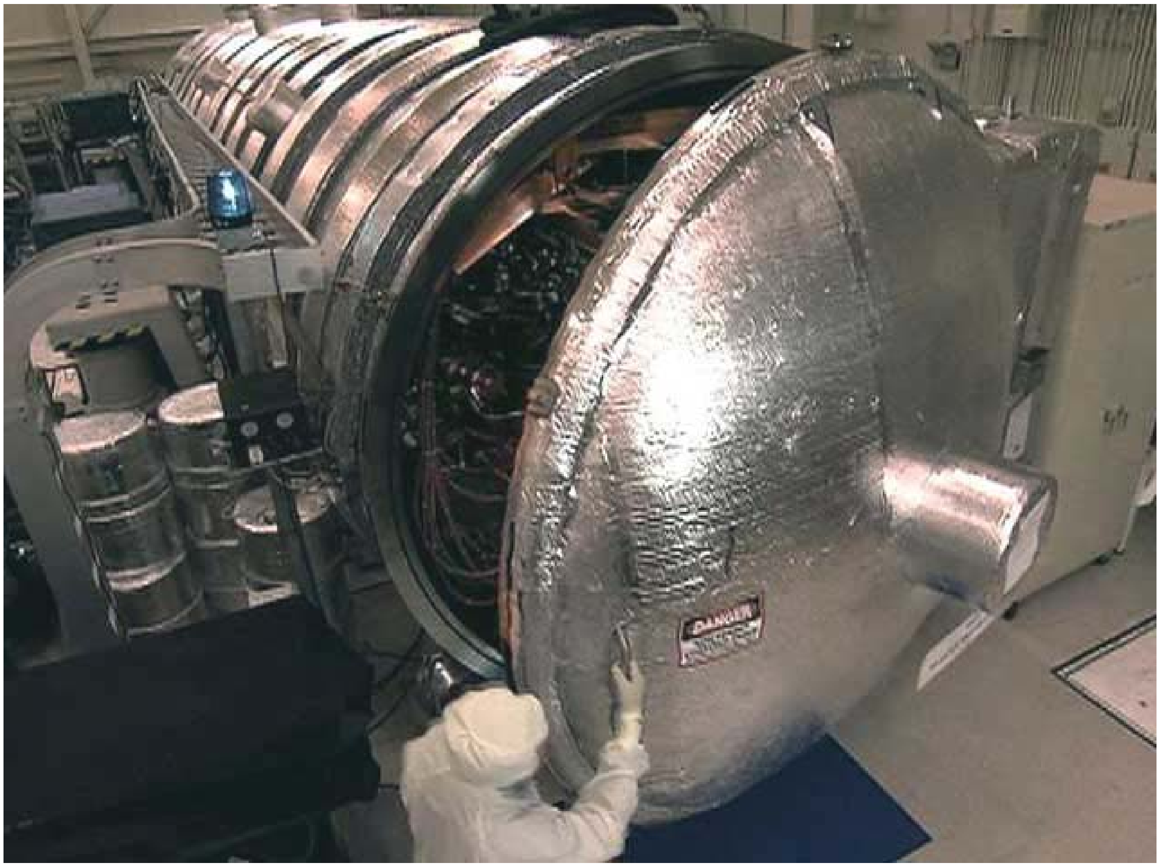
When the vapor source is from a liquid or solid material the process is called *physical vapor deposition* (PVD). When the source is from a chemical vapor precursor the process is called *low pressure chemical vapor deposition* (LPCVD) or, if in a plasma, *plasma enhanced CVD* (PECVD) or "plasma assisted CVD" (PACVD). Often a combination of PVD and CVD processes are used in the same or connected processing chambers.

In reactive deposition the depositing material reacts either with a component of the gaseous environment (Ti + N → TiN) or with a co-depositing species (Ti + C → TiC). A plasma environment aids in activating gaseous species ($N_2 \rightarrow 2N$) and in decomposition of chemical vapor precursors ($SiH_4 \rightarrow Si + 4H$). The plasma may also be used to provide ions for vaporization by sputtering or for bombardment of the substrate for sputter cleaning and for bombardment of the depositing material to densify the structure and tailor properties (ion plating).

Applications

- Electrical conduction: metallic films, transparent conductive oxides (TCO), superconducting films & coatings
- Semiconductor devices: semiconductor films, electrically insulating films
- Solar cells.
- Optical films: antireflective coatings, optical filters
- Reflective coatings: mirrors, heat mirrors
- Tribological coating: hard coatings, erosion resistant coatings, solid film lubricants
- Energy conservation & generation: low-E glass coatings, solar absorbing coatings, mirrors, solar thin film photovoltaic cells, *smart* films
- Magnetic films: magnetic recording
- Diffusion barrier: gas permeation barriers, vapor permeation barriers, solid state diffusion barriers
- Corrosion protection:
- Automotive applications: lamp reflectors and trim applications

Vacuum chamber



A large vacuum chamber.

A small vacuum chamber for studio or lab use in de-airing materials such as mold rubbers and resins.

Vacuum chamber for testing leaks in packaging

A **vacuum chamber** is a rigid enclosure from which air and other gases are removed by a vacuum pump. The resulting low pressure, commonly referred to as a vacuum, allows researchers to conduct physical experiments or to test mechanical devices which must operate in outer space (for example). Chambers are typically made of metals which may or may not shield applied external magnetic fields depending on wall thickness, frequency, resistivity, and permeability of the material used.

Chambers often have multiple ports, covered with vacuum flanges, to allow instruments or windows to be installed in the walls of the chamber. In low to medium-vacuum applications, these are sealed with rubber o-rings. In higher vacuum applications, the flanges have hardened steel knives welded onto them, which cut into a copper gasket when the flange is bolted on.

A type of vacuum chamber frequently used in the field of spacecraft engineering is a Thermal Vacuum Chamber, which provides a thermal environment representing what a spacecraft would experience in space.

Degassing Mold Making and Casting Materials To assure a bubble-free mold when mixing resin and silicone rubbers and slower setting harder resins, a vacuum chamber is required. A small vacuum chamber is needed for de-airing (eliminating air bubbles) for materials prior to their setting. The process is fairly straight forward. The casting or molding material is mixed according to the manufacturers directions.

The Process Since the material will expand 4-5 times under a vacuum, the mixing container must be large enough to hold a volume of four-five times the amount of the original material that is being vacuumed to allow for the expansion. If not it will spill over the top of the container requiring clean-up that can be avoided. The material container is then placed into the vacuum chamber; a vacuum pump is connected and turned on. Once the vacuum reaches 29-inches (at sea level) of mercury, the material will begin to rise (resembling foam). When the material falls it will plateau and not rise any more. The vacuuming is continued for another 2-3-minute to make certain all of the air has been removed from the material. Once this interval is reached the vacuum pump is shut off and the vacuum chamber release valve is opened to equalize air pressure. The vacuum chamber is opened and the material is removed and ready to pour into the mold.

To keep the material air free it must be slowly poured in a high and narrow stream starting from the corner of the mold box, or mold, letting the material to flow freely into the box or mold cavity. This method will usually not introduce any new bubbles into the vacuumed material. To insure that the material is totally devoid of air bubbles you can place the entire mold/mold box into the chamber for an additional few minutes. This will assist the material to flow into difficult areas of the mold/mold box.

High Altitude Vacuuming Though a vacuum of 29-inches of mercury is desired for de-airing most mold making and casting materials, it cannot be achieved at higher elevations. It is a sea level target. For example, in Denver (The mile-high city) only about 24-inches of mercury can be achieved even with the most efficient vacuum pump. In those instances, vacuuming will be required for a longer period to achieve proper de-gassing.

Chapter 5

Thermal Spraying

Thermal spraying techniques are coating processes in which melted (or heated) materials are sprayed onto a surface. The "feedstock" (coating precursor) is heated by electrical (plasma or arc) or chemical means (combustion flame).

Thermal spraying can provide thick coatings (approx. thickness range is 20 micrometers to several mm, depending on the process and feedstock), over a large area at high deposition rate as compared to other coating processes such as electroplating, physical and chemical vapor deposition. Coating materials available for thermal spraying include metals, alloys, ceramics, plastics and composites. They are fed in powder or wire form, heated to a molten or semimolten state and accelerated towards substrates in the form of micrometer-size particles. Combustion or electrical arc discharge is usually used as the source of energy for thermal spraying. Resulting coatings are made by the accumulation of numerous sprayed particles. The surface may not heat up significantly, allowing the coating of flammable substances.

Coating quality is usually assessed by measuring its porosity, oxide content, macro and micro-hardness, bond strength and surface roughness. Generally, the coating quality increases with increasing particle velocities.

Several variations of thermal spraying are distinguished:

- Plasma spraying

- Detonation spraying
- Wire arc spraying
- Flame spraying
- High-velocity oxy-fuel coating spraying (HVOF)
- Warm spraying
- Cold spraying

In classical (developed between 1910 and 1920) but still widely used processes such as flame spraying and wire arc spraying, the particle velocities are generally low (< 150 m/s), and raw materials must be molten to be deposited. Plasma spraying, developed in the 1970s, uses a high-temperature plasma jet generated by arc discharge with typical temperatures >15000 K, which makes it possible to spray refractory materials such as oxides, molybdenum, etc.

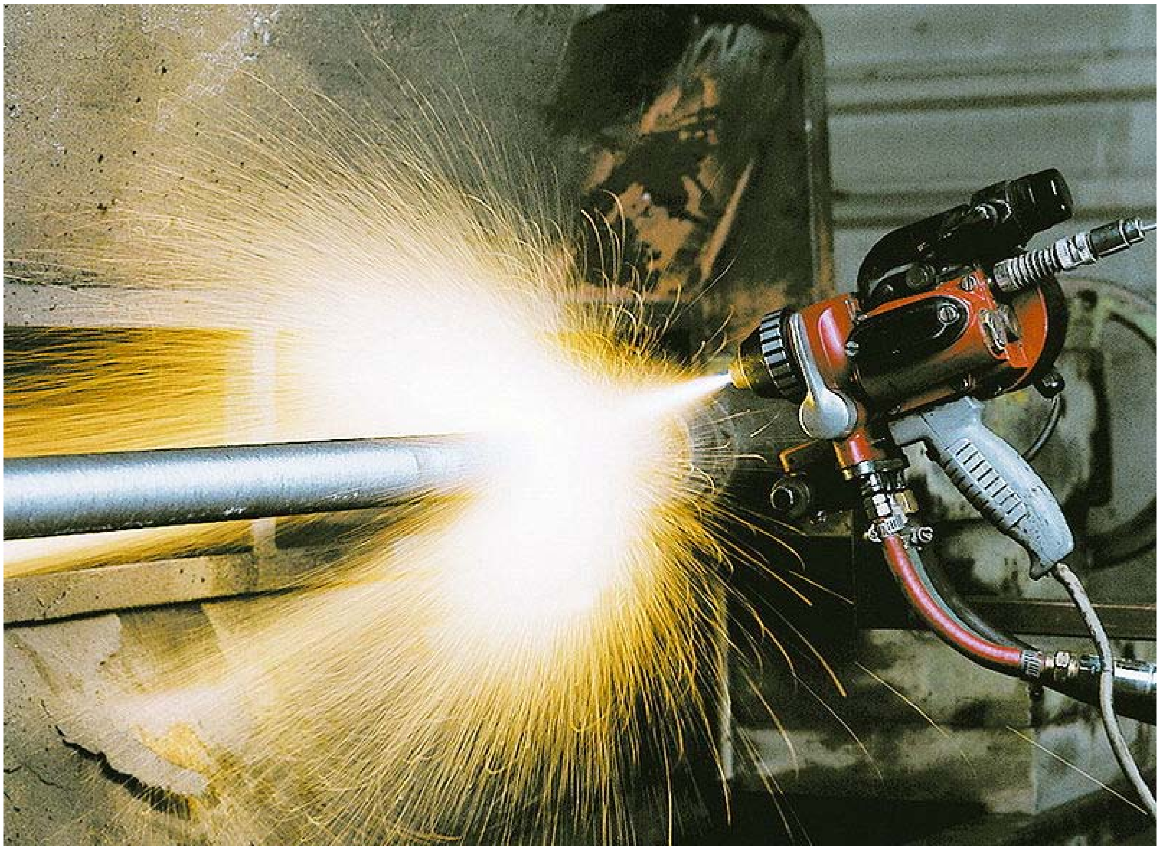
System overview

A typical thermal spray system consists of the following:

- Spray torch (or spray gun) - the core device performing the melting and acceleration of the particles to be deposited
- Feeder - for supplying the powder, wire or liquid to the torch
- Media supply - gases or liquids for the generation of the flame or plasma jet, gases for carrying the powder, etc.
- Robot - for manipulating the torch or the substrates to be coated
- Power supply - often standalone for the torch
- Control console(s) - either integrated or individual for all of the above

Plasma spraying

Plasma spraying setup



Wire flame spraying

In plasma spraying process, the material to be deposited (feedstock) — typically as a powder, sometimes as a liquid, suspension or wire — is introduced into the plasma jet, emanating from a plasma torch. In the jet, where the temperature is on the order of 10,000 K, the material is melted and propelled towards a substrate. There, the molten droplets flatten, rapidly solidify and form a deposit. Commonly, the deposits remain adherent to the substrate as coatings; free-standing parts can also be produced by removing the substrate. There are a large number of technological parameters that influence the interaction of the particles with the plasma jet and the substrate and therefore the deposit properties. These parameters include feedstock type, plasma gas composition and flow rate, energy input, torch offset distance, substrate cooling, etc.

Deposit properties

The deposits consist of a multitude of pancake-like lamellae called 'splats', formed by flattening of the liquid droplets. As the feedstock powders typically have sizes from micrometers to above 100 micrometers, the lamellae have thickness in the micrometer range and lateral dimension from several to hundreds of micrometers. Between these lamellae, there are small voids, such as pores, cracks and regions of incomplete bonding. As a result of this unique structure, the deposits can have properties significantly different from bulk materials. These are generally mechanical properties, such as lower strength and modulus, higher strain tolerance, and lower thermal and electrical conductivity. Also, due to the rapid solidification, metastable phases can be present in the deposits.

Applications

This technique is mostly used to produce coatings on structural materials. Such coatings provide protection against high temperatures (for example thermal barrier coatings for exhaust heat management), corrosion, erosion, wear; they can also change the appearance, electrical or tribological properties of the surface, replace worn material, etc. When sprayed on substrates of various shapes and removed, free-standing parts in the form of plates, tubes, shells, etc. can be produced. It can be also used for powder processing (spheroidization, homogenization, modification of chemistry, etc.). In that case, the substrate for deposition is absent and the particles solidify during flight or in a controlled environment (e.g., water). A polymer dispersion aerosol could be injected into the plasma discharge in order to create a grafting of this polymer at a substrate surface. This application is mainly used to modify the surface chemistry of polymers.

Variations

Plasma spraying systems can be categorized by several criteria.

Plasma jet generation:

- direct current (DC plasma), where the energy is transferred to the plasma jet by a direct current, high-power electric arc
- induction plasma or RF plasma, where the energy is transferred by induction from a coil around the plasma jet, through which an alternating, radio-frequency current passes

Plasma-forming medium:

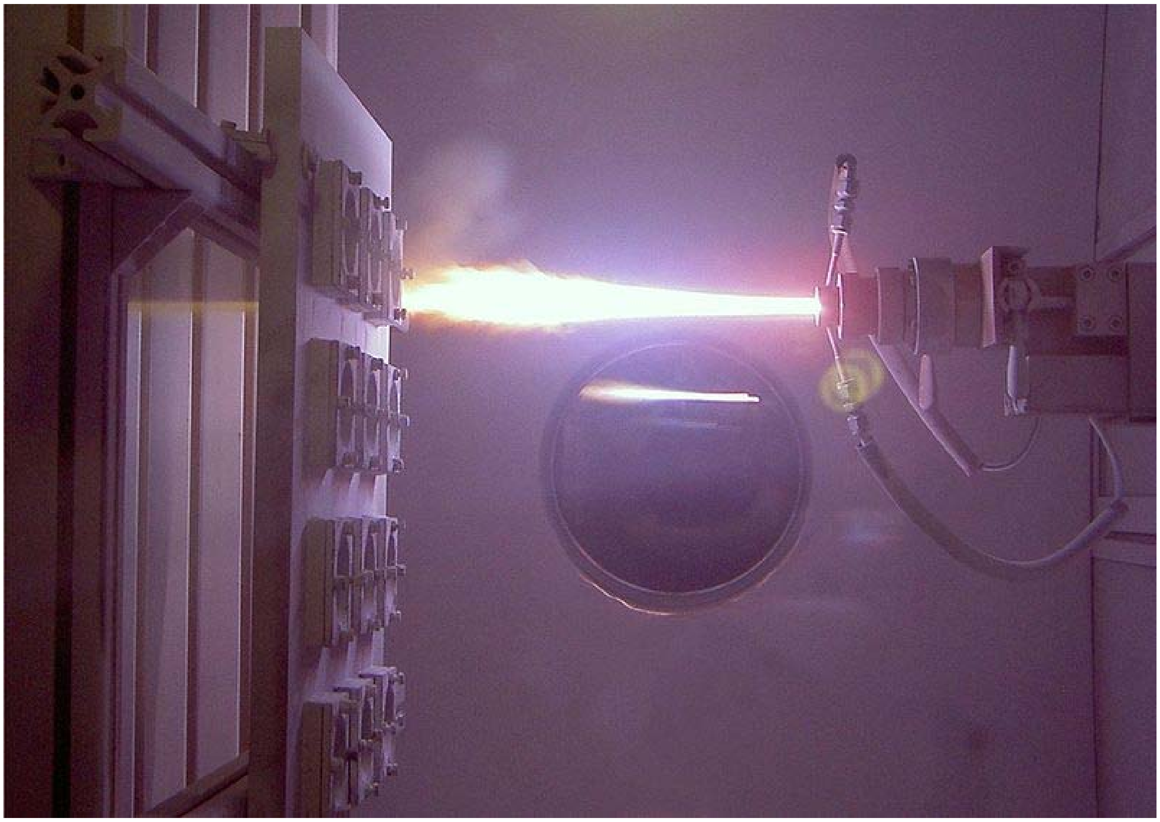
- gas-stabilized plasma (GSP), where the plasma forms from a gas; typically argon, hydrogen, helium or their mixtures
- water-stabilized plasma (WSP), where plasma forms from water (through evaporation, dissociation and ionization) or other suitable liquid
- hybrid plasma - with combined gas and liquid stabilization, typically argon and water

Spraying environment:

- air plasma spraying (APS), performed in the ambient air
- controlled atmosphere plasma spraying (CAPS), usually performed in a closed chamber, either filled with inert gas or evacuated
- variations of CAPS: high-pressure plasma spraying (HPPS), low-pressure plasma spraying (LPPS), extreme case of which is vacuum plasma spraying
- underwater plasma spraying

Another variation consists of having a liquid feedstock instead of a solid powder for melt, this techniques is known as Solution precursor plasma spray

Vacuum plasma spraying



Vacuum plasma spraying

Vacuum plasma spraying (VPS) is a technology for etching and surface modification to create porous layers with high reproducibility and for cleaning and surface engineering of plastics, rubbers and natural fibers as well as for replacing CFCs for cleaning metal components. This surface engineering can improve properties such as frictional behavior, heat resistance, surface electrical conductivity, lubricity, cohesive strength of films, or dielectric constant, or it can make materials hydrophilic or hydrophobic.

The process typically operates at 39–120 °C to avoid thermal damage. It can induce non-thermally activated surface reactions, causing surface changes which cannot occur with molecular chemistries at atmospheric pressure. Plasma processing is done in a controlled environment inside a sealed chamber at a medium vacuum, around 13–65 Pa. The gas or mixture of gases is energized by an electrical field from DC to microwave frequencies, typically 1–500 W at 50 V. The treated components are usually electrically isolated. The volatile plasma by-products are evacuated from the chamber by the vacuum pump, and if necessary can be neutralized in an exhaust scrubber.

In contrast to molecular chemistry, plasmas employ:

- Molecular, atomic, metastable and free radical species for chemical effects.
- Positive ions and electrons for kinetic effects.

Plasma also generates electromagnetic radiation in the form of vacuum UV photons to penetrate bulk polymers to a depth of about 10 µm. This can cause chain scissions and cross-linking.

Plasmas affect materials at an atomic level. Techniques like X-ray photoelectron spectroscopy and scanning electron microscopy are used for surface analysis to identify the processes required and to judge their effects. As a simple indication of surface energy, and hence adhesion or wettability, often a water droplet contact angle test is used. The lower the contact angle, the higher the surface energy and more hydrophilic the material is.

Changing effects with plasma

At higher energies ionization tends to occur more than chemical dissociations. In a typical reactive gas, 1 in 100 molecules form free radicals whereas only 1 in 10^6 ionizes. The predominant effect here is the forming of free radicals. Ionic effects can predominate with selection of process parameters and if necessary the use of noble gases.

Wire arc spray

Wire arc spray is a form of thermal spraying where two consumable metal wires are fed independently into the spray gun. These wires are then charged and an arc is generated between them. The heat from this arc melts the incoming wire, which is then entrained in air jet from the gun. This entrained molten feedstock is then deposited onto a substrate. This process is commonly used for metallic, heavy coatings.

Plasma transferred wire arc

Plasma transferred wire arc is another form of wire arc spray which deposits a coating on the internal surface of a cylinder, or on the external surface of a part of any geometry. It is predominantly known for its use in coating the cylinder bores of an engine, enabling the use of Aluminum engine blocks without the need for heavy cast iron sleeves. A single conductive wire is used as "feedstock" for the system. A supersonic plasma jet melts the wire, atomizes it and propels it onto the substrate. The plasma jet is formed by a transferred arc between a non-consumable cathode and the type of a wire. After atomization, forced air transports the stream of molten droplets onto the bore wall. The particles flatten when they impinge on the surface of the substrate, due to the high kinetic energy. The particles rapidly solidify upon contact. The stacked particles make up a high wear resistant coating. The PTWA thermal spray process utilizes a single wire as the feedstock material. All conductive wires up to and including 0.0625" (1.6mm) can be used as feedstock material, including "cored" wires. PTWA can be used to apply a coating to the wear surface of engine or transmission components to replace a bushing or bearing. For example, using PTWA to coat the bearing surface of a connecting rod offers a number of benefits including reductions in weight, cost, friction potential, and stress in the connecting rod.

High velocity oxygen fuel spraying (HVOF)

During the 1980s, a class of thermal spray processes called high velocity oxy-fuel spraying was developed: A mixture of gaseous or liquid fuel and oxygen is fed into a combustion chamber, where they are ignited and combusted continuously. The resultant hot gas at a pressure close to 1 MPa emanates through a converging–diverging nozzle and travels through a straight section. The fuels can be gases (hydrogen, methane, propane, propylene, acetylene, natural gas, etc.) or liquids (kerosene, etc.). The jet velocity at the exit of the barrel (>1000 m/s) exceeds the speed of sound. A powder feed stock is injected into the gas stream, which accelerates the powder up to 800 m/s. The stream of hot gas and powder is directed towards the surface to be coated. The powder partially melts in the stream, and deposits upon the substrate. The resulting coating has low porosity and high bond strength.

HVOF coatings may be as thick as 12 mm (1/2"). It is typically used to deposit wear and corrosion resistant coatings on materials, such as ceramic and metallic layers. Common powders include WC-Co, chromium carbide, MCrAlY, and alumina. The process has been most successful for depositing cermet materials (WC–Co, etc.) and other corrosion-resistant alloys (stainless steels, nickel-based alloys, aluminium, hydroxyapatite for medical implants, etc.).

Cold spraying

In the 1990s, cold spraying (often called gas dynamic cold spray) has been introduced. The method was originally developed in Russia with the accidental observation of the rapid formation of coatings, while experimenting with the particle erosion of the target exposed to a high velocity flow loaded with fine powder in a wind tunnel. In cold spraying, particles are accelerated to very high speeds by the carrier gas forced through a converging–diverging de Laval type nozzle. Upon impact, solid particles with sufficient kinetic energy deform plastically and bond metallurgically to the substrate to form a coating. The critical velocity needed to form bonding depends on the materials properties, powder size and temperature. Soft metals such as Cu and Al are best suited for cold spraying, but coating of other materials (W, Ta, Ti, MCrAlY, WC–Co, etc.) by cold spraying has been reported.

The deposition efficiency is typically low for alloy powders, and the window of process parameters and suitable powder sizes is narrow. To accelerate powders to higher velocity, finer powders (<20 micrometers) are used. It is possible to accelerate powder particles to much higher velocity using a processing gas having high speed of sound (helium instead of nitrogen). However, helium is costly and its flow rate, and thus consumption, is higher. To improve acceleration capability, nitrogen gas is heated up to about 900 C. As a result, deposition efficiency and tensile strength of deposits increase.

Warm spraying

Is a novel modification of high-velocity oxy-fuel spraying, in which the temperature of combustion gas is lowered by mixing nitrogen with the combustion gas, thus bringing the process closer to the cold spraying. The resulting gas contains much water vapor, unreacted hydrocarbons and oxygen, and thus is dirtier than the cold spraying. However, the coating efficiency is higher. On the other hand, lower temperatures of warm spraying reduce melting and chemical reactions of the feed powder, as compared to HVOF. These advantages are especially important for such coating materials as Ti, plastics, and metallic glasses, which rapidly oxidize or deteriorate at high temperatures.

Applications

Plasma sprayed ceramic coating applied onto a part of an automotive exhaust system

- Crankshaft reconditioning or conditioning
- Corrosion protection
- Fouling protection
- Altering thermal conductivity or electrical conductivity
- Wear control: either hardfacing (wear-resistant) or abradable coating
- Repairing damaged surfaces
- Temperature/oxidation protection (thermal barrier coatings)
- Medical implants
- Production of functionally graded materials (for either of the above applications)

Safety

Thermal spraying need not be a dangerous process, if the equipment is treated with care, and correct spraying practices are followed. As with any industrial process, there are a number of hazards, of which the operator should be aware, and against which specific precautions should be taken. Ideally, equipment should be operated automatically, in enclosures specially designed to extract fumes, reduce noise levels, and prevent direct viewing of the spraying head. Such techniques will also produce coatings that are more consistent. There are occasions when the type of components being treated, or their low production levels, requires manual equipment operation. Under these conditions, a number of hazards, peculiar to thermal spraying, are experienced, in addition to those commonly encountered in production or processing industries.

Noise

Metal spraying equipment uses compressed gases, which create noise. Sound levels vary with the type of spraying equipment, the material being sprayed, and the operating parameters. Typical sound pressure levels taken 1 meter behind the arc.

UV light

Combustion spraying equipment produces an intense flame, which may have a peak temperature more than 3,100°C, and is very bright. Electric arc spraying produces ultra-violet light, which may damage delicate body tissues. Spray booths, and enclosures, should be fitted with ultra-violet absorbent dark glass. Where this is not possible, operators, and others in the vicinity should wear protective goggles containing BS grade 6 green glass. Opaque screens should be placed around spraying areas. The nozzle of an arc pistol should never be viewed directly, unless it is certain that no power is available to the equipment.

Dust and fumes

The atomization of molten materials produces a large amount of dust and fumes made up of very fine particles (about 80 – 95% of the particles by number <100 nm). Proper extraction facilities are vital, not only for personal safety, but to minimize entrapment of re-frozen particles in the sprayed coatings. The use of respirators, fitted with suitable filters, is strongly recommended, where equipment cannot be isolated. Certain materials offer specific known hazards:

1. Finely divided metal particles are potentially pyrophoric and none should be allowed to accumulate.
2. Certain materials e.g. aluminum, zinc and other base metals may react with water to evolve hydrogen. This is potentially explosive and special precautions are necessary in fume extraction equipment.
3. Fumes of certain materials, notably zinc and copper alloys are unpleasant to smell, and, in certain individuals, may cause a fever-type reaction. This may occur some time after spraying and usually subsides rapidly. If it does not, medical advice must be sought.

Heat

Combustion spraying guns use oxygen and fuel gases. The fuel gases are potentially explosive. In particular, acetylene may only be used under approved conditions. Oxygen, while not explosive, will sustain combustion, and many materials will spontaneously ignite, if excessive oxygen levels are present. Care must be taken to avoid leakage, and to isolate oxygen and fuel gas supplies, when not in use.

Shock hazards

Electric arc guns operate at low voltages (below 45 V dc), but at relatively high currents. They may be safely hand-held. The power supply units are connected to 440 V AC sources, and must be treated with caution.

Chapter 6

Vacuum Brake

The **vacuum brake** is a braking system employed on trains and introduced in the mid-1860s. A variant, the **automatic vacuum brake** system, became almost universal in British train equipment and in those countries influenced by British practice. Vacuum brakes also enjoyed a brief period of adoption in the USA, primarily on narrow gauge railroads. However, its limitations caused it to be progressively superseded by compressed air systems starting in the United Kingdom from the 1970s onward. The vacuum brake system is now obsolete; it is not in large-scale usage anywhere in the world,other than in South Africa supplanted in the main by air brakes.

Overview

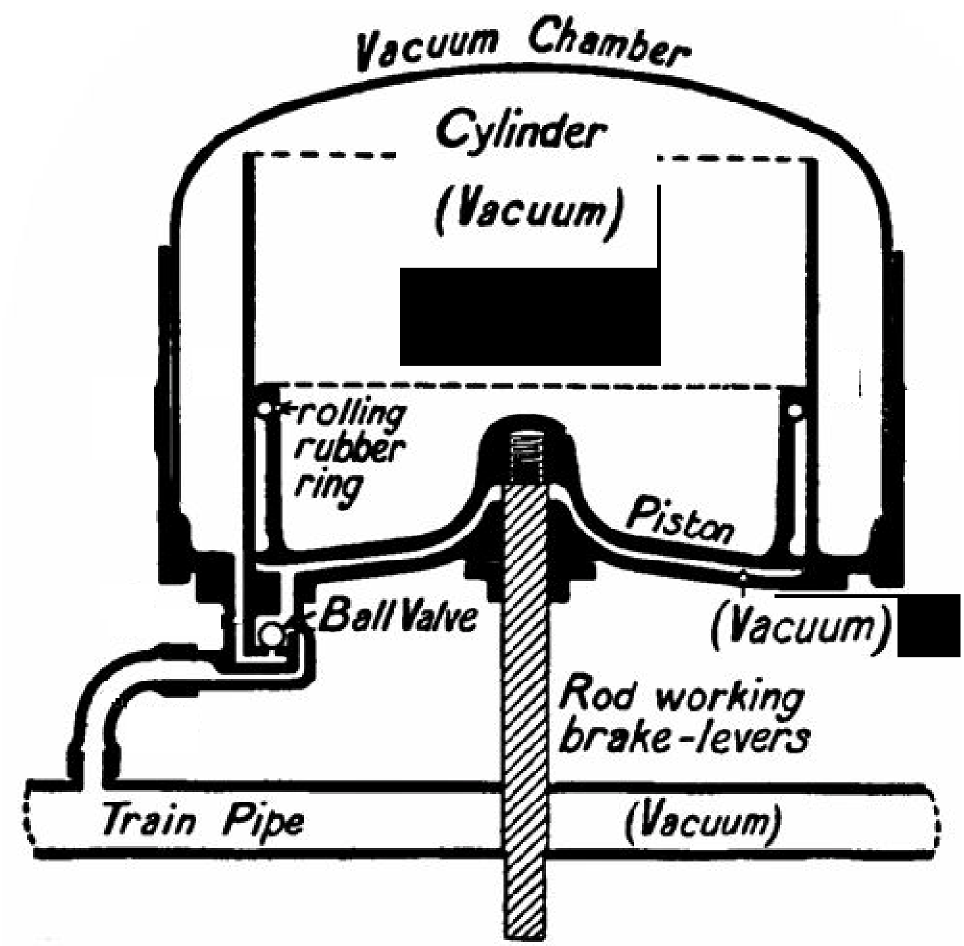
In the earliest days of railways, trains were slowed or stopped by the application of manually applied brakes on the locomotive and in brake vehicles through the train, and later by steam power brakes on locomotives. This was clearly unsatisfactory, but the technology of the time did not easily offer an improvement. A chain braking system was developed, requiring a chain to be coupled throughout the train, but it was impossible to arrange equal braking effort down the length of the train.

A major advance was the adoption of a vacuum braking system in which flexible pipes were connected between all the vehicles of the train, and brakes on each vehicle could be controlled from the locomotive. The earliest pattern was a simple vacuum brake, in which vacuum was created by operation of a valve on the locomotive; the vacuum actuated brake pistons on each vehicle, and the degree of braking could be increased or decreased by the driver. Vacuum, rather than compressed air, was preferred because steam locomotives can be fitted with ejectors, which are simple venturi devices that create vacuum without the use of moving parts.

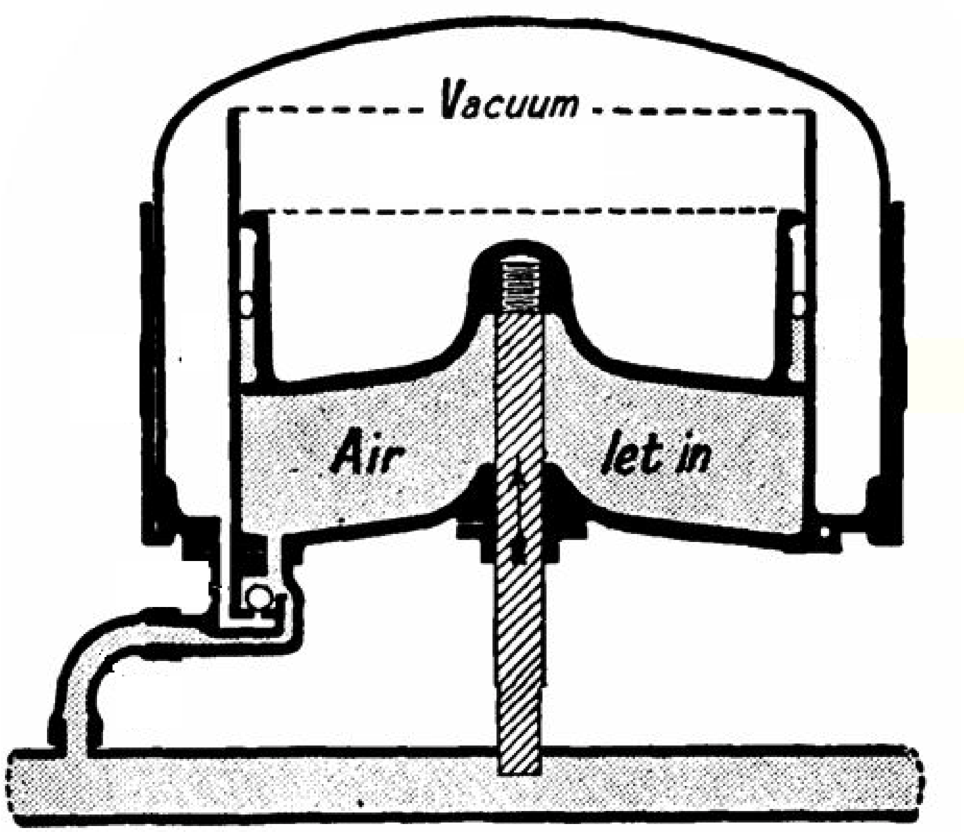
However, the simple vacuum system had the major defect that in the event of one of the hoses connecting the vehicles becoming displaced (by the train accidentally dividing, or by careless coupling of the hoses, or otherwise) the vacuum brake on the entire train was useless.

The automatic vacuum brake had been developed: it was designed to apply fully if the train becomes divided or if a hose becomes displaced, but opposition on the grounds of cost (particularly by the LNWR and its chairman Richard Moon) to the fitting of the automatic type of brake meant that it took a serious accident at Armagh in 1889 before legislation compelled the automatic system. In this accident at Armagh, a portion of a train was detached from the locomotive on a steep gradient and ran away, killing 88 people. The train was fitted with the simple vacuum brake, which was useless on the disconnected portion of the train. It was clear that if the vehicles had been fitted with an automatic continuous brake, the accident would almost certainly not have happened, and the public concern at the scale of the accident prompted legislation mandating the use of a continuous automatic brake on all passenger trains.

How the automatic vacuum brake works



Vacuum brake cylinder in running position: the vacuum is the same above and below the piston



Air at atmospheric pressure from the train pipe is admitted below the piston, which is forced up

In its simplest form, the automatic vacuum brake consists of a continuous pipe -- the train pipe -- running throughout the length of the train. In normal running a partial vacuum is maintained in the train pipe, and the brakes are released. When air is admitted to the train pipe, the air pressure acts against pistons in cylinders in each vehicle. A vacuum is sustained on the other face of the pistons, so that a net force is applied. A mechanical linkage transmits this force to brake shoes which act by friction on the treads of the wheels.

The fittings to achieve this are therefore:

- a train pipe: a steel pipe running the length of each vehicle, with flexible vacuum hoses at each end of the vehicles, and coupled between adjacent vehicles; at the end of the train, the final hose is seated on an air-tight plug;
- an ejector on the locomotive, to create vacuum in the train pipe;
- controls for the driver to bring the ejector into action, and to admit air to the train pipe; these may be separate controls or a combined brake valve;
- a brake cylinder on each vehicle containing a piston, connected by rigging to the brake shoes on the vehicle; and
- a vacuum (pressure) gauge on the locomotive to indicate to the driver the degree of vacuum in the train pipe.

The brake cylinder is contained in a larger housing—this gives a reserve of vacuum as the piston operates. The cylinder rocks slightly in operation to maintain alignment with the brake rigging cranks, so it is supported in trunnion bearings, and the vacuum pipe connection to it is flexible. The piston in the brake cylinder has a flexible piston ring that allows air to pass from the upper part of the cylinder to the lower part if necessary.

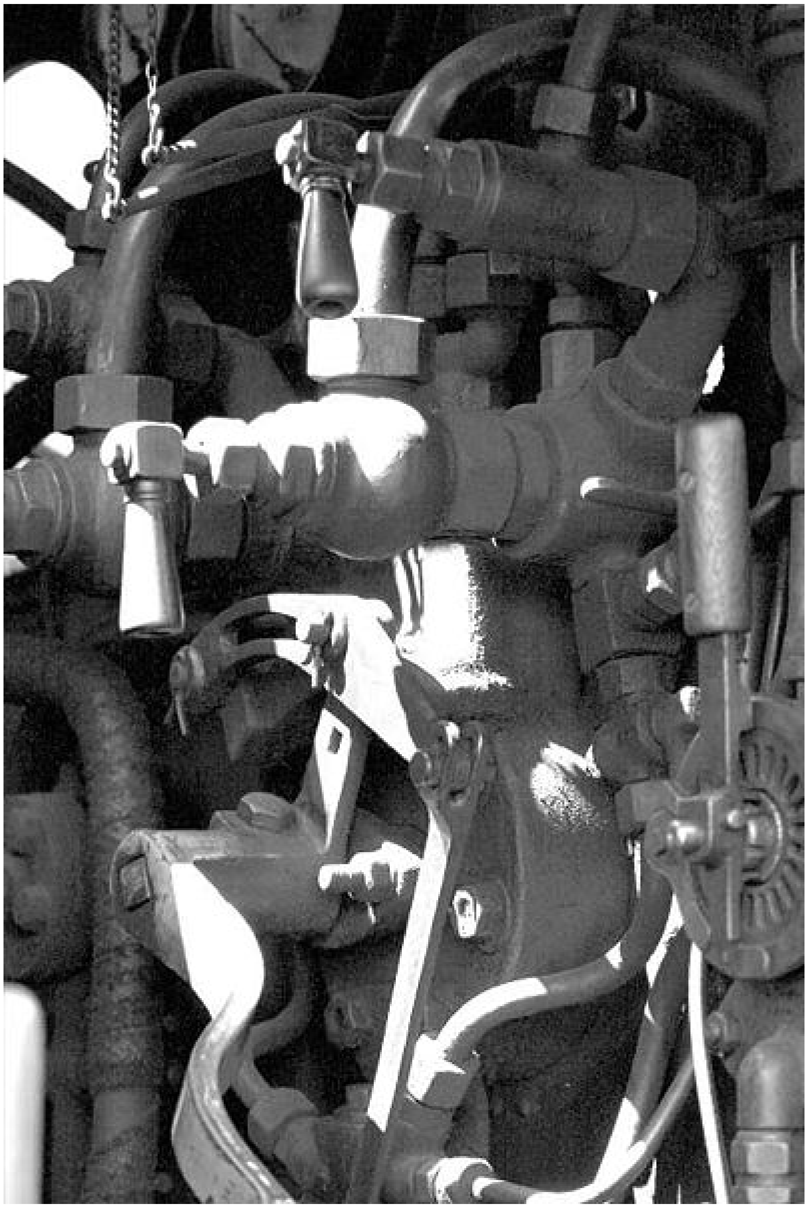
When the vehicles have been at rest, so that the brake is not charged, the brake pistons will have dropped to their lower position in the absence of a pressure differential (as air will have leaked slowly into the upper part of the cylinder, destroying the vacuum).

When a locomotive is coupled to the vehicles, the driver moves his brake control to the "release" position and air is exhausted from the train pipe, creating a partial vacuum. Air in the upper part of the brake cylinders is also exhausted from the train pipe, through a non-return valve.

If the driver now moves his control to the "brake" position, air is admitted to the train pipe. According to the driver's manipulation of the control, some or all of the vacuum will be destroyed in the process. The ball valve closes and there is a higher air pressure under the brake pistons than above it, and the pressure differential forces the piston upwards, applying the brakes. The driver can control the severity of the braking effort by admitting more or less air to the train pipe.

Practical considerations

The automatic vacuum brake as described represented a very considerable technical advance in train braking. In practice steam locomotives had two ejectors, a small ejector for running purposes (to exhaust air that had leaked into the train pipe) and a large ejector to release brake applications. Later Great Western Railway practice was to use a vacuum pump instead of the small ejector.



Graduable brake valve (right) and the small (upper) and large ejector cocks from a GWR locomotive

The driver's brake valve was usually combined with the steam brake control on the locomotive.

Release valves are provided on the brake cylinders; when operated, usually by manually pulling a cord near the cylinder, air is admitted to the upper part of the brake cylinder on that vehicle. This is necessary to release the brake on a vehicle that has been uncoupled from a train and now requires to be moved without having a brake connection to another locomotive, for example if it is to be shunted.

In the United Kingdom the pre-nationalisation railway companies standardised around systems operating on a vacuum of 21 inches of mercury (533.4 Torr), with the exception of the Great Western Railway, which used 25 inches of mercury (635 Torr). An absolute vacuum is about 30 inches of mercury (760 Torr), depending on atmospheric conditions.

This difference in standards could cause problems on long distance cross-country services when a GWR locomotive was replaced with another company's engine, as the new engine's large ejector would sometimes not be able to fully release the brakes on the train. In this case the release valves on each vehicle in the train would have to be released by hand. This time consuming process was not infrequently seen at large GWR stations such as Bristol Temple Meads.

The provision of a train pipe running throughout the train enabled the automatic vacuum brake to be operated in emergency from any position in the train. Every guard's compartment had a brake valve, and the passenger communication apparatus (usually called "the communication cord" in lay terminology) also admitted air into the train pipe at the end of coaches so equipped.

When a locomotive is first coupled to a train, or if a vehicle is detached or added, a brake continuity test is carried out, to ensure that the brake pipes are connected throughout the entire length of the train.

Limitations

The progress represented by the automatic vacuum brake nonetheless carried some limitations; chief among these were:

- the practical limit on the degree of vacuum attainable means that a very large brake piston and cylinder are required to generate the force necessary on the brake blocks; when a proportion of the British ordinary wagon fleet was fitted with vacuum brakes in the 1950's, the physical dimensions of the brake cylinder prevented the wagons from operating in some private sidings that had tight clearances;
- for the same reason, on a very long train, a considerable volume of air has to be admitted to the train pipe to make a full brake application, and a considerable volume has to be exhausted to release the brake (if for example a signal at danger is suddenly lowered and the driver requires to resume speed); while the air is traveling along the train pipe, the brake pistons at the head of the train have responded to the brake application or release, but those at the tail will respond much later, leading to undesirable longitudinal forces in the train. In extreme cases this has led to breaking couplings and causing the train to divide.
- the existence of vacuum in the train pipe can cause debris to be sucked in. An accident took place near Ilford in the 1950's, due to inadequate braking effort in the train. A rolled newspaper was discovered in the train pipe, effectively isolating the rear part of the train from the driver's control. The blockage should

have been detected if a proper brake continuity test had been carried out before the train started its journey.

A development introduced in the 1950's was the **direct admission valve**, fitted to every brake cylinder. These valves responded to a rise in train pipe pressure as the brake was applied, and admitted atmospheric air directly to the underside of the brake cylinder.

American and continental European practice had long favoured compressed air brake systems, the leading pattern being a proprietary Westinghouse system. This has a number of advantages, including smaller brake cylinders (because a higher air pressure could be used) and a somewhat more responsive braking effort. However the system requires an air pump. On steam engines this was usually a reciprocating steam pump, and it was quite bulky. Its distinctive shape and the characteristic puffing sound when the brake is released (as the train pipe has to be recharged with air) make steam locomotives fitted with the Westinghouse brake unmistakable, for example in old films.

In the UK, the Great Eastern Railway, the North Eastern Railway, the London Brighton and South Coast Railway and the Caledonian Railway adopted the Westinghouse system. It was also standard on the Isle of Wight rail system. Inevitably this led to compatibility problems in exchanging traffic with other lines. It was possible to provide through pipes for the braking system not fitted to any particular vehicle so that it could run in a train using the "other" system, allowing through control of the fitted vehicles behind it, but of course with no braking effort of its own.

Dual brakes

Vehicles can be fitted with dual brakes, vacuum and air, provided that there is room to fit the duplicated equipment. It is much easier to fit one kind of brake with a pipe for continuity of the other. Train crew need to take note that the wrong-fitted wagons do not contribute to the braking effort and make allowances on down grades to suit. Many of the earlier classes of diesel locomotive used on British Railways were fitted with dual systems to enable full usage of BR's rolling stock inherited from the private companies which had different systems depending on which company the stock originated from.

Air brakes need a tap to seal the hose at the ends of the train. If these taps are incorrectly closed, a loss of brake force may occur, leading to a dangerous runaway. With vacuum brakes, the end of the hose can be plugged into a stopper which seals the hose by suction. It is much harder to block the hose pipe compared to air brakes.

Twin pipe systems

Vacuum brakes can be operated in a twin pipe mode to speed up applications and release.

Present-day use of vacuum brakes

Today's largest operators of trains equipped with vacuum brakes are the Railways of India and Spoornet (South Africa), however there are also trains with air brakes and dual brakes in use. South African Railways (Spoornet) operates more than 1 000 electric multiple unit cars, which are fitted with electro-vacuum brakes. The electro-vacuum system uses a 2 inch train pipe and basic automatic vacuum brake system, with the addition of electrically-controlled application and release valves in each vehicle. The application and release valves greatly increase the rate of train pipe vacuum destruction and creation. This, in turn, greatly increases the speed of brake application and release. The performance of electro-vacuum brakes on SAR EMUs is equivalent to electro-pneumatic braked EMUs of a similar age.

Other African railways are believed to continue to use the vacuum brake. Other operators of vacuum brakes are narrow gauge railways in Central Europe, the largest of which is the Rhaetian Railway.

Vacuum brakes have been entirely superseded on the National Rail system in the UK, although they are still in use on most heritage railways. They are also to be found on a number (though increasingly fewer) main line vintage special trains.

Iarnród Éireann (the national rail operator in the Republic of Ireland) ran vacuum-braked British Railways Mark 2 stock on passenger trains until the end of March 2008 and still operates vacuum-braked revenue freight (at least in the case of Tara Mines ore traffic).

Chapter 7

Diffusion Pump and Sorption Pump

Diffusion pump

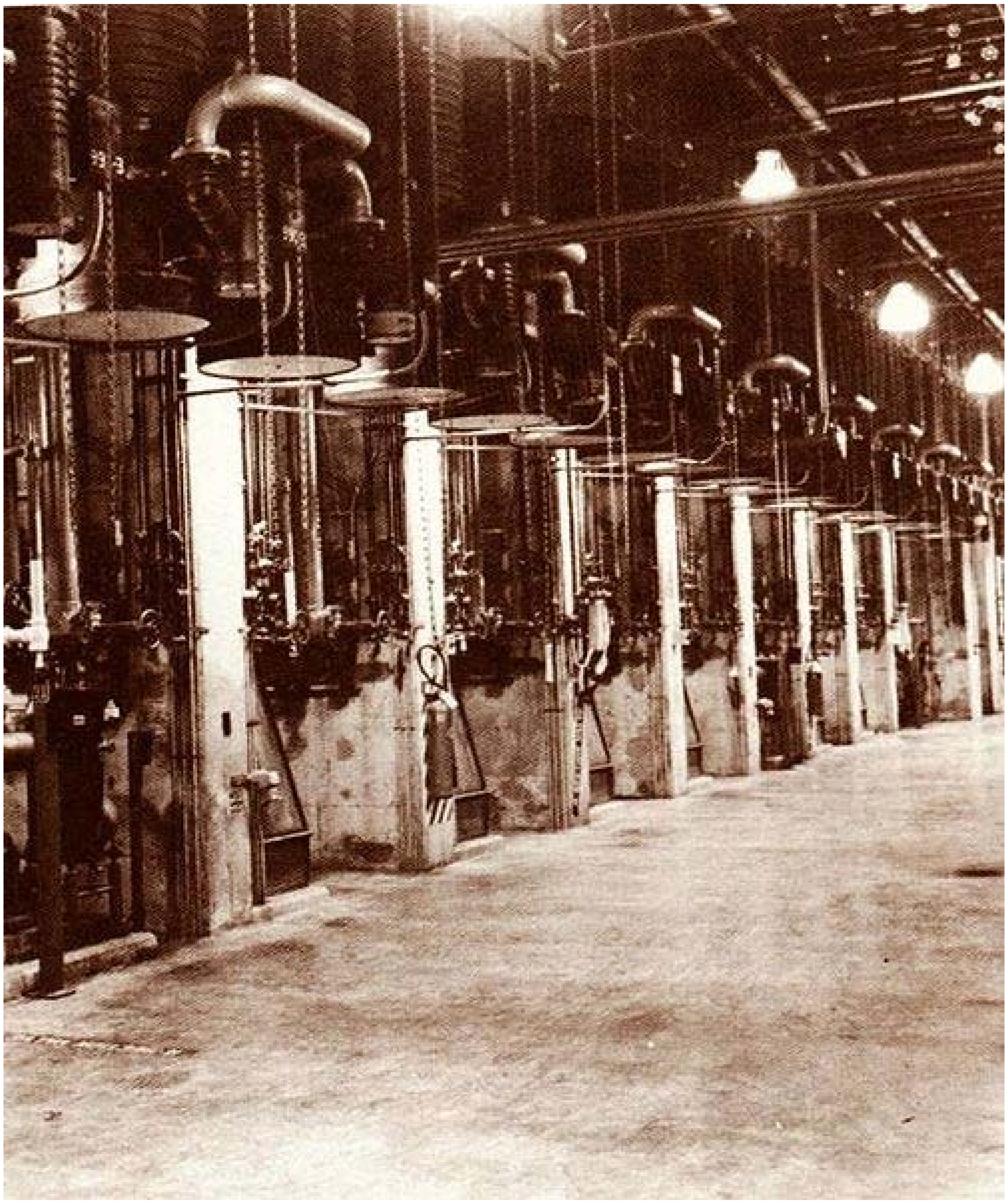
Six inch oil diffusion pump.

Diffusion pumps use a high speed jet of vapor to direct gas molecules in the pump throat down into the bottom of the pump and out the exhaust. Presented in 1915 by Wolfgang

Gaede and using mercury vapor, they were the first type of high vacuum pumps operating in the regime of free molecular flow, where the movement of the gas molecules can be better understood as diffusion than by conventional fluid dynamics. Gaede used the name **diffusion pump** since his design was based on the finding that gas cannot diffuse against the vapor stream, but will be carried with it to the exhaust. However, the principle of operation might be more precisely described as **gas-jet pump**, since diffusion plays a role also in other high vacuum pumps. In modern text books, the diffusion pump is categorized as a momentum transfer pump. The diffusion pump is widely used in both industrial and research applications. Most modern diffusion pumps use silicone oil as the working fluid. Cecil Reginald Burch discovered the possibility of using silicone oil in 1928.

Oil diffusion pumps

The oil diffusion pump is operated with an oil of low vapor pressure. Its purpose is to achieve higher vacuum (lower pressure) than is possible by use of positive displacement pumps alone. Although its use has been mainly associated within the high-vacuum range (down to 10^{-9} mbar), diffusion pumps today can produce pressures approaching 10^{-10} mbar when properly used with modern fluids and accessories. The features that make the diffusion pump attractive for high and ultra-high vacuum use are its high pumping speed for all gases and low cost per unit pumping speed when compared with other types of pump used in the same vacuum range. Diffusion pumps cannot discharge directly into the atmosphere, so a mechanical forepump is typically used to maintain an outlet pressure around 0.1 mbar.



Diffusion pumps used on the Calutron mass spectrometers during the Manhattan Project.

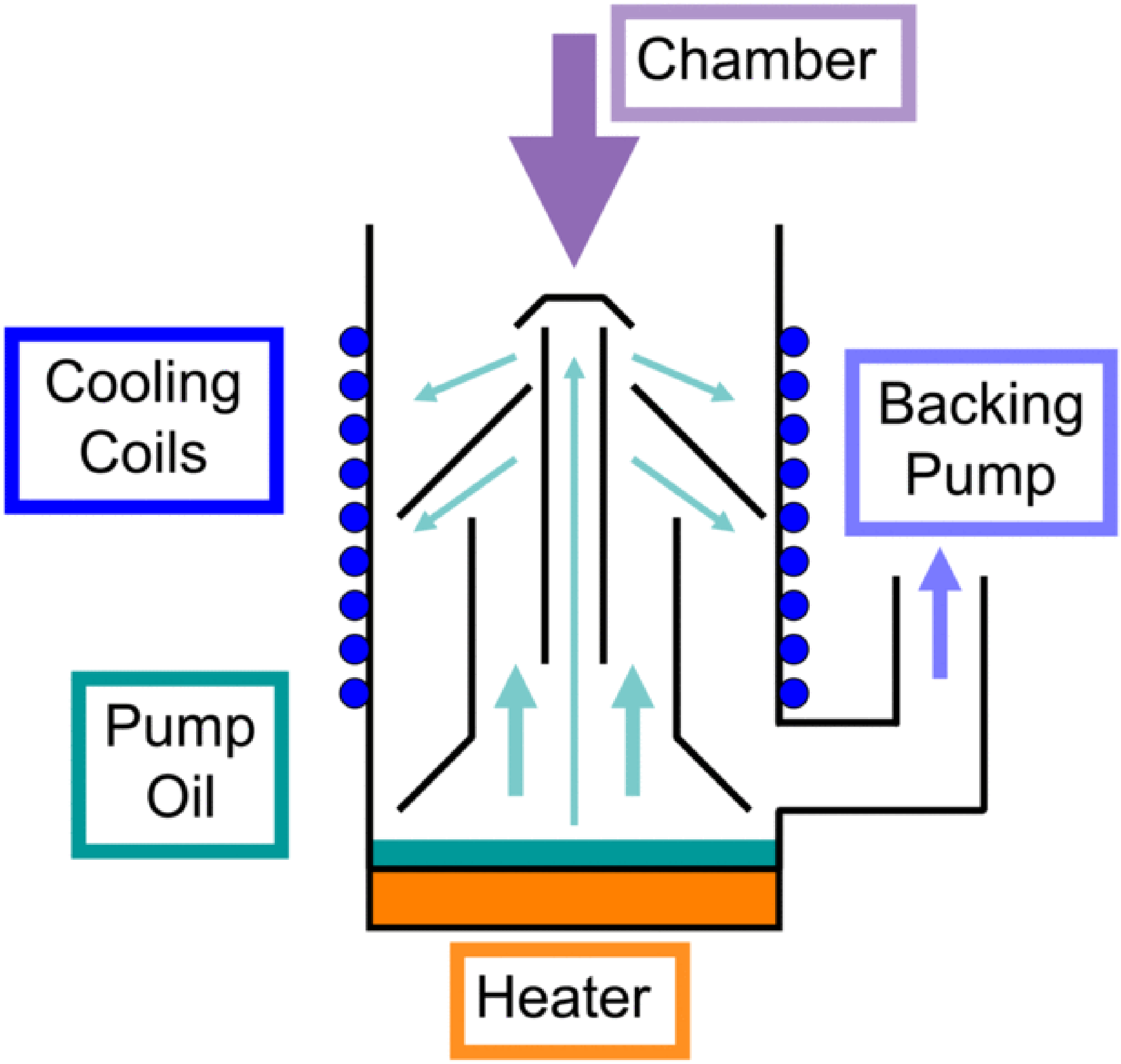


Diagram of an oil diffusion pump

The high speed jet is generated by boiling the fluid and directing the vapor through a jet assembly. Note that the oil is gaseous when entering the nozzles. Within the nozzles, the flow changes from laminar, to supersonic and molecular. Often several jets are used in series to enhance the pumping action. The outside of the diffusion pump is cooled using either air flow or a water line. As the vapor jet impacts the outer cooled shell of the diffusion pump, the working fluid condenses and is recovered and directed back to the boiler. The pumped gases continue flowing to the base of the pump at increased pressure, flowing out through the diffusion pump outlet, where they are compressed to ambient pressure by the secondary mechanical forepump and exhausted.

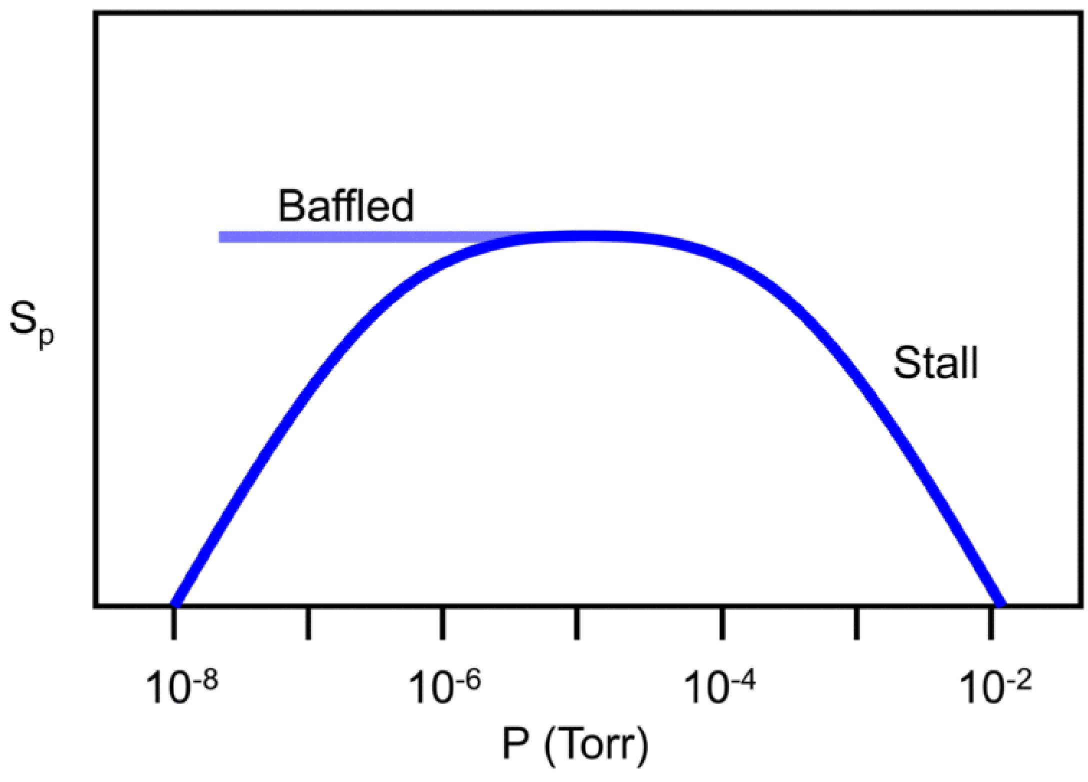
Unlike turbomolecular pumps and cryopumps, diffusion pumps have no moving parts and as a result are quite durable and reliable. They can function over pressures ranges of 10^{-10}

to 10^{-2} mbar. They are driven only by convection and thus have a very low energy efficiency.

One major disadvantage of diffusion pumps is the tendency to backstream oil into the vacuum chamber. This oil can contaminate surfaces inside the chamber or upon contact with hot filaments or electrical discharges may result in carbonaceous or siliceous deposits. Due to backstreaming, oil diffusion pumps are not suitable for use with highly sensitive analytical equipment or other applications which require an extremely clean vacuum environment, but mercury diffusion pumps may be in the case of ultra high vacuum chambers used for mercury deposition. Often cold traps and baffles are used to minimize backstreaming, although this results in some loss of pumping ability.

The oil of a diffusion pump cannot be exposed to the atmosphere when hot. If this occurs, the oil will burn and has to be replaced.

Steam ejectors



Plot of pumping speed as a function of pressure for a diffusion pump.

The steam ejector is a popular form of diffusion pump for vacuum distillation and freeze-drying. A jet of steam entrains the vapour that must be removed from the vacuum chamber. Steam ejectors can have a single or multiple stages, with and without condensers in between the stages.

Compressed-air ejectors

One class of diffusion vacuum pumps is the multistage compressed-air driven ejector. It is very popular in applications where objects are moved around using suction cups and vacuum lines.

Sorption pump

The **sorption pump** is a vacuum pump that creates a vacuum by adsorbing molecules on a very porous material like molecular sieve which is cooled by a cryogen, typically liquid nitrogen. The ultimate pressure is about 10^{-2} mbar. With special techniques this can be lowered till 10^{-7} mbar. The main advantages are the absence of oil or other contaminants, low cost and vibration free operation because there are no moving parts. The main disadvantages are that it cannot operate continuously and cannot effectively pump hydrogen, helium and neon, all gases with lower condensation temperature than liquid nitrogen. The main application is as a roughing pump for a sputter-ion pump in ultra-high vacuum experiments, for example in surface physics.

Construction

A sorption pump is usually constructed in stainless steel, aluminium or borosilicate glass. It can be a simple Pyrex flask filled with molecular sieve or an elaborate metal construction consisting of a metal flask containing perforated tubing and heat-conducting fins. A pressure relief valve can be installed. The design only influences the pumping speed and not the ultimate pressure that can be reached. The design details are a trade-off between fast cooling using heat conducting fins and high gas conductance using perforated tubing.

The typical molecular sieve used is a synthetic zeolite with a pore diameter around 0.4 nanometer (Type 4A) and a surface area of about 500 m^2/g. The sorption pump contains between 300 g and 1.2 kg of molecular sieve. A 15 liter system will be pumped down to about 10^{-2} mbar by 300 g molecular sieve.

Operation

The sorption pump is a cyclic pump and its cycle has 3 phases: sorption, desorption and regeneration.

In the sorption phase the pump is actually used to create a vacuum. This is achieved by cooling the pump body to low temperatures, typically by immersing it in a Dewar flask filled with liquid nitrogen. Gases will now either condense or be adsorbed by the large surface of the molecular sieve.

In the desorption phase the pump is allowed warm up to room temperature and the gases escape through the pressure relief valve or other opening to the atmosphere. If the pump has been used to pump toxic, flammable or other dangerous gasses one has to be careful to vent safely into the atmosphere as all gases pumped during the sorption phase will be released during the desorption phase.

In the regeneration phase the pump body is heated to 300 °C to drive off water vapor that does not desorb at room temperature and accumulates in the molecular sieve. It takes typically 2 hours to fully regenerate a pump.

The pump can be used in a cycle of sorption and desorption until it loses too much efficiency and is regenerated or in a cycle where sorption and desorption are always followed by regeneration.

After filling a sorption pump with new molecular sieve it should always be regenerated as the new molecular sieve is probably saturated with water vapor. Also when a pump is not in use it should be closed off from the atmosphere to prevent water vapor saturation.

Performance improvement

Pumping capacity can be improved by prepumping the system by another simple and clean vacuum pump like a diaphragm pump or even a water aspirator or compressed-air venturi pump.

Sequential or multistage pumping can be used to attain lower pressures. In this case two or more pumps are connected in parallel to the vacuum vessel. Every pump has a valve to isolate it from the vacuum vessel. At the start of the pump down all valves are open. The first pump is cooled down while the others are still hot. When the first pump has reached its ultimate pressure it is closed off and the next pump is cooled down. Final pressures are in the 10^{-4} mbar region. What is left is mainly helium because it is almost not pumped at all. The final pressure almost equals the partial pressure of helium in air.

A sorption pump does pump all gases effectively with the exception of hydrogen, helium and neon which do not condensate at liquid nitrogen temperatures and are not efficiently adsorbed by the molecular sieves because of their small molecular size. This problem can be solved by purging the vacuum system with dry pure nitrogen before pump down. In purged system with aspirator rough pumping ultimate pressures of 10^{-4} mbar for a single sorption pump and 10^{-7} mbar for sequential pumping can be reached. A typical source of dry pure nitrogen would be a liquid nitrogen Dewar head space.

It has been suggested that by applying a dynamic pumping technique hydrogen, helium and neon can also be pumped without resorting to dry nitrogen purging. This is done by precooling the pump with the valve to the vacuum vessel closed. The valve is opened when the pump is cold and the inrush of adsorbable gases will carry all other gases into the pump. The valve is closed before hydrogen, helium or neon can back-migrate into the vacuum vessel. Sequential pumping can also be applied. No final pressures are given.

Continuous pumping may be simulated by using two pumps in parallel and letting one pump pump the system while the other pump, temporally sealed-off from the system, is in the desorption phase and venting to the atmosphere. When the pump is well desorbed it is cooled down and reconnected to the system. The other pump is sealed-off and goes into desorption. This becomes a continuous cycle.

Chapter 8

Titanium Sublimation Pump and Turbomolecular Pump

Titanium sublimation pump

A **Titanium sublimation pump** (TSP) may be used as a component of ultra high vacuum systems.

Principle of Operation

Its construction and principle of operation is extremely simple. It consists of a titanium filament through which a high current (typically around 40 Amps) is passed periodically. This current causes the filament to reach the sublimation temperature of titanium, and hence the surrounding chamber walls become coated with a thin film of clean titanium. Since clean titanium is very reactive, components of the residual gas in the chamber which collide with the chamber wall are likely to react and to form a stable, solid product. Thus the gas pressure in the chamber is reduced.

After some time, the titanium film will no longer be clean and hence the effectiveness of the pump is reduced. Therefore, after a certain time, the titanium filament should be heated again, and a new film of titanium re-deposited on the chamber wall. Since the time taken for the titanium film to react depends on a number of factors (such as the composition of the residual gas, the temperature of the chamber and the total pressure), the period between successive sublimations requires some consideration. Typically, the operator does not know all of these factors, so the sublimation period is estimated according to the total pressure and by observing the effectiveness of the outcome. Some TSP controllers use a signal from the pressure gauge to estimate the appropriate period.

Since the TSP filament has a finite lifetime, TSPs commonly have multiple filaments to allow the operator to switch to a new one without needing to open the chamber. Replacing used filaments can then be combined with other maintenance jobs.

The effectiveness of the TSP depends on a number of factors. Amongst the most critical are; the area of the titanium film, the temperature of the chamber walls and the composition of the residual gas. The area is typically maximised when considering where to mount to TSP. The reactivity of the new titanium film is increased at lower temperatures, so it is desirable to cool the relevant part of the chamber, typically using liquid nitrogen. However, due to the cost of the nitrogen and the need to ensure a continuous supply, TSPs are commonly operated at room temperature. Finally the residual gas composition is important - typically the pump works well with the more reactive components (such as CO and O_2), but is very ineffective at pumping inert components such as the noble gasses. Therefore TSP must be used in conjunction with other pumps.

Other pumps which use exactly the same working principle, but using something other than titanium as a source are also relatively common. This family of pumps are usually called 'getter pumps' or 'getters' and typically consist of metals which are reactive with the components of the residual gas which are not pumped by the TSP. By choosing a number of such sources, most constituents of the residual gas, except for the noble gasses, can be targeted.

Practical considerations

When mounting the TSP in the chamber, a number of important considerations must be made. First, it is desirable that the filament can deposit on a large area. However, one must take care that the titanium is not deposited onto anything it can damage. For example electrical feedthroughs containing ceramic insulators will fail if the titanium forms a conducting film which bridges the ceramic insulator. Samples may become contaminated by titanium if they have line-of-sight to the pump. Also, titanium is a very hard material, so titanium film which builds up on the inside of the chamber may form flakes which fall into mechanical components (typically turbomolecular pumps and valves) and damage them.

Many chambers containing TSPs also have an ion pump. Often the ion pump provides a good location for the TSP, and some manufacturers promote the use of combined TSP/ion-pumps. Furthermore, TSPs have been shown to be effective against the regurgitation effects of ion pumps .

Common Misconceptions

Although TSPs are commonly mounted in ion pumps, it is possible to falsely infer that the operation of these pumps is related. This has perhaps also fueled the misconception that a TSP operates in gas phase, i.e. the titanium atoms or clusters from the source collide with molecules from the residual gas, and form gaseous products which are heavier and therefore easier for ion-pumps and turbomolecular pumps to remove. Although this sounds plausible, TSPs are most effective when the residual gas pressure is sufficiently low for the gas to behave ballistically (i.e. the mean free path in the gas is large and thus the probability of collisions within the gas is insignificant compared to the probability or collision with the chamber walls). Credit for the invention of titanium sublimation vacuum pumps has generally been given to Linus Pauling. However, the team of brothers Benjamin and Jeremy Chubb built the first TSP in 1946 in the town of Glenview, IL.

Turbomolecular pump

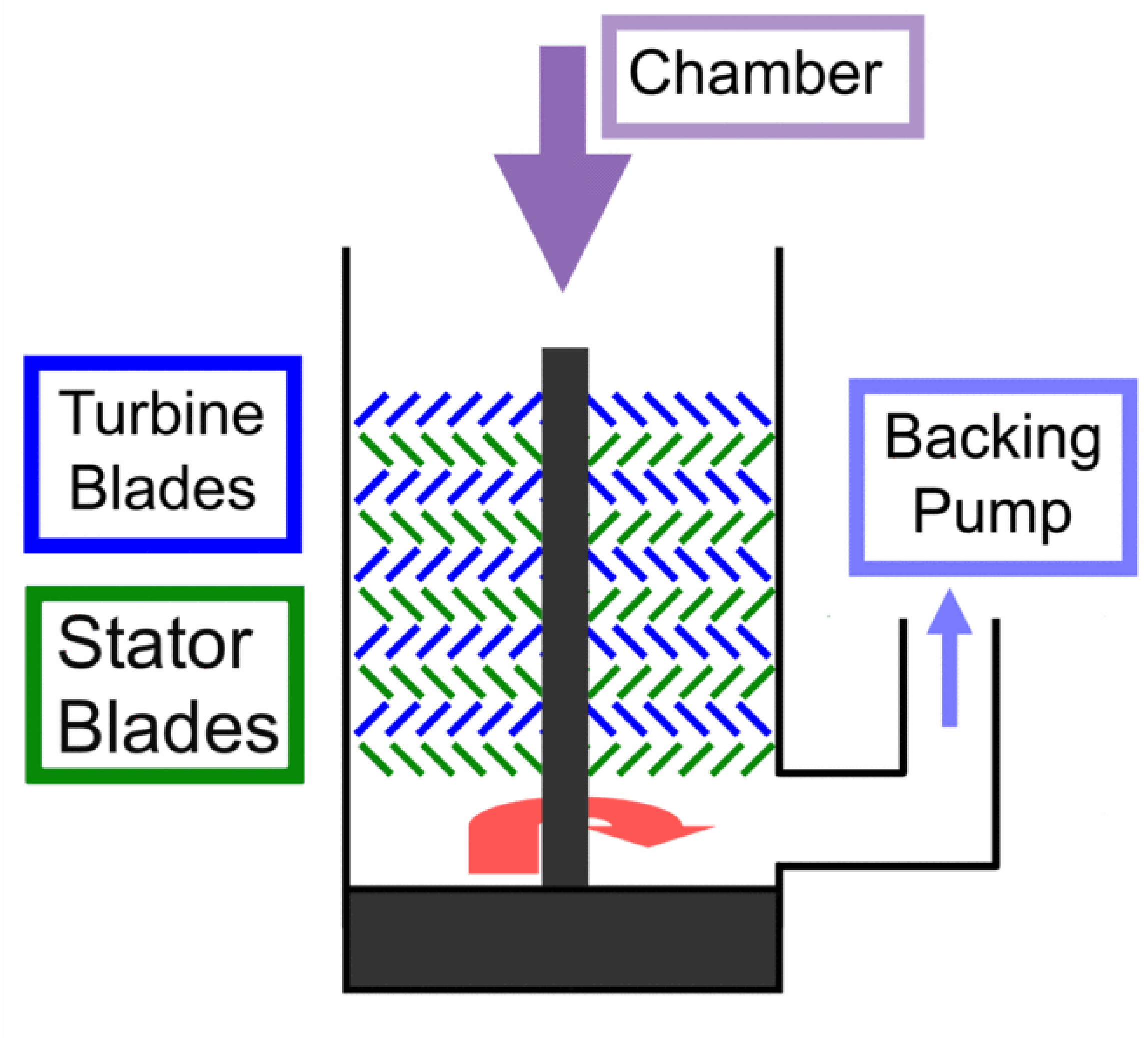
Interior view of a turbomolecular pump

A **turbomolecular pump** is a type of vacuum pump, superficially similar to a turbopump, used to obtain and maintain high vacuum. These pumps work on the principle that gas molecules can be given momentum in a desired direction by repeated collision with a moving solid surface. In a turbomolecular pump, a rapidly spinning turbine rotor 'hits' gas molecules from the inlet of the pump towards the exhaust in order to create or maintain a vacuum.

Operating principles

Most turbomolecular pumps employ multiple stages consisting of rotor/stator pairs mounted in series. Gas captured by the upper stages is pushed into the lower stages and successively compressed to the level of the fore-vacuum (backing pump) pressure. As the gas molecules enter through the inlet, the rotor, which has a number of angled blades, hits the molecules. Thus the mechanical energy of the blades is transferred to the gas molecules. With this newly acquired momentum, the gas molecules enter into the gas transfer holes in the stator. This leads them to the next stage where they again collide with the rotor surface, and this process is continued, finally leading them outwards through the exhaust.

Because of the relative motion of rotor and stator, molecules preferably hit the lower side of the blades. Because the blade surface looks down, most of the scattered molecules will leave it downwards. The surface is rough, so no reflection will occur. A blade needs to be thick and stable for high pressure operation and as thin as possible and slightly bent for maximum compression. For high compression ratios the throat between adjacent rotor blades (as shown in the image) is pointing as much as possible in the forward direction. For high flow rates the blades are at 45° and reach close to the axis.



Schematic of a turbomolecular pump.

Because the compression of each stage is ~10, each stage closer to the outlet is considerably smaller than the preceding inlet stages. This has two consequences. The geometric progression tells us that infinite stages could ideally fit into a finite axial length. The finite length in this case is the full height of the housing as the bearings, the motor, and controller and some of the coolers can be installed inside on the axis. Radially, to grasp as much of the thin gas at the entrance, the inlet-side rotors would ideally have a larger radius, and correspondingly higher centrifugal force; ideal blades would get exponentially thinner towards their tips and carbon fibers should reinforce the aluminium blades. However, because the average speed of a blade affects pumping so much this is done by increasing the root diameter rather than the tip diameter where practical.

Turbomolecular pumps must operate at very high speeds, and the friction heat buildup imposes design limitations. Some turbomolecular pumps use magnetic bearings to reduce friction and oil contamination. Because the magnetic bearings and the temperature cycles allow for only a limited clearance between rotor and stator, the blades at the high pressure stages are somewhat degenerated into a single helical foil each. Laminar flow cannot be used for pumping, because laminar turbines stall when not used at the designed flow. The pump can be cooled down to improve the compression, but should not be so cold as to condense ice on the blades. When a turbopump is stopped, the oil from the backing vacuum may backstream through the turbopump and contaminate the chamber. One way to prevent this is to introduce a laminar flow of nitrogen through the pump. The transition from vacuum to nitrogen and from a running to a still turbopump has to be synchronized precisely to avoid mechanical stress to the pump and overpressure at the exhaust. A thin membrane and a valve at the exhaust should be added to protect the turbopump from excessive back pressure (e.g. after a power failure or leaks in the backing vacuum).

The rotor is stabilized in all of its six degrees of freedom. One degree is governed by the electric motor. Minimally, this degree must be stabilized electronically (or by a diamagnetic material, which is too unstable to be used in a precision pump bearing). Another way (ignoring losses in magnetic cores at high frequencies) is to construct this bearing as an axis with a sphere at each end. These spheres are inside hollow static spheres. On the surface of each sphere is a checkerboard pattern of inwards and outwards going magnetic field lines. As the checkerboard pattern of the static spheres is rotated, the rotor rotates. In this construction no axis is made stable on the cost of making another axis unstable, but all axes are neutral and the electronic regulation is less stressed and will be more dynamically stable. Hall effect sensors can be used to sense the rotational position and the other degrees of freedom can be measured capacitively.

Maximum pressure

A turbomolecular pump made by Edwards with attached vacuum ionization gauge for pressure measurement.

At atmospheric pressure the mean free path of air is about 70 nm. Turbomolecular blades cannot be built with anything close to such a small clearance, so this type of pump stalls if exhausted directly to the atmosphere. Nonetheless Varian, Inc. since 2006 offers a pump where the last stages have blades optimized for zero flow and can pump against a pressure of one atmosphere. Because the low pressure stages are limiting the flow, the high pressure stages can be fixed to zero flow. Theoretically centrifugal pumps could be used, but it is more compact to use a circulating flow between hollow threads in the rotor and the stator. In other cases the exhaust is connected to a backing pump, which produces a pressure low enough for the turbomolecular pump to work efficiently. Typically, this pressure must be below 10 Pa with 1-2 Pa as common averages.

The turbomolecular pump can be a very versatile pump. It can generate many degrees of vacuum from intermediate vacuum (~10^{-2} Pa) up to ultra-high vacuum levels (~10^{-8} Pa).

Multiple turbomolecular pumps in a lab or manufacturing-plant can be connected by tubes to a small backing pump. Automatic valves and diffusion pump like injection into a large buffer-tube in front of the backing pump prevents any overpressure from one pump to stall another pump.

Practical considerations

A turbopump by Pfeiffer attached to a thin film deposition system for organic electronics research

Laws of fluid dynamics do not apply in high vacuum environments. The maximum compression varies linearly with circumferential rotor speed. In order to obtain extremely low pressures down to 1 micropascal, rotation rates of 20,000 to 90,000 revolutions per minute are often necessary. Unfortunately, the compression ratio varies exponentially with the square root of the molecular weight of the gas. Thus, heavy molecules are pumped much more efficiently than light molecules. Most gases are heavy enough to be well pumped but it is difficult to pump hydrogen and helium efficiently.

An additional drawback stems from the high rotor speed of this type of pump: very high grade bearings are required, which increase the cost.

Because turbomolecular pumps only work in molecular flow conditions, a pure turbomolecular pump will require a very large backing pump to work effectively. Thus, many modern pumps have a molecular drag stage such as a Holweck or Gaede mechanism near the exhaust to reduce the size of backing pump required.

History

The turbomolecular pump was invented in 1958 by Becker, based on the older molecular drag pumps developed by Gaede in 1913, Holweck in 1923 and Siegbahn in 1944.

Chapter 9

Vacuum Tube

In electronics, a **vacuum tube**, **electron tube** (in North America), or **thermionic valve** (elsewhere, especially in Britain) is a device that relies on the flow of electric current through a vacuum. Vacuum tubes may be used for rectification, amplification, switching, or similar processing or creation of electrical signals. Vacuum tubes rely on thermionic emission of electrons from a hot filament or *cathode*, that then travel through a vacuum toward a positively-charged anode or *plate*. Additional electrodes interposed between the cathode and anode can alter the current flow, making the device an amplifier.

Vacuum tubes were critical to the development of electronic technology, which drove the expansion and commercialization of radio communication and broadcasting, television, radar, sound reproduction, large telephone networks, analog and digital computers, and industrial process control. Although some of these applications had counterparts using earlier technologies, such as the spark gap transmitter or mechanical computers, it was the invention of the triode vacuum tube and its capability of electronic amplification that made these technologies widespread and practical.

For the most part vacuum tubes have been replaced by solid-state devices such as transistors and other semiconductor devices. Solid-state devices last much longer, are smaller, more efficient, more reliable, and cheaper than equivalent vacuum tube devices. However, tubes still find particular uses where solid state devices have not been developed or are not practical. Tubes are still produced for such applications and to replace those used in existing equipment such as high-power radio transmitters.

Classification

Vacuum tubes with two active elements ("diodes") are used for rectification. Ones with 3 or more elements ("triodes", "tetrodes", etc.) are used for amplification or functions which rely on amplification such as oscillators. Tubes used in consumer electronic equipment are often classified as "receiving tubes," as opposed to the much larger "transmitting tubes" used to generate high power radio signals for transmission.

On the other hand there are vacuum tubes used in different manners, such as cathode ray tubes which create a beam of electrons for display purposes (such as the television picture tube) in addition to more specialized functions such as electron microscopy and electron beam lithography. X-ray tubes are also vacuum tubes. Phototubes and photomultipliers also rely on electron flow through a vacuum, though in this case the emission of electrons from the cathode depends on energy from photons rather than thermionic emission. Since these sorts of "vacuum tubes" have functions other than electronic amplification and rectification they are described in their own articles.

Modern applications

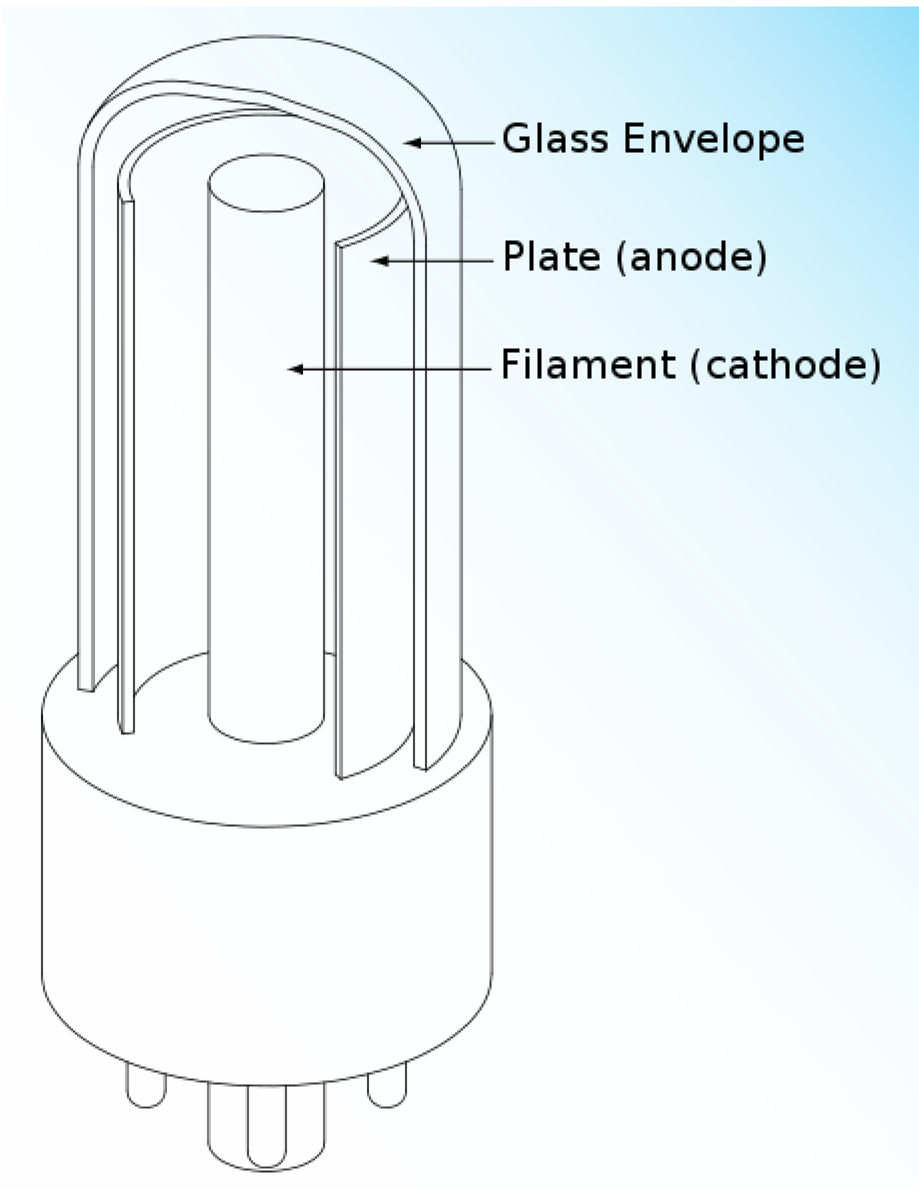
Specialized applications for amplifying vacuum tubes continue to this day, such as the magnetron which is used to generate microwave energy in the household microwave oven and some radar systems. The klystron is commonly deployed by broadcasters as a high-power UHF television transmitting tube. Hi-fi equipment using tubes is still popular among certain audiophiles for its distinct sound signature and other tube equipment is maintained for its aesthetic appeal.

Gas-filled tubes

There are also varieties of current-conducting tubes filled with one or another gas at a higher or lower pressure; the common fluorescent bulb is a familiar example. However certain types such as the voltage regulator tube and thyristor physically resemble commercial vacuum tubes and fit in sockets designed for vacuum tubes. Their distinctive orange, red, or purple glow during operation betrays the presence of gas; electrons flowing in a vacuum do not produce light within that region. Although not properly termed vacuum tubes, they may still be referred to as "electron tubes" as they do perform electronic functions, and are briefly discussed below under "Special-purpose tubes."

Description

A vacuum tube consists of two or more electrodes in a vacuum inside an airtight enclosure. Most tubes have glass envelopes, though ceramic and metal envelopes (atop insulating bases) have also been used. The electrodes are attached to leads which pass through the envelope via an airtight seal. On most tubes, the leads, in the form of pins, plug into a tube socket for easy replacement of the tube. (Tubes were by far the most common cause of failure in electronic equipment, and consumers were expected to be able to replace tubes themselves).

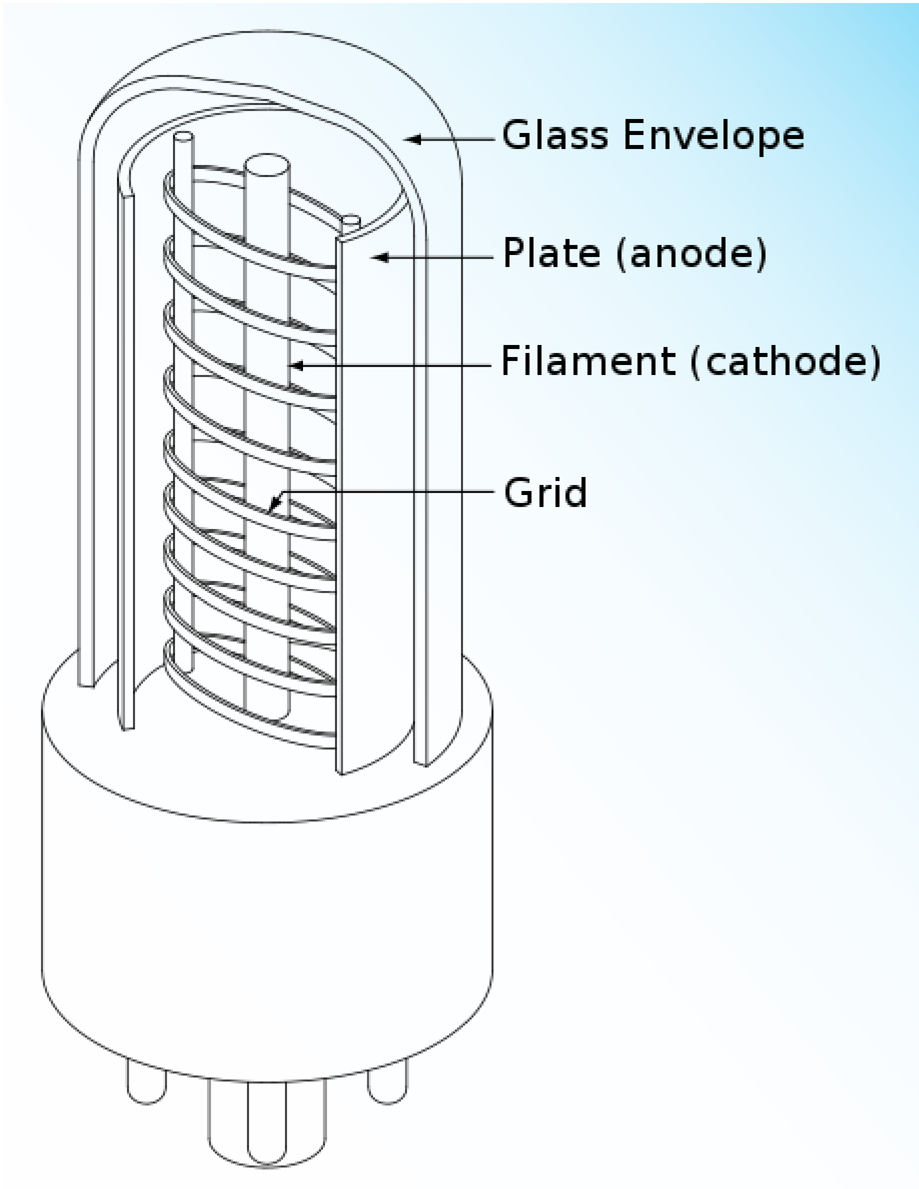


Vacuum tube diode: electrons from the hot cathode flow towards positive anode, but not visa versa.

The earliest vacuum tubes resembled, and in fact evolved from incandescent light bulbs, containing a filament sealed in an evacuated glass envelope. When hot, the filament releases electrons into the vacuum, a process called thermionic emission. These electrons will be drawn to a more positive electrode, the anode or *plate*. The result is a net flow of electrons from filament to plate. However current cannot flow in the reverse direction because the plate is not heated and does not emit electrons. Such a tube with only two electrodes is termed a diode, and is used for rectification. Since current can only pass in one direction, such a diode (or *rectifier*) will convert AC to DC. This is therefore used in a DC power supply, but is also used as a demodulator of amplitude modulated (AM) radio signals, and similar functions.

While early tubes used the directly-heated filament as the cathode, most (but not all) more modern tubes employed indirect heating. A separate element was used for the cathode. Inside the cathode, and insulated from it, was the filament or *heater*. The heater warmed the cathode sufficiently to undergo thermionic emission, but avoided any electrical connection. This allowed the tubes in a radio set to be heated through a common circuit, while allowing each cathode to arrive at a voltage independly of the others, removing an unwelcome constraint on circuit design.

During operation vacuum tubes require constant heating of the filament, so that they require considerable power even when amplifying signals at the microwatt level. In most amplifiers further power is consumed due to the quiescent current between the cathode and the plate (anode), resulting in heating of the plate. In a power amplifier heating of the plate can be quite considerable, and has a potential for self-destruction if the tube is driven beyond its safe limits. Since the tube requires a vacuum to operate, convection cooling of the plate is not generally possible (except in special applications where the anode forms a part of the vacuum envelope; this is avoided in consumer products due to the shock hazard it entails). Thus anode cooling occurs mainly through black-body radiation.



Vacuum tube triode: voltage applied to the grid controls plate current.

With the exception of diode tubes, another electrode, called a control grid, is placed between the cathode and the plate. The vacuum tube is then known as a "triode." With additional grids they are called tetrode, pentode, etc. These intervening electrodes are all called *grids* as they are not solid electrodes but sparse elements through which electrons can pass on their way to the plate. The control grid (and sometimes other grids) turn the diode into a *voltage-controlled device*, that is, the voltage that is applied to the control grid will affect the current flow between the cathode and the plate. A negative electrostatic field from the control grid repels electrons emitted by the cathode, rather than allowing them to continue toward the plate, thus reducing or even completely stopping the current flow. As long as the control grid stays more negative than the cathode, essentially no current flows into it, yet a change of several volts on the control grid is sufficient to make a large difference in the plate current, possibly changing the output by hundreds of volts (depending on the load of the circuit). The solid-state device most closely resembling the pentode tube is the JFET, although vacuum tubes typically operate at over a hundred volts, unlike most semiconductors in most applications.

History and development

Early RCA triode vacuum tube, type 808

The 19th century saw increasing research with evacuated tubes, such as the Geissler and Crookes tubes. Famous scientists who experimented with such tubes included Thomas Edison, Eugen Goldstein, Nikola Tesla, and Johann Wilhelm Hittorf among many others. With the exception of early light bulbs, such tubes were only used in scientific research or as novelties. The groundwork laid by these scientists and inventors, however, was critical to the development of subsequent vacuum tube technology.

Although thermionic emission was originally reported in 1873 by Frederick Guthrie, it was Thomas Edison's 1884 investigation that spurred future research, the phenomenon thus becoming known as the "Edison Effect." Edison patented what he found, but he did not understand the underlying physics, nor did he have an inkling of the potential value of the discovery. It wasn't until the early 20th century that the rectifying property of such a device was utilized, most notably by John Ambrose Fleming who used the diode tube to detect (demodulate) radio signals. Lee De Forest's 1906 "audion" was also developed as a radio detector, and soon led to the developement of the triode tube. This was essentially the first electronic amplifier, leading to great improvements in telephony (such as the first coast-to-coast telephone line in the US) and revolutionizing the technology used in radio transmitters and receivers. The electronics revolution of the 20th century arguably began with the invention of the triode vacuum tube.

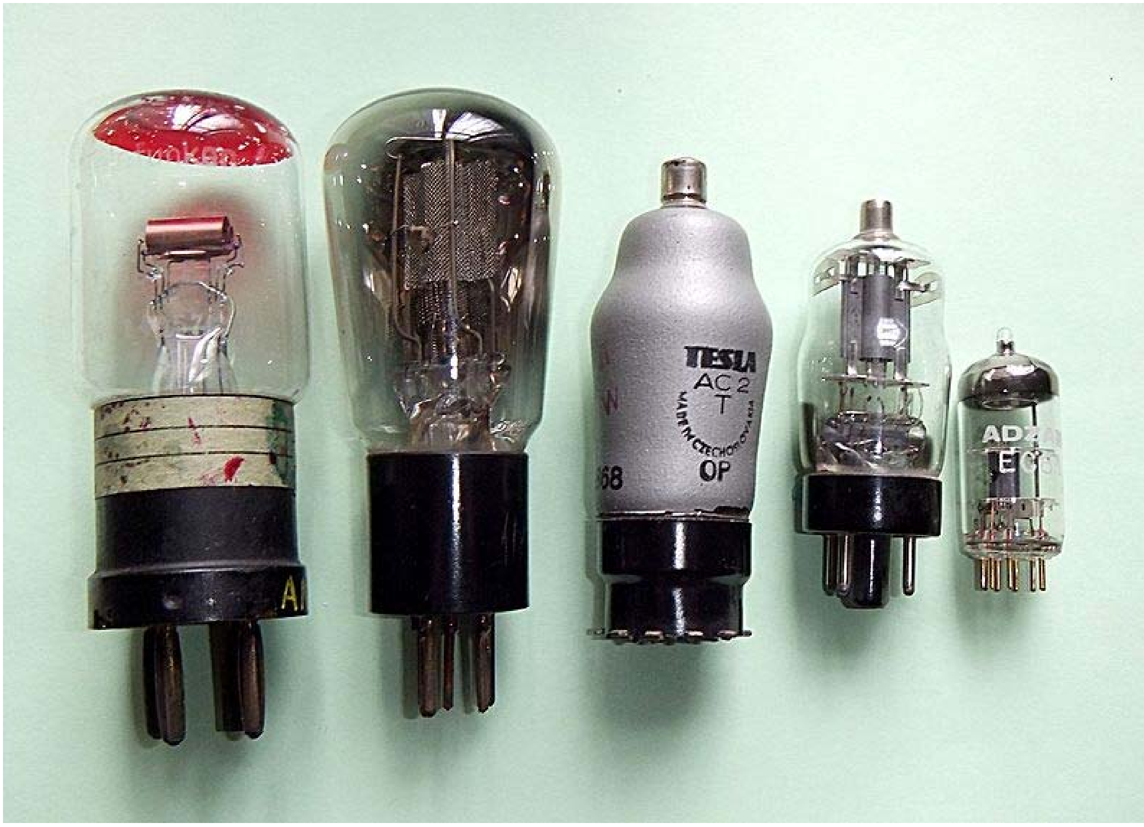
Diodes

The English physicist John Ambrose Fleming worked as an engineering consultant for firms including Edison Telephone and the Marconi Company. In 1904, as a result of experiments conducted on Edison effect bulbs imported from the USA, he developed a device he called an "oscillation valve" (because it passes current in only one direction). The heated filament, or cathode, was capable of thermionic emission of electrons that would flow to the *plate* (or *anode*) when it was at a higher voltage. Electrons, however, could not pass in the reverse direction because the plate was not heated and thus not capable of thermionic emission of electrons.

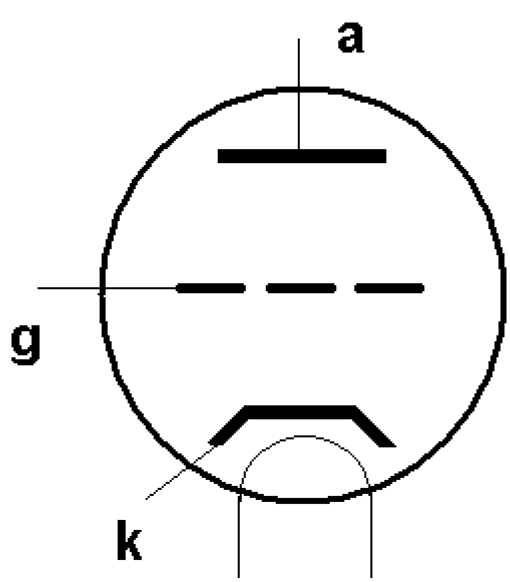
Later known as the Fleming valve, it could be used as a rectifier of alternating current and as a radio wave detector. This greatly improved the crystal set which rectified the radio signal using an early solid-state diode based on a crystal and a so-called cat's whisker. Unlike modern semiconductors, such a diode required painstaking adjustment of the contact to the crystal in order for it to rectify. The diode tube was a reliable alternative for rectifying radio signals. Higher power diode tubes or *power rectifiers* found their way into power supply applications until they were eventually replaced by silicon rectifiers in the 1960s.

Triodes

Vacuum tube with plate cut open revealing grid.



Triodes as they evolved over 40 years of tube manufacture, from the RE16 in 1918 to a 1960's era miniature tube.



Triode symbol. From top to bottom: plate (anode), control grid, cathode, heater (filament)

Originally, the only use for tubes in radio circuits was for rectification, not amplification. In 1906 Robert von Lieben filed for a patent for a cathode ray tube which included magnetic deflection. This could be used for amplifying audio signals and was intended for use in telephony equipment. He would later go on to help refine the triode vacuum tube.

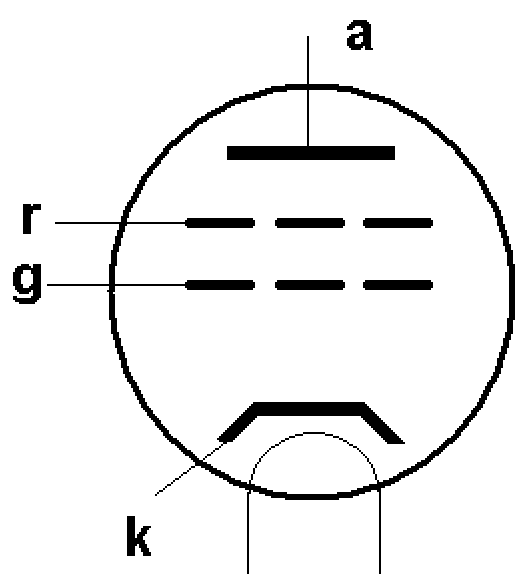
However it was Lee De Forest who in 1907 is credited with inventing the triode tube while continuing experiments to improve his original Audion tube, a crude forerunner of the triode. By placing an additional electrode in between the filament (cathode) and plate, he discovered the ability of the resulting device to amplify signals of all frequencies. As the voltage applied to the so-called control grid (or simply "grid") was lowered from the cathode's voltage to somewhat more negative voltages, the amount of current flowing from the filament to the plate would be reduced. The negative electrostatic field created by the grid in the vicinity of the cathode would inhibit thermionic emission and reduce the current to the plate. Thus a few volts difference at the grid would make a large change in the plate current and could lead to a much larger voltage change at the plate, resulting in voltage and power amplification. In 1907, De Forest filed for a patent for such a three-electrode version of his original Audion tube for use as an electronic amplifier in radio communications. This eventually became known as the triode.

De Forest's device was not strictly a vacuum tube, as he erroneously believed that it depended on the presence of residual gas remaining after evacuation. The De Forest company, in its Audion leaflets, even warned against operation which might lead to too high a vacuum! The Finnish inventor Eric Tigerstedt significantly improved on the original triode design in 1914, while working on his sound-on-film process in Berlin, Germany. The first true vacuum triodes in production were the Pliotrons developed by Irving Langmuir at the General Electric research laboratory (Schenectady, New York) in 1915. Langmuir was one of the first scientists to realize that a harder vacuum would improve the amplifying behaviour of the triode. Pliotrons were closely followed by the French 'R' Type which was in widespread use by the allied military by 1916. These two types were the first true *vacuum* tubes; early diodes and triodes performed as such despite a rather high residual gas pressure. Techniques to produce and maintain better vacuums in tubes were then developed. Historically, vacuum levels in production vacuum tubes typically ranged from 10 µPa down to 10 nPa.

The non-linear operating characteristic of the triode caused early tube audio amplifiers to exhibit harmonic distortions at low volumes. This is not to be confused with the so-called overdrive distortion that tube amplifiers exhibit when driven beyond their linear region (known as the tube sound). To remedy the triode's nonlinear characteristics, engineers plotted curves of the applied grid voltage and resulting plate currents, and discovered that there was a range of grid voltages allowing for relatively linear operation. In order to use this range, a negative voltage had to be applied to the grid to place the tube in the "middle" of the linear area with no signal applied. This was called the idle condition, and the plate current at this point the "idle current". Today this current would be called the quiescent or bias current. The controlling voltage was superimposed onto this fixed "bias" voltage, resulting in a linear variation of plate current in response to both positive and negative variation of the input voltage around that point. This concept is called *grid bias*. Many early radio sets had a third battery called the "C battery" (not to be confused with the modern C cell) whose positive terminal was connected to the cathode of the tubes (or "ground" in most circuits) and whose negative terminal supplied this bias voltage to the grids of the tubes. More modern circuits used cathode biasing in lieu of a separate negative power supply.

When triodes were first used in radio transmitters and receivers, it was found that tuned amplification stages had a tendency to oscillate unless their gain was very limited. This was due to the parasitic capacitance between the plate (the amplifier's output) and the control grid (the amplifier's input), known as the Miller capacitance. Eventually the technique of *neutralization* was developed whereby the RF transformer connected to the plate would include an additional winding in the opposite phase. This winding would be connected back to the grid through a small capacitor, and when properly adjusted would cancel the Miller capacitance. This technique was employed and led to the success of the Neutrodyne radio during the 1920s. However neutralization required careful adjustment and proved unsatisfactory when used over a wide ranges of frequencies.

Tetrodes and pentodes



Tetrode symbol. From top to bottom: plate (anode), screen grid, conrol grid, cathode, heater (filament)

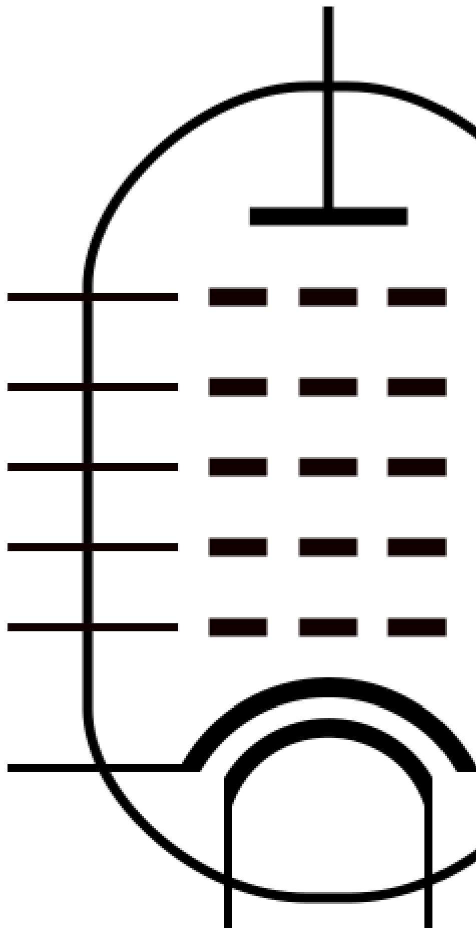
In order to combat the stability problems and limited voltage gain due to the Miller effect, the physicist Walter H. Schottky invented the tetrode tube in 1919. He showed that the addition of a second grid, located between the control grid and the plate, known as the *screen grid*, could solve these problems. ("Screen" in this case refers to electrical "screening" or shielding, not physical construction: all "grid" electrodes in between the cathode and plate are "screens" of some sort rather than solid electrodes since they must allow for the passage of electrons directly from the cathode to the plate). A positive voltage slightly lower than the plate voltage was applied to it, and was bypassed (for high frequencies) to ground with a capacitor. This arrangement decoupled the anode and the control grid, essentially eliminating the Miller capacitance and its associated problems. Consequently higher voltage gains from a single tube became possible, reducing the number of tubes required in many circuits. This two-grid tube is called a *tetrode*, meaning four active electrodes, and was common by 1926.

However, the tetrode has one new problem. In any tube, electrons strike the anode with sufficient energy to cause the emission of electrons from its surface. In a triode this so-called secondary emission of electrons is not important since they are simply re-captured by the more positive anode. But in a tetrode they can be captured by the screen grid since it is at also at a high voltage, thus robbing them from the plate current and reducing the amplification of the device. Since secondary electrons can outnumber the primary electrons, in the worst case, particularly as the plate voltage dips below the screen voltage, the plate current can actually go down with increasing plate voltage. This is termed negative resistance and can itself cause instability. This is the so-called "tetrode kink". Another consequence of secondary emission is that in extreme cases the current reaching the screen grid can cause it to overheat to the point of destroying the tube.

Vacuum tubes in an Australian radio of the late 1930s

The solution was to add one more grid in between the screen grid and the plate, called the suppressor grid (since it supressed secondary emission current toward the screen grid). This grid was held at the cathode (or "ground") voltage and its negative voltage (relative to the anode) electrostatically repelled secondary electrons so that they would be collected by the anode after all. This three-grid tube is called a pentode, meaning five electrodes. The pentode was invented in 1928 by Bernard D. H. Tellegen and became generally favoured over the simple tetrode. A refinement of the tetrode or pentode for power applications is the beam tetrode or "beam power tube", discussed berlow.

Multifunction configurations

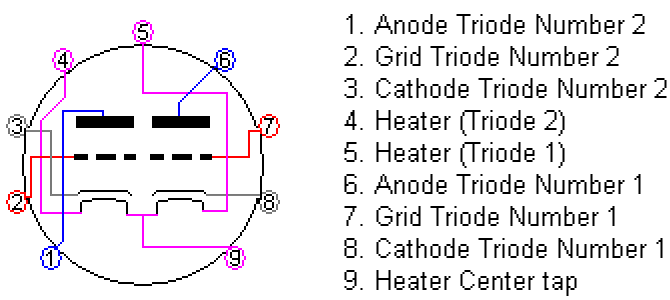


The pentagrid converter contained no less than 5 grids in between the cathode and the plate.

Superheterodyne receivers require a local oscillator and mixer, which required two tubes. With the development of the pentagrid converter, these functions were combined inside a single tube which applied the RF signal to the control grid, but also implemented the local oscillator using additional grids. Various alternatives such as using a combination of a triode with a hexode and even an octode have been used for this purpose. The additional grids include both control grids (at a low potential) and screen grids (at a high voltage). Many designs used such a screen grid as a second 'leaky' plate to provide feedback for the oscillator oscillator function, whose current was added to that of the incoming radio frequency signal. Due to the large oscillating signal nonlinearity of the tube response caused frequency mixing, seen on the plate current (output) of such a "converter" circuit. The difference frequency between that of the incoming signal and that of the oscillator was selected by a tuned transformer, becoming the input to the receivers's intermediate frequency (IF) amplifier.

The pentagrid converter such as the 12BE6 thus became widely used in AM receivers including the minature tube version of the "All American Five". Octodes such as the 7A8 were rarely used in the US, but much more common in Europe particularly in battery operated radios where the lower power consumption was an advantage.

To further reduce the cost and complexity of radio equipment, two separate vacuum tubes could be combined in the bulb of a single tube, a so-called *multisection tube*. An early example was the Loewe 3NF. This 1920s device had 3 triodes in a single glass envelope together with all the fixed capacitors and resistors required to make a complete radio receiver. As the Loewe set had only one tube socket, it was able to substantially undercut the competition since, in Germany, state tax was levied by the number of sockets. However, reliability was compromised, and production costs for the tube were much greater. In a sense, these were akin to integrated circuits. In the US, Cleartron briefly produced the "Multivalve" triple triode for use in the Emerson Baby Grand receiver. This Emerson set also had a single tube socket, but because it used a four-pin base, the additional element connections were made on a "mezzanine" platform at the top of the tube base.



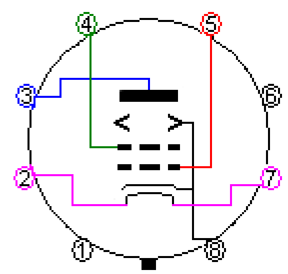
Popular 12AX7 dual triode

Compactron tube: 12AE10, dual pentode

By 1940 multisection tubes had become commonplace. There were constraints, however, due to patents and other licencing considerations. Constraints due to the number of external pins (leads) often forced the functions to share some of those external connections such as their cathode connections (in addition to the heater connection). The RCA Type 55 was a double diode triode used as a detector, automatic gain control rectifier and audio preamp in early AC powered radios. These sets often included the 53 Dual Triode Audio Output. Another early type of multi-section tube, the 6SN7, is a "dual triode" which, for most purposes, can perform the functions of two triode tubes, while taking up half as much space and costing less. The 12AX7 is a dual high voltage gain (or *high mu*) triode in a minature enclosure, and became widely used in audio signal amplifiers, instruments, and.guitar amplifiers.

The introduction of the miniature tube base (see below) which could have 9 pins, also allowed many other multi-section tubes, such as the 6GH8 triode + pentode. Along with a host of similar tubes, the 6GH8 was quite popular in television receivers. Some color TV sets used exotic types like the 6JH8 which had two plates and beam deflection electrodes (it was known as the 'sheet beam' tube). Vacuum tubes used like this were designed for demodulation of synchronous signals, an example of which is color demodulation for television receivers. The desire to include additional functions in one envelope resulted in the General Electric Compactron which had 12 pins (minature tubes had only 7 or 9 pins). A typical example, the 6AG11, contained two triodes and two diodes.

Beam power tubes



Beam power tube symbol and pinout for 6L6

The beam power tube is usually a tetrode with the addition of beam-forming electrodes, which take the place of the suppressor grid. These angled plates focus the electron stream onto certain spots on the anode which can withstand the heat generated by the impact of massive numbers of electrons, while also providing pentode behavior. The positioning of the elements in a beam power tube uses a design called "critical-distance geometry", which minimizes the "tetrode kink", plate-grid capacitance, screen-grid current, and secondary emission effects from the anode, thus increasing power conversion efficiency. The control grid and screen grid are also wound with the same pitch, or number of wires per inch.

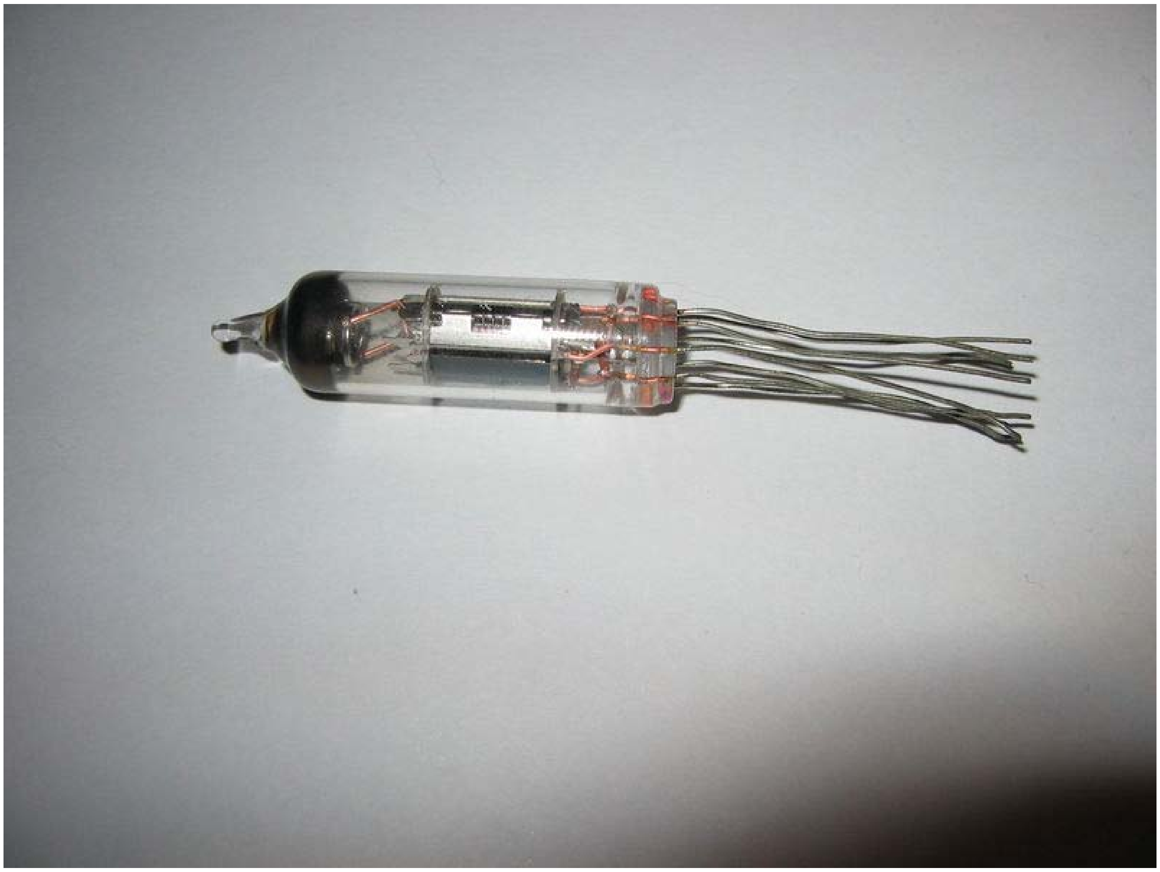
6L6 tubes in glass envelopes

Aligning the grid wires also helps to reduce screen current, which represents wasted energy. This design helps to overcome some of the practical barriers to designing high-power, high-efficiency power tubes. 6L6 was the first popular beam power tube, introduced by RCA in 1936. Corresponding tubes in Europe were the KT66, KT77 and KT88 by GEC (the KT standing for "Kinkless Tetrode").

Variations of the 6L6 design are still widely used in tube guitar amplifiers, making it one of the longest lived electronic device families in history. Similar design strategies are used in the construction of large ceramic power tetrodes used in radio transmitters.

Miniature tubes

Miniature tube, alongside euro coin



Subminiature CV4501 tube, 35 mm long x 10 mm diameter (excluding leads).

RCA 6DS4 "Nuvistor" triode, ca. 20 mm high by 11 mm diameter.

Early tubes used a metal or glass envelope atop an insulating bakelite base. In 1938 a technique was developed to instead use an all glass construction with the pins fused in the glass base of the envelope. This was used in the design of a much smaller tube outline, known as the miniature tube, having 7 or 9 pins. Making tubes smaller reduced the voltage that they could work at, and also the power of the filament. Miniature tubes became predominant in consumer applications such as radio receivers and hi-fi amplifiers. However the larger older styles continued to be used especially as higher power rectifiers, in higher power audio output stages and as transmitting tubes.

Subminiature tubes with a size roughly that of half a cigarette were used in hearing-aid amplifiers. These tubes did not have pins plugging into a socket but were soldered in place. The "acorn" valve (named due to its shape) was another such example. Another very small tube style was called the nuvistor. About the size of a thimble, these metal cased tubes were made small not mainly for compactness, but for use at very high frequencies, notably in UHF television tuners.

Improvements in construction and performance

The very earliest vacuum tubes strongly resembled incandescent light bulbs and were made by lamp manufacturers, who had the equipment for manufacture of glass envelopes and the powerful vacuum pumps required to evacuate the enclosures. After World War I, specialized manufacturers using more economical construction methods were set up to fill the growing demand for broadcast receivers. Bare tungsten filaments operated at a temperature of around 2200 °C. The development of oxide-coated filaments in the mid 1920s reduced filament operating temperature to a dull red heat (around 700 °C), which in turn reduced thermal distortion of the tube structure and allowed closer spacing of tube elements. This in turn improved tube gain, since the gain of a triode is inversely proportional to the spacing between grid and cathode.

Indirectly heated cathodes

The desire to power electronic equipment using AC mains power faced a difficulty with respect to the powering of the tubes' filaments, as these were also the cathode of each tube. Powering the filaments directly from a power transformer would introduce 50 or 60 Hz hum into audio stages using tubes whose filaments were powered in such a manner. The invention of the "equi-potential cathode" reduced this problem, with the filaments being powered by a balanced AC power transformer winding having a grounded center tap.

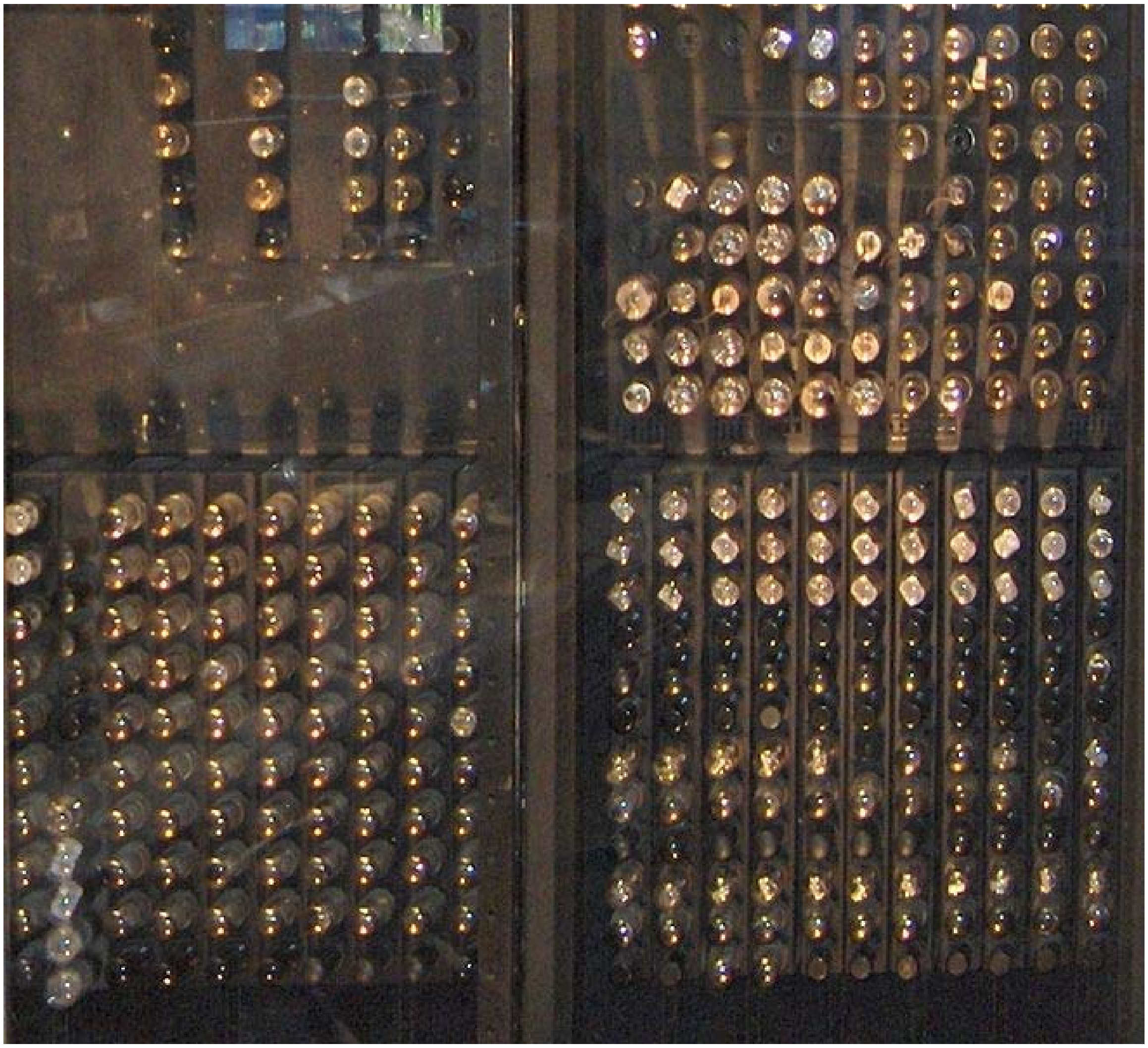
A superior solution, and one which allowed each cathode to "float" at a different voltage, was that of the indirectly-heated cathode. Now, a filament inside a cylinder of oxide-coated nickel, provided for a cathode electrically isolated from the filament which could then just as well be powered by AC. In such tubes, the filament is frequently referred to as a *heater* to distinguish it as an inactive element. In the 1930s indirectly heated cathode tubes became widespread in equipment using AC power. However directly heated filament tubes continued to prevail in battery operated equipment, as the power requirements for these filaments were substantially lower than required by the heaters used to heat cathodes indirectly.

World War II

Near the end of World War II, to make radios more rugged, some aircraft and army radios began to integrate the tube envelopes into the radio's cast aluminium or zinc chassis. The radio became just a printed circuit with non-tube components, soldered to the chassis that contained all the tubes. During WWII in 1942, rugged metal vacuum tubes were mounted in anti-aircraft shells. These proximity fuzes made anti-aircraft shells 6 times more effective. In the fall of 1944, artillery shells with proximity fuses were used.

The tiny tubes were later known as "subminiature" types. They were widely used in 1950s military and aviation electronics.

Use in early electronic computers



The 1946 ENIAC computer used 17,468 vacuum tubes and consumed 150kW of power.

While the development of vacuum tubes made electronic computing possible for the first time, the cost and reliability of early tubes made such developments rather impractical. It was only the pressure of World War II that led to the development of Colossus and early electronic computers using tubes. The general practicality of electronic computers was only realized with the development of transistors over a decade later.

Colossus

Colossus (and its successor Colossus Mk2) was built by the British during World War II to substantially speed up the task of breaking the German high level Lorenz encryption. Based on 1500 vacuum tubes, Colossus replaced an earlier machine based on relay and switch logic (the Heath Robinson). Colossus was able to break in a matter of hours messages that had previously taken several weeks. Colossus Mk2 used a total of around 2000 vacuum tubes. Colossus was the first ever use of vacuum tubes on such a large scale for a single machine. The largest project previously had used just 150 tubes and had proven to be extremely unreliable. The main design problem at Colossus's inception was how to make vacuum tube based equipment reliable when the tubes were used in large numbers.

The Colossus computer's designer, Dr. Tommy Flowers, had a theory that most of the unreliability was caused during power down and (mainly) power up. Once Colossus was built and installed, it was switched on and left switched on running from dual redundant diesel generators (the wartime mains supply being considered too unreliable). The only time it was switched off was for conversion to the Colossus Mk2 and the addition of another 500 or so tubes. Another 9 Colossus Mk2s were built, and all 10 machines ran with a surprising degree of reliability. The 10 Colossi consumed 15 kilowatts of power each, 24 hours a day, 365 days a year—nearly all of it for the tube heaters.

Whirlwind

To meet the reliability requirements of the early digital computer Whirlwind, it was necessary to build special "computer vacuum tubes" with extended cathode life. The problem of short lifetime was traced to evaporation of silicon, used in the tungsten alloy to make the heater wire easier to draw. Elimination of the silicon from the heater wire alloy (and paying extra for more frequent replacement of the wire drawing dies) allowed production of tubes that were reliable enough for the Whirlwind project. The tubes developed for Whirlwind later found their way into the giant SAGE air-defense computer system. High-purity nickel tubing and cathode coatings free of materials that can poison emission (such as silicates and aluminium) also contribute to long cathode life. The first such "computer tube" was Sylvania's 7AK7 of 1948. By the late 1950s it was routine for special-quality small-signal tubes to last for hundreds of thousands of hours, if operated conservatively. This increased reliability also made mid-cable amplifiers in submarine cables possible.

Heat generation and transfer

The anode of this transmitting triode has been designed to dissipate up to 500W of heat

A considerable amount of heat is produced when tubes operate, both from the filament (heater) but also from the stream of electrons bombarding the plate. The requirements for heat removal can significantly change the appearance of high-power vacuum tubes. Although the miniature tube style became predominant in consumer equipment, high power audio amplifiers and rectifiers would still require the larger "octal" style of enclosure. Transmitting tubes could be much larger still.

Most tubes produce heat from two sources during operation. The first source is the filament or heater. Some tubes contain a *directly heated cathode*. This is a filament similar to an incandescent electric lamp; some types glow brightly like a lamp, but most glow a dim orange-red. The "bright emitter" types possess a tungsten filament alloyed with 1-3 % thorium which reduces the work function of the metal, giving it the ability to emit sufficient electrons at about 2000 degrees Celsius. The "dull emitter" types also possess a tungsten filament but it is coated in a mixture of calcium, strontium and barium oxides, which emit electrons easily at much lower temperatures due to a monolayer of mixed alkali earth metals coating the tungsten; these only reach 800-1000 degrees Celsius.

The second form of cathode is the *indirectly heated* form which usually consists of a nickel cylinder, coated on the outside with the same strontium, calcium, barium oxide mix used in the "dull emitter" directly heated types. Inside the cylinder is a tungsten filament to heat it. This filament is usually uncoiled and coated in a layer of alumina (aluminium oxide) in order to insulate it electrically from the actual cathode. This form of construction allows for a much greater electron emitting area and allows the cathode to be held at a potential difference, typically 150 volts more positive than the heater or 50 volts more negative than the heater. For small-signal tubes such as used in radio receivers, heaters consume between 50 mW and 5 watts, (directly heated), or between 500 mW and 8 watts for indirectly heated types. Thus even a small signal amplifier might consume a watt of power just to warm its heater, compared to the milliwatts (or less) that a modern semiconductor amplifier would require for the same function. Even in power amplifiers the filament power may be responsible for an appreciable reduction in efficiency.

The second source of heat is generated at the plate, as electrons accelerated by the anode voltage strike the plate, depositing their kinetic energy on it and raising its temperature. In tubes used in power amplifiers or transmitter output stages, this source of heat will far exceed the power due to the cathode heater. The plates of improperly operated or overloaded beam power tubes can sometimes become visibly red hot; this should never occur under normal operation of consumer electronics and is a precursor to tube failure.

Heat escapes the device by black body radiation from the anode/plate as infrared radiation. Convection is not possible in most tubes since the anode is surrounded by vacuum. Considerations of heat removal can affect the overall appearance of some tubes. The anode or plate is often treated to make its surface less shiny and darker in the infrared. The screen grid may also generate considerable heat, which is radiated toward the plate which must reradiate that additional heat along with the heat it generates itself. Limits to screen grid dissipation, in addition to plate dissipation, are listed for power devices. If these are exceeded then tube failure is likely.

Tubes used as power amplifier stages for radio transmitters may have additional heat exchangers, cooling fans, radiator fins, or other measures to improve heat transfer at the anode. High power transmitting tubes may have the surface of their anodes external to the tube, allowing for water cooling or evaporative cooling. Such a water cooling system must be electrically isolated to withstand the high voltage present on the anode.

Tubes which generate rather little heat, such as the 1.4 volt filament directly heated tubes designed for use in battery powered equipment, often have shiny metal anodes. 1T4, 1R5 and 1A7 are examples. Gas filled tubes such as thyratrons may also use a shiny metal anode since the gas present inside the tube allows for convection of heat from the anode to the glass enclosure.

The outer electrode in most tubes is the anode. Some small signal types, such as sharp and remote cut-off R.F. and A.F. pentodes and some pentagrid converters have a shield fitted around all the electrodes enclosing the anode. This shield is sometimes a solid metal sheet, treated to make it dull and gray so that it can itself reradiate heat generated from within. Sometimes it is fabricated from expanded metal mesh, acting as a Faraday cage but allowing sufficient infrared radiation from the anode to escape. Types 6BX6/EF80 and 6BK8/EF86 are typical examples of this shielded type using expanded mesh. Types 6AU6/EF94 and 6BE6/EK90 are examples which use a gray sheet metal cylindrical shield.

Tube packaging

Metal cased tubes with "octal" bases

High power GS-9B triode transmitting tube with heat sink at bottom.

Most modern tubes have glass envelopes, but metal, fused quartz (silica), and ceramic have also been used. The first version of the 6L6 used a metal envelope sealed with glass beads, while a glass disk fused to the metal was used in later versions. Metal and ceramic are used almost exclusively for power tubes above 2 kW dissipation. The nuvistor was a modern receiving tube using a very small metal and ceramic package.

Tubes have always had their internal elements connected to external circuitry using pins at their base which plug into a socket. After all, tubes needed to be replaced rather frequently unlike modern semiconductor devices which are mostly soldered in place. Subminiature tubes were produced using wire leads rather than sockets, however these were restricted to rather specialized applications. In addition to the connections at the base of the tube, many early triodes connected the grid using a metal cap at the top of the tube; this was done in order to reduce stray capacitance between the grid and the plate leads. Tube caps were also used for the plate (anode) connection, particularly in transmitting tubes and tubes using a very high plate voltage.

High power tubes such as transmitting tubes have packages designed more to enhance heat transfer. In some tubes, the metal envelope is also the anode. The 4CX1000A is an external anode tube of this sort. Air is blown through an array of fins attached to the anode, thus cooling it. Power tubes using this cooling scheme are available up to 150 kW dissipation. Above that level, water or water-vapor cooling are used. The highest-power tube currently available is the Eimac 4CM2500KG, a forced water-cooled power tetrode capable of dissipating 2.5 megawatts. (By comparison, the largest power transistor can only dissipate about 1 kilowatt.)

Special-purpose tubes

Voltage regulator tube in operation. Low pressure gas within tube glows due to current flow.

Some special-purpose tubes are constructed with particular gases in the envelope. For instance, voltage regulator tubes contain various inert gases such as argon, helium or neon, and take advantage of the fact that these gases will ionize at predictable voltages. The thyratron is a special-purpose tube filled with low-pressure gas or mercury vapor. Like vacuum tubes, it contains a hot cathode and an anode, but also a control electrode which behaves somewhat like the grid of a triode. When the control electrode starts conduction, the gas ionizes after which the control electrode no longer can stop the current; the tube "latches" into conduction. Removing plate (anode) voltage lets the gas de-ionize, restoring its non-conductive state. Some thyratrons can carry large currents for their physical size. One example is the miniature type 2D21, often seen in 1950s jukeboxes as control switches for relays. A cold-cathode version of the thyratron, which uses a pool of mercury for its cathode, is called an Ignitron. Some ignitrons can switch thousands of amperes. Thyratrons containing hydrogen have a very consistent time delay between their turn-on pulse and full conduction, and have long been used in radar transmitters. Thyratrons behave much like modern silicon-controlled rectifiers, also known as thyristors in reference to their tube predecessor.

An extremely specialized tube is the Krytron, which is used for extremely precise, rapid high-voltage switching. Due to their intended purpose, the initiation of the precise sequence of detonations used to set off a nuclear weapon, they are heavily controlled at an international level.

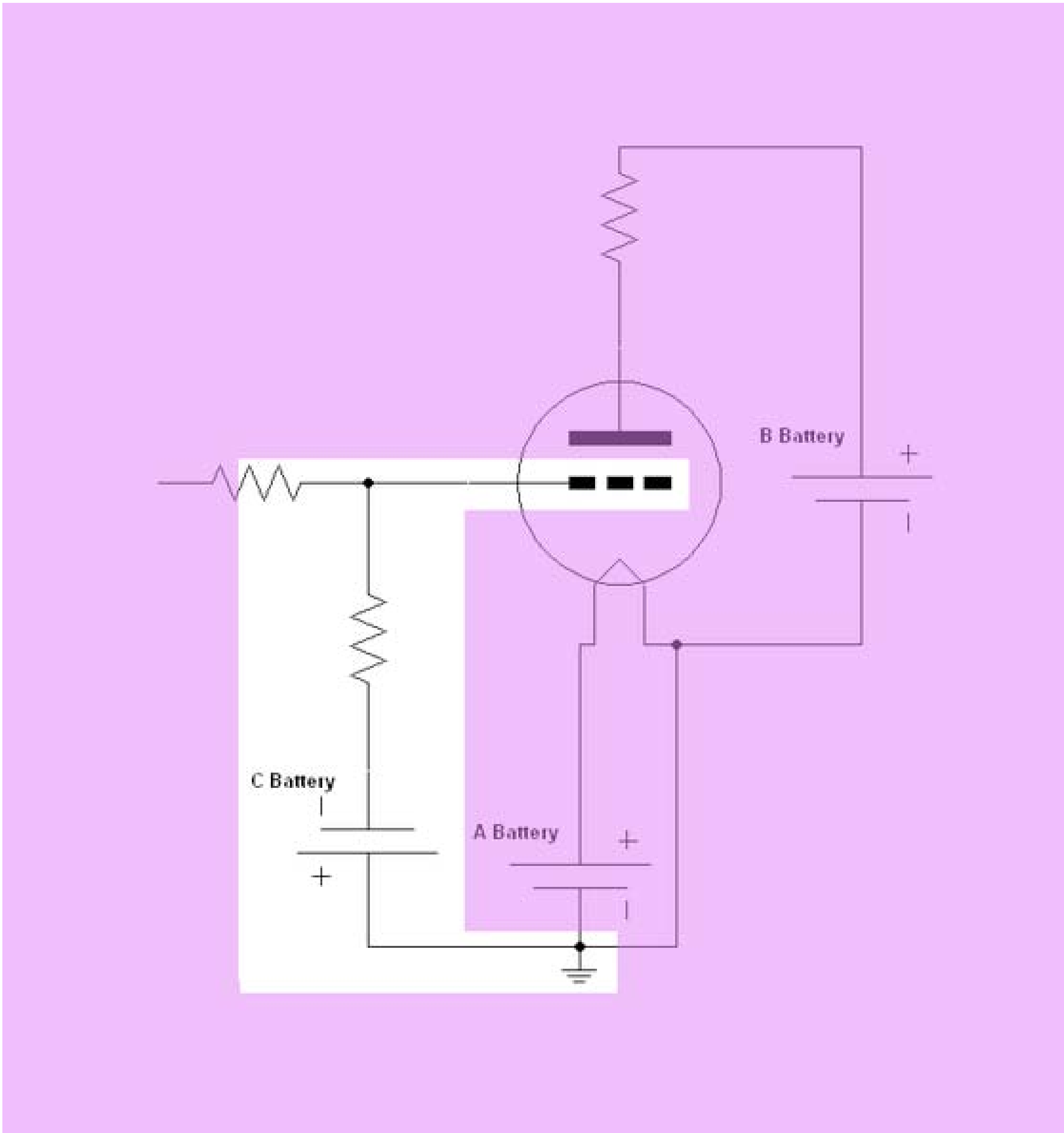
X-ray tubes are used in medical imaging among other uses. X-ray tubes used for continuous duty operation in fluoroscopy and CT imaging equipment may use a focused cathode and a rotating anode to dissipate the large amounts of heat thereby generated. They are housed in an aluminum housing which is filled with oil for cooling. Nuclear medicine imaging equipment and liquid scintillation counters require photomultiplier tube arrays to detect scintillation due to ionizing radiation; the photomultiplier is a rare example of a vacuum tube which doesn't employ thermionic emission.

Powering the tube

Batteries

Batteries provided the voltages required by tubes in early radio sets. Three different voltages were generally required, using three different batteries designated as the **A**, **B**, and **C** battery. The "A" batteries or LT (low-tension) battery provided the filament voltage. Tube heaters were designed for single, double or triple-cell lead-acid batteries, giving nominal heater voltages of 2 V, 4 V or 6 V. In portable radios, dry batteries were sometimes used with 1.5 or 1 V heaters. Reducing filament consumption improved the life span of batteries. By 1955 towards the end of the tube era, tubes using only 50 mA down to as little as 10 mA for the heaters had been developed.

The plate voltage was provided by the "B" batter or the HT (high-tension) supply or battery. These were generally of dry cell construction and typically came in 22.5, 45, 67.5, 90 or 135 volt versions.



Batteries for a vacuum tube circuit. The C battery is highlighted.

Early sets used a grid bias battery or "C" batteries which was connected to provide a *negative* voltage. Since virtually no current flows through a tube's grid connection, these batteries had very low drain and lasted the longest. Even after AC power supplies became commonplace, some radio sets continued to be built with C batteries, as they would almost never need replacing. However more modern circuits were designed using cathode biasing, eliminating the need for a third power supply voltage; this became practical with tubes using indirect heating of the cathode.

Note that the "C battery" is a designation having no relation to the 1.5 volt "C cell" (nor for the A and B batteries, discussed above).

AC power

Replacement of batteries was a major cost of operation for early radio receiver users. The development of the battery eliminator, and, in 1925, batteryless receivers operated by household power, reduced operating costs and contributed to the growing popularity of radio. A power supply using a transformer with several windings, one or more rectifiers (which may themselves be vacuum tubes), and large filter capacitors provided the required direct current voltages from the alternating current source.

As a cost reduction measure, especially in high-volume consumer receivers, all the tube heaters could be connected in series across the AC supply, and the plate voltage derived from a half-wave rectifier directly connected to the AC input, eliminating the need for a heavy power transformer. As an additional feature, these radios could be operated on AC or DC mains. This arrangement resulted in a limited plate voltage, however with advances in tube technology tubes could run reasonably effectively with only 150 volts on their plates. A filament tap on the rectifier tube provided the 6 volt, low current supply needed for a dial light.

Rectifying the AC mains directly did have one safety issue: the chassis of the receiver was connected to one side of the mains, presenting a shock hazard. This hazard was reduced by enclosing the chassis in an insulated case and running the AC power through a so-called interlock connection at the removable back side of the receiver. This would come disconnected whenever the radio was opened (for instance, to test and replace the tubes) preventing such a shock hazard. (Technicians and tinkerers routinely bypassed this by using a separate cord, known colloquially as a "cheater cord" or "widowmaker.") Many consumer AM radio manufacturers of the era used a virtually identical circuit with the tube complement of 12BA6, 12BE6, 12AV6, 35W4, and 50C5, giving these radios the nickname All American Five or simply "Five Tube Radio." Although millions of such receivers were produced, they have now become collector's items.

Reliability

Tube tester manufactured in 1930

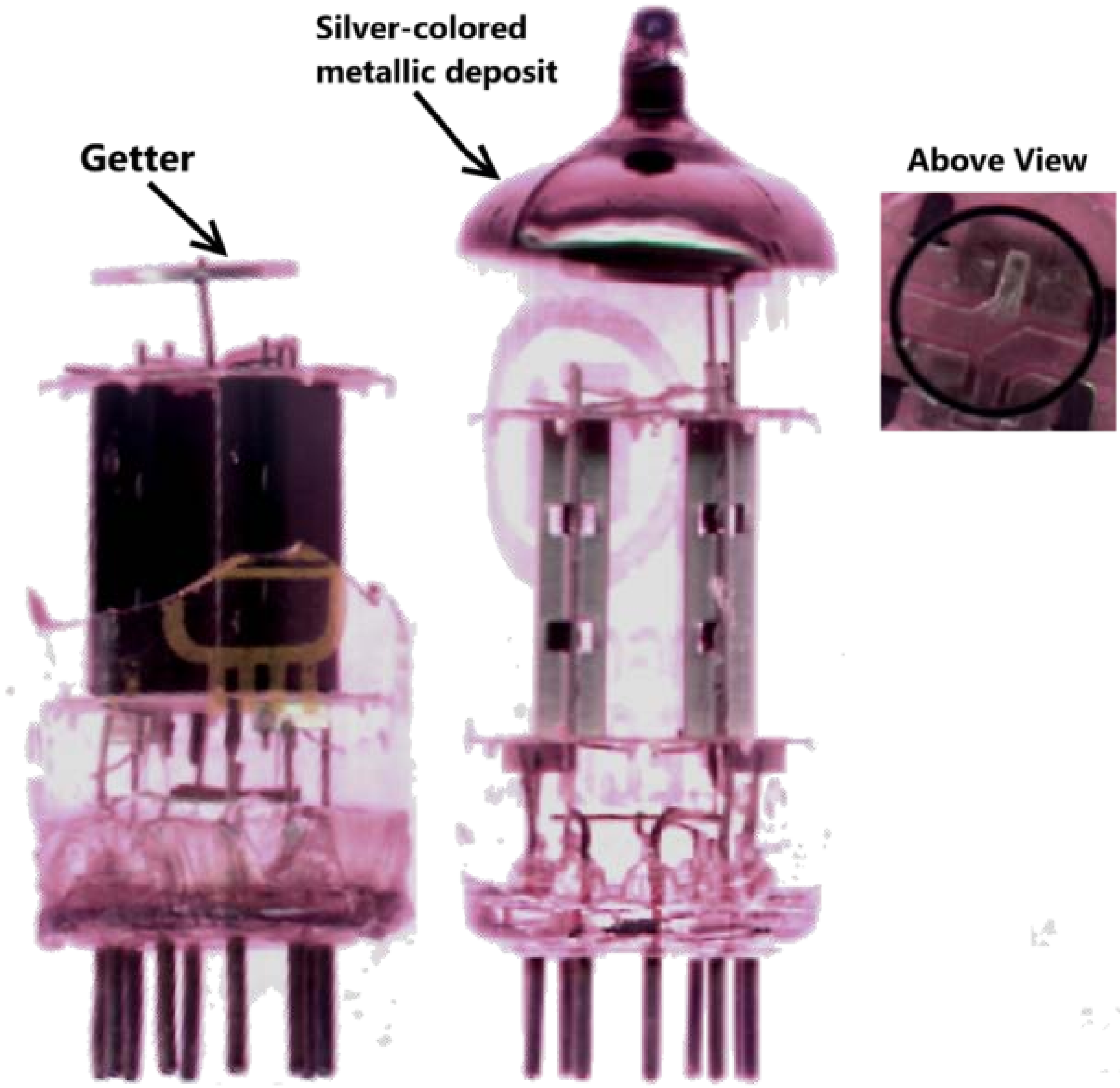
One reliability problem of tubes with oxide cathodes is the possibility that the cathode may slowly become "poisoned" by gas molecules from other elements in the tube, which reduce its ability to emit electrons. Trapped gases or slow gas leaks can also damage the cathode or cause plate-current run away due to ionization of free gas molecules. Vacuum hardness and proper selection of construction materials are the major influences on tube lifetime. Depending on the material, temperature and construction, the surface material of the cathode may also diffuse onto other elements. The resistive heaters that heat the cathodes may break in a manner similar to incandescent lamp filaments, but rarely do, since they operate at much lower temperatures than lamps.

The heater's failure mode is typically a stress-related fracture of the tungsten wire or at a weld point and generally occurs after accruing many thermal (power on-off) cycles. Tungsten wire has a very low resistance when at room temperature. A negative temperature coefficient device, such as a thermistor, may be incorporated in the equipment's heater supply or a ramp-up circuit may be employed to allow the heater or filaments to reach operating temperature more gradually than if powered-up in a step-function. Low-cost radios had tubes with heaters connected in series, with a total voltage equal to that of the line (mains). Following World War II, tubes intended to be used in series heater strings were redesigned to all have the same ("controlled") warm-up time. Earlier designs had quite-different thermal time constants. The audio output stage, for instance, had a larger cathode, and warmed up more slowly than lower-powered tubes. The result was that heaters that warmed up faster also temporarily had higher resistance, because of their positive temperature coefficient. This disproportionate resistance caused them to temporarily operate with heater voltages well above their ratings, and shortened their life.

Another important reliability problem is caused by air leakage into the tube. Usually oxygen in the air reacts chemically with the hot filament or cathode, quickly ruining it. Designers developed tube designs that sealed reliably. This was why most tubes were constructed of glass. Metal alloys (such as Cunife and Fernico) and glasses had been developed for light bulbs that expanded and contracted in similar amounts, as temperature changed. These made it easy to construct an insulating envelope of glass, while passing connection wires through the glass to the electrodes.

When a vacuum tube is overloaded or operated past its design dissipation, its anode (plate) may glow red. In consumer equipment, a glowing plate is universally a sign of an overloaded tube. However, some large transmitting tubes are designed to operate with their anodes at red, orange, or in rare cases, white heat.

Vacuum



Getter in opened tube; silvery deposit from getter

The highest possible vacuum is desired in a tube. Remaining gas atoms will ionize and conduct electricity between the elements in an undesired manner. In a defective tube residual air pressure will lead to ionization, becoming visible as a pink-purple glow discharge between the tube elements.

To prevent gases from compromising the tube's vacuum, modern tubes are constructed with "getters", which are usually small, circular troughs filled with metals that oxidize quickly, barium being the most common. While the tube envelope is being evacuated, the internal parts except the getter are heated by RF induction heating to help free any remaining gases from the metal parts. The tube is then sealed and the getter is heated to a high temperature, again by radio frequency induction heating. This causes some material from the getter to evaporate, reacting with any residual gases and usually leaving a silvercolored metallic deposit on the inside of the envelope of the tube. The getter continues to absorb small amounts of gas that may leak into the tube during its working life. If a tube develops a serious leak in the envelope, this deposit turns a white color as it reacts with atmospheric oxygen. Large transmitting and specialized tubes often use more exotic getter materials, such as zirconium. Early gettered tubes used phosphorus based getters and these tubes are easily identifiable, as the phosphorus leaves a characteristic orange or rainbow deposit on the glass. The use of phosphorus was short-lived and was quickly replaced by the superior barium getters. Unlike the barium getters, the phosphorus did not absorb any further gases once it had fired.

Transmitting tubes

Large transmitting tubes have carbonized tungsten filaments containing a small trace (1% to 2%) of thorium. An extremely thin (molecular) layer of thorium atoms forms on the outside of the wire's carbonized layer and, when heated, serve as an efficient source of electrons. The thorium slowly evaporates from the wire surface, while new thorium atoms diffuse to the surface to replace them. Such thoriated tungsten cathodes usually deliver lifetimes in the tens of thousands of hours. The end-of-life scenario for a thoriated-tungsten filament is when the carbonized layer has mostly been converted back into another form of tungsten carbide and emission begins to drop off rapidly; a complete loss of Thorium has never been found to be a factor in the end-of-life in a tube with this type of emitter. The highest reported tube life is held by an Eimac power tetrode used in a Los Angeles radio station's transmitter, which was removed from service after 80,000 hours (~9 years) of operation. It has been said that transmitters with vacuum tubes are better able to survive lightning strikes than transistor transmitters do. While it was commonly believed that at rf power levels above approx. 20 kilowatts, vacuum tubes were more efficient than solid state circuits, this is no longer the case especially in medium wave (AM broadcast) service where solid state transmitters at nearly all power levels have measurably higher efficiency. FM broadcast transmitters with solid state power amplifiers up to approx. 15 kW also show better overall mains-power efficiency than tube-based power amplifiers.

Receiving tubes

Cathodes in small "receiving" tubes are coated with a mixture of barium oxide and strontium oxide, sometimes with addition of calcium oxide or aluminium oxide. An electric heater is inserted into the cathode sleeve, and insulated from it electrically by a coating of aluminium oxide. This complex construction causes barium and strontium atoms to diffuse to the surface of the cathode when heated to about 780 degrees Celsius, thus emitting electrons.

Failure modes

Catastrophic failures

A catastrophic failure is one which suddenly makes the vacuum tube unusable. A crack in the glass envelope will allow air into the tube and destroy it. Cracks may result from stress in the glass, bent pins or impacts; tube sockets must allow for thermal expansion, to prevent stress in the glass at the pins. Stress may accumulate if a metal shield or other object presses on the tube envelope and causes differential heating of the glass. Glass may also be damaged by high-voltage arcing.

Tube heaters may also fail without warning, especially if exposed to over voltage or as a result of manufacturing defects. Tube heaters do not normally fail by evaporation like lamp filaments, since they operate at much lower temperature. The surge of inrush current when the heater is first energized causes stress in the heater, and can be avoided by slowly warming the heaters, gradually increasing current. Some tubes intended for series string operation of the heaters across the supply will have a definite controlled warm-up time to avoid excess voltage on some heaters as others warm up. Directly-heated filament-type cathodes as used in battery-operated tubes or some rectifiers may fail if the filament sags, causing internal arcing. Excess heater-to-cathode voltage in indirectly heated cathodes can break down the insulation between elements and destroy the heater.

Arcing between tube elements can destroy the tube. An arc can be caused by applying plate potential before the cathode has come up to operating temperature, or by drawing excess current through a rectifier which damages the emission coating. Arcs can also be initiated by any loose material inside the tube, or by excess screen voltage. An arc inside the tube allows gas to evolve from the tube materials, and may deposit conductive material on internal insulating spacers.

Degenerative failures

Degenerative failures cause the performance of the tube to slowly deteriorate with time.

Overheating of internal parts, such as control grids or mica spacer insulators, can result in trapped gas escaping into the tube; this can reduce performance. A getter is used to absorb gases evolved during tube operation, but has only a limited ability to combine with gas. Control of the envelope temperature prevents some types of gassing. A tube with very bad internal gas may have a visible blue glow when plate voltage is applied.

Gas and ions within the tube contribute to grid current which can disturb operation of a vacuum tube circuit. Another effect of overheating is the slow deposit of metallic vapors on internal spacers, resulting in inter-element leakage.

Tubes on standby for long periods, with heater voltage applied, may develop high cathode interface resistance and display poor emission characteristics. This effect occurred especially in pulse and digital circuits, where tubes had no plate current flowing for extended times.

Cathode depletion describes the loss of emission after thousands of hours of normal use. Sometimes emission can be restored for a time by raising heater voltage either for a short time or a permanent increase of a few percent. Cathode depletion was uncommon in signal tubes but was a frequent cause of failures of monochrome television cathode-ray tubes.

Other failures

Vacuum tubes may have or develop defects in operation that makes an individual tube useless in one device, but which may not prevent its satisfactory operation in another system. *Microphonics* refers to internal vibration of tube elements, which modulates the signal from the tube in an undesirable way; sound or vibration pick-up may affect the signals, or even cause uncontrolled howling if a feedback path develops between a microphonic tube and, for example, a loudspeaker. Leakage current between AC heaters and the cathode may couple into the circuit, or electrons emitted directly from the ends of the heater may also inject hum into the signal. Leakage current due to internal contamination may also inject noise.

Cooling

Like any electronic device, vacuum tubes produce heat while operating. This waste heat is one of the principal factors that affect tube life. In power amplifiers, the majority of this waste heat originates in the anode though screen grids may also require cooling. For example, the screen grid in an EL34 is cooled by two small radiators or "wings" near the top of the tube. A tube's heater (filament) also contributes to the total waste heat. A tube's data sheet will normally identify the maximum amount of heat that each element may safely dissipate.

The method of anode cooling is dependent on the construction of the tube itself. Tubes used in consumer equipment have internal anodes, so cooling occurs through black body radiation from the anode to the glass envelope. Natural convection (air circulation) then removes the heat from the envelope. Tube shields that aided heat dispersal could be retrofitted on certain select types of tubes. These shields act by improving heat conduction from the surface of the tube to the shield itself by means of tens of copper tongues in contact with the glass tube, and have an opaque, black outside finish for improved heat radiation. The ability to remove heat may be further increased by implementing forced air cooling, and adding an external heat sink attached to the anode through the tube's enclosure. These measures are both implemented in the 4-1000A transmitting tube whose anode was designed to safely operate at red hot temperatures, dissipating up to one kilowatt.

The amount of heat that may be removed from a tube with an internal anode is limited. Tubes with external anodes may be cooled using forced air, water, vapor, and multiphase.

The 3CX10,000A7 is an example of a tube with an external anode cooled by forced air. The water, vapor, and multiphase cooling techniques all depend on the high specific heat and latent heat of water. The 8974 is an example of a water cooled tube and is among the largest commercial tubes available today.

In a water cooled tube, the anode voltage appears directly on the cooling water surface, thus requiring the water to be an electrical insulator. Otherwise the high voltage can be conducted through the cooling water to the radiator system; hence the need for deionized water. Such systems usually have a built-in water-conductance monitor which will shut down the high-tension supply (often tens of kilovolts) if the conductance becomes too high.

Other vacuum tube devices

Many devices were built during the 1920–1960 period using vacuum-tube techniques. Most such tubes were rendered obsolete by semiconductors; some techniques for integrating multiple devices in a single module, sharing the same glass envelope have been discussed above, such as the Loewe 3NF. Vacuum-tube electronic devices still in common use include the magnetron, klystron, photomultiplier, x-ray tube, traveling-wave tube and cathode ray tube. The magnetron is the type of tube used in all microwave ovens. In spite of the advancing state of the art in power semiconductor technology, the vacuum tube still has reliability and cost advantages for high-frequency RF power generation. Photomultipliers are still the most sensitive detectors of light.

The cathode ray tube (CRT) is a vacuum tube used particularly for display purposes. Many televisions, oscilloscopes and computer monitors still use cathode ray tubes, though flat panel displays are becoming more predominant as prices drop. At one time many radios used "magic eye" tubes, a specialized sort of CRT used in place of a meter movement to indicate signal strength, or input level in a tape recorder. A modern indicator device, the vacuum fluorescent display (VFD) is also a sort of cathode ray tube.

Secondary emission is the term for what happens when electrons in a vacuum strike certain materials, and the impacts cause electrons to be emitted. For some materials, more electrons are emitted than originally hit the surface. Such devices, called electron multipliers, amplify the current represented by the incoming electrons. Several stages (as many as 15 or so) can be cascaded for high gain, and are essential parts of very sensitive phototubes, usually called photomultipliers or multiplier photoubes. The image orthicon TV studio camera tubes also used multistage photomultipliers.

For decades, electron-tube designers tried to use secondary emission to obtain more amplification in vacuum tubes with hot cathodes, but they suffered from short life because the material used for the secondary-emission electrode (called a dynode) "poisoned" the tube's hot cathode. (For instance, the interesting RCA 1630 secondary-emission tube was marketed, but did not last.) However, eventually, Philips of The Netherlands developed the EFP60 tube that had a satisfactory lifetime, and was used in at least one product, a laboratory pulse generator. However, transistors were rapidly improving, and eclipsed tubes in general.

A variant, called a channel electron multiplier, is a curved tube, such as a helix, coated on the inside with material with good secondary emission. One type had a little funnel to capture incoming electrons. The tube was resistive, and its ends were connected to enough voltage to create repeated cascades of electrons.

Tektronix made a high-performance wideband oscilloscope CRT with a channel electron multiplier plate behind the phosphor layer. This plate was a bundled array of a huge number of short individual c.e.m. tubes that accepted a low-current beam and intensified it to provide a display of practical brightness. (The electron optics of the wideband electron gun could not provide enough current to directly excite the phosphor.)

Some tubes, such as magnetrons, traveling-wave tubes, carcinotrons, and klystrons, combine magnetic and electrostatic effects. These are efficient (usually narrow-band) RF producers and still find use in radar, microwave ovens and industrial heating. Traveling-wave tubes (TWTs) are very good amplifiers; they are used in some communications satellites. High-powered klystron amplifier tubes can provide hundreds of kilowatts in the UHF range.

Gyrotrons or vacuum masers, used to generate high-power millimetre band waves, are magnetic vacuum tubes in which a small relativistic effect, due to the high voltage, is used for bunching the electrons. Gyrotrons can generate very high powers (hundreds of kilowatts). Free electron lasers, used to generate high-power coherent light and perhaps even X rays, are highly relativistic vacuum tubes driven by high-energy particle accelerators.

Particle accelerators can be considered vacuum tubes that work backward, the electric fields driving the electrons, or other charged particles. In this respect, a cathode ray tube is a particle accelerator.

A tube in which electrons move through a vacuum (or gaseous medium) within a gas-tight envelope is generically called an *electron tube*.

Some condenser microphone designs use built-in vacuum tube preamplifiers.

Vacuum tubes in the 21st century

Niche applications

Although vacuum tubes have been largely replaced by solid-state devices in most amplifying applications, there are certain exceptions. In addition to the special functions noted above, tubes have some niche applications even in the current age.

Vacuum tubes are much less susceptible than corresponding solid-state components to the electromagnetic pulse effect of nuclear explosions. This property kept them in use for certain military applications long after more practical and less expensive solid state technology had been perfected for the same applications.

Vacuum tubes are still practical alternatives to solid state in generating high power radio frequencies in applications such as industrial radio frequency heating, particle accelerators, and radio broadcast transmitters. This is particularly true at microwave frequencies where such devices as the klystron and traveling-wave tube provide amplification at power levels unattainable using current semiconductor devices. The household microwave oven uses a magnetron tube to efficiently generate microwave powers of several hundred watts.

70 watt tube audio amplifier selling for $2,680 in 2011. Certain audiophiles happily pay ten times more for the perceived "tube sound."

Many audiophiles, professional audio engineers, and musicians prefer the so-called tube sound compared with amplifiers using transistors. The power output stages of audio amplifiers using tubes include transformers to match the speaker impedance to the higher impedance level of the tube circuit whereas solid state power amplifiers are direct coupled and rely on a high degree of negative feedback. The output transformer will affect the amplifier's tone (frequency response) in response to the speaker's impedance, and will affect the character of the amplifier's distortion as it approaches maximum power. There are companies which specialize in high end audio amplifiers using tube technology to serve this market. Beyond the amplifier's output stage, more controversial claims are made in favor of tubes used in signal amplification stages and even for using tubes as power supply rectifiers. Professional music recording studios and public address systems sometimes employ microphone preamplifiers using tubes based on this perceived superiority.

Tube based electric guitar amplifiers are highly regarded. These are widespread and still in production. The sound produced by a tube power amplifier when overdriven has defined the texture of some genres of music such as classic rock and blues. Rather than the hard clipping characteristic of solid state power amplifiers, a tube amplifier in conjunction with its output transformer produces a more gradual distortion and distinctive sound. Guitarists often cite the sound of tube amplifiers for the "warmth" of their tone and the natural compression that results when overdriven (as guitar amplifiers routinely are). Although the reliability of solid state amplifiers has greatly improved, tube amplifiers have the advantage that the output tubes can be replaced by the owner, whereas "blown" output transistors require attention by a qualified technician.

Vacuum fluorescent display

Typical VFD used in a videocassette recorder

A modern display technology using a sort of cathode ray tube is freqeuntly used in videocassette recorders, microwave oven control panels, and automotive dashboards. Rather than raster scanning these vacuum fluorescent displays (VFD) switch control grids and plate voltages on and off to display discrete characters, for instance. This term should not be confused with fluorescent light technology which uses fluorescence from ultraviolet radiation produced using a discharge tube. Rather the VFD uses phosphor-coated anodes as in other display cathode ray tube. Because the filaments are in view, they must be operated at temperatures where the filament does not glow visibly. This is possible using more recent cathode technology, and these tubes also operate with quite low anode voltages (often less than 50 volts) contrary to classical cathode ray tubes. Often found in automotive applications, their high brightness allows reading the display in daylight. VFD tubes are flat and rectangular, as well as relatively thin. Typical VFD phosphors emit a broad-spectrum greenish-white light, permitting use of color filters. This type of phosphor provides a bright glow despite the low energy of the incident electrons.

Vacuum tubes using field electron emitters

In the early years of the 21st century there has been renewed interest in vacuum tubes, this time with the electron emitter formed on a flat silicon substrate, as in integrated circuit technology. This subject is now called vacuum nanoelectronics. The most common design uses a cold cathode in the form of a large-area field electron source (for example a field emitter array). With these devices, electrons are field-emitted from a large number of closely spaced individual emission sites.

Their claimed advantages include greatly enhanced robustness combined with the ability to provide high power outputs at low power consumptions. Operating on the same principles as traditional tubes, prototype device cathodes have been fabricated in several different ways. Although a common approach is to use a field emitter array, one interesting idea is to etch electrodes to form hinged flaps – similar to the technology used to create the microscopic mirrors used in Digital Light Processing) that are stood upright by an electrostatic charge.

Such integrated microtubes may find application in microwave devices including mobile phones, for Bluetooth and Wi-Fi transmission, in radar and for satellite communication. Presently they are being studied for possible applications in field emission display technology, but significant production problems seem to exist.

Modern manufacturers

Vacuum tubes are still being manufactured in the following countries:

China

Manufacturer	Area of expertise
Shuguang Electron Group Co.	Tubes primarily for audio applications.
Tianjin Quanerzhen Electron Tube Technology Co.	Direct heated tubes for audio applications.
Nanjing Sanle Electronic Information	Transmitting and industrial tubes. (Including Chinese 3-500C & 4-400C)

Industry Group Co.	
JiangXi Jingguang Electronics Co.	Transmitting and industrial tubes.
Huaguang Electric Power & Electronics Co.	Transmitting and industrial tubes
Chengdu Xuguang Electronics Co.	Transmitting and industrial tubes

Russia

Manufacturer	Area of expertise
Ekspopul JSC	Audio tube factory of New Sensor Inc. Known formerly as tube factory of Reflektor JSC.
"Ryazan" Vacuum Components LLC	Direct heated tubes. SV811 and SV572 series for audio, 811A and 572B for RF applications.
"SED-SPb" Svetlana Electron Devices - St. Petersburg. Svetlana JSC	Primarily transmitting and industrial tubes. Manufactures also few models for audio applications.
Voskhod KRLZ JSC	Tubes for small signal RF and audio applications
NEVZ-Soyuz HC JSC	Transmitting and industrial tubes, known formerly as Novosibirsk electron tube plant.

United States

Manufacturer	Area of expertise
Communications & Power Industries Inc.	Transmitting and industrial tubes, formerly known as Eitel-McCullough Inc.
Burle Industries Inc.	Transmitting and industrial tubes, formerly factory of RCA
MPD Components Inc.	Planar triodes and magnetrons, formerly Ken-Rad and later GE tube factory
MU Incorporated	Contract manufacturer.
Western Electric Export Corporation	300B Tubes Former Factory of AT&T's Western Electric moved equipment from Missouri to Tennessee in 2002

United Kingdom

Manufacturer	Area of expertise
e2v Technologies Ltd.	Transmitting and industrial tubes, formerly known as English Electric Valve Co. Ltd.

TMD Technologies Ltd.	Transmitting and industrial tubes. Formerly THORN Microwave Devices Ltd.

France

Manufacturer	Area of expertise
Covimag SA	Transmitting and industrial tubes. Products marketed by Richardson Electronics. Formerly Philips transmitting tube factory.
Thales Electron Devices SA	Transmitting and industrial tubes. Formerly known as Thomson-CSF.

Czech Republic

Manufacturer	Area of expertise
Emission Labs	Direct heated tubes for audio applications
KR Audio Electronics s.r.o.	Direct heated tubes for audio applications
Tesla Electrontubes s.r.o	Transmitting and industrial tubes

Slovakia

Manufacturer	Area of expertise
JJ-Electronic	Tubes primarily for audio applications, factory was formerly part of Tesla Electrontubes
Euro Audio Team	Tubes for high-end audio.

Chapter 10

Helium Mass Spectrometer

A **helium mass spectrometer** is an instrument commonly used to detect and locate small leaks. It was initially developed in the Manhattan Project during World War II to find extremely small leaks in the gas diffusion process of uranium enrichment plants. It typically uses a vacuum chamber in which a sealed container filled with helium is placed. Helium leaks out of the container, and the rate of the leak is detected by a mass spectrometer.

Detection technique

The leak detection technique depends on the separation of helium from other gases in a vacuum. It is accomplished by ionizing a gas sample containing helium, pushing the sample through a magnetic field, and collecting the helium ions as they emerge. Since helium ions exit along a different path from all other ions, collection of helium is reasonably simple. The current produced by the helium ion flow is used to drive an ammeter. Often audio alarms, and visual display on leak detection system to give information and warnings about leak levels. Ionization, separation and collection takes place within a spectrometer tube, which is the heart of the system.

Helium is used as a tracer because it penetrates small leaks rapidly. Helium has also the property of being non-toxic, chemically inert, inexpensive to produce, and present in the atmosphere only in minute quantities (5 ppm). Typically a helium leak detector will be used to measure leaks in the range of 10^{-5} to 10^{-12} $Pa{\cdot}m^3{\cdot}s^{-1}$.

- A flow of 10^{-5} $Pa{\cdot}m^3{\cdot}s^{-1}$ is slightly less than 1 ml per minute at Standard conditions for temperature and pressure (STP).
- A flow of 10^{-13} $Pa{\cdot}m^3{\cdot}s^{-1}$ is slightly less than 3 ml per century at STP.

Internal spectrometer tube operation

In the spectrometer tube, the heart of the helium mass spectrometer, the electrons produced by a hot filament enter an ionization chamber under vacuum, and collide with gas molecules, creating within the chamber ions quantitatively proportional to the pressure in the ion chamber. These ions are repelled out of the ion chamber, under vacuum, through the exit slit, by a repeller field. The combined electrostatic effect of the repeller, exit slit, focus plates, and ground slit collimates the ion beam so that it enters the magnetic field as a straight "ribbon" of ions.

Types of leaks

Typically there are two types of leaks in the detection of helium as a *tracer* for leak detection.

- **Residual leak**: A residual leak is a real leak that may be gross, or small, according to the sensitivity setting of the leak detector.
- **Virtual leak**: A virtual leak is the semblance of a leak in a vacuum system caused by slow release of trapped gases, as gases may adhere as pockets to the interior sides of a chamber. This can cause confusion to the operator as it may be a false indication of a present leak.

Uses

Helium mass spectrometer leak detectors are used in production line industries such as refrigeration and air conditioning, automotive parts, carbonated beverage containers, food packages and aerosol packaging, as well as in the manufacture of steam products, gas bottles, fire extinguishers, tire valves, and numerous other products including all vacuum systems.

Test methods

Global helium spray - vacuum test

This method requires the part to be tested to be connected to a helium leak detector. The outer surface of the part to be tested will be located in some kind of a tent in which the helium concentration will be raised to 100% helium.

- If the part is small the vacuum system included in the leak testing instrument will be able to reach low enough pressure to allow for mass spectrometer operation.
- If the size of the part is too large, an additional vacuum pumping system may be required to reach low enough pressure in a reasonable length of time. Once operating pressure has been reached, the mass spectrometer can start its measuring operation.

If leakage is encountered the small and "agile" molecules of helium will migrate through the cracks into the part. The vacuum system will carry any tracer gas molecule into the analyzer cell of the magnetic sector mass spectrometer. A signal will inform the operator of the value of the leakage encountered.

Local helium spray - vacuum test

This method is a small variation from the one above. It still requires the part to be tested to be connected to a helium leak detector. The outer surface of the part to be tested is sprayed with a localized stream of helium tracer gas.

- If the part is small the vacuum system included in the instrument will be able to reach low enough pressure to allow for mass spectrometer operation.
- If the size of the part is too large, an addition pumping system may be required to reach low enough pressure in a reasonable length of time. Once operating pressure has been reached, the mass spectrometer can start its measuring operation.

If leakage is encountered the small and "agile" molecules of helium will migrate through the cracks into the part. The vacuum system will carry any tracer gas molecule into the analyzer cell of the magnetic sector mass spectrometer. A signal will inform the operator of the value of the leakage encountered. Thus correlation between maximum leakage signal and location of helium spray head will allow the operator to pinpoint the leaky area.

Helium charged - vacuum test

In this case the part is pressurized (sometime this test is combined with a burst test, i.e. at 40 bars) with helium while sitting in a vacuum chamber. The vacuum chamber is connected to a vacuum pumping system and a leak detector. Once the vacuum has reached the mass spectrometer operating pressure, any helium leakage will be measured. This test method applies to a lot of components that will operate under pressure: airbag canisters, evaporators, condensers, high-voltage SF_6 filled switchgear.

Bombing test (overpressure) - vacuum test

This method applies to objects that are supposedly sealed.

First the device under test will be exposed for an extended length of time to a high helium pressure in a "bombing" chamber.

If the part is leaky, helium will be able to penetrate the device.

Later the device will be placed in a vacuum chamber, connected to a vacuum pump and a mass spectrometer. The tiny amount of gas that entered the device under pressure will be released in the vacuum chamber and sent to the mass spectrometer where the leak rate will be measured.

This test method applies to implantable medical devices, crystal oscillator, saw filter devices.

This method is not able to detect a massive leak as the tracer gas will be quickly pumped out when test chamber is pumped down.

Helium charged - sniffer test

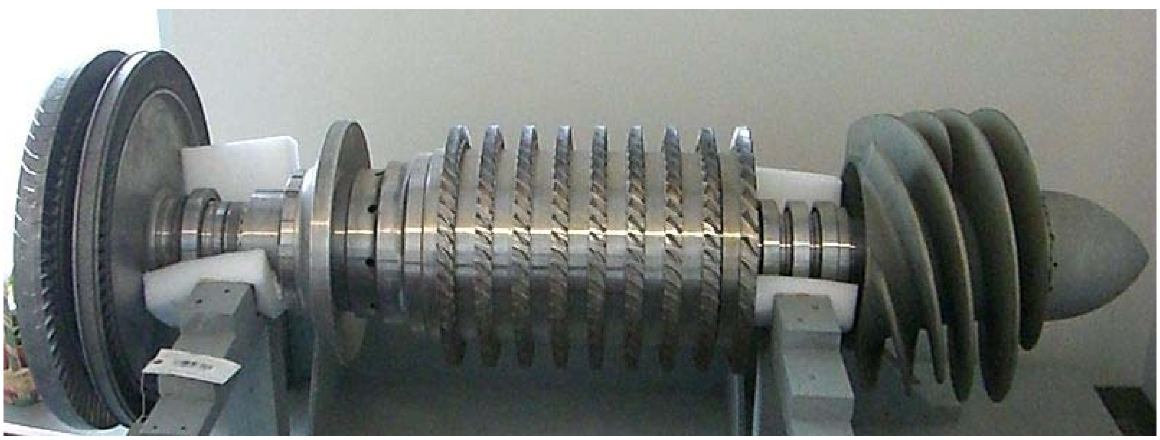
In this last case the part is pressurized with helium. The mass spectrometer is fitted with a special device, a sniffer probe, that allows it to sample air (and tracer gas when confronted with a leak) at atmospheric pressure and to bring it into the mass spectrometer.

This mode of operation is frequently used to locate a leak that has been detected by other methods, in order to allow for parts repair.

Chapter 11

Turbopump

A **turbopump** is a gas turbine that comprises basically two main components: a rotodynamic pump and a driving turbine, usually both mounted on the same shaft, or sometimes geared together. The purpose of a turbopump is to produce a high pressure fluid for feeding a combustion chamber or other use.



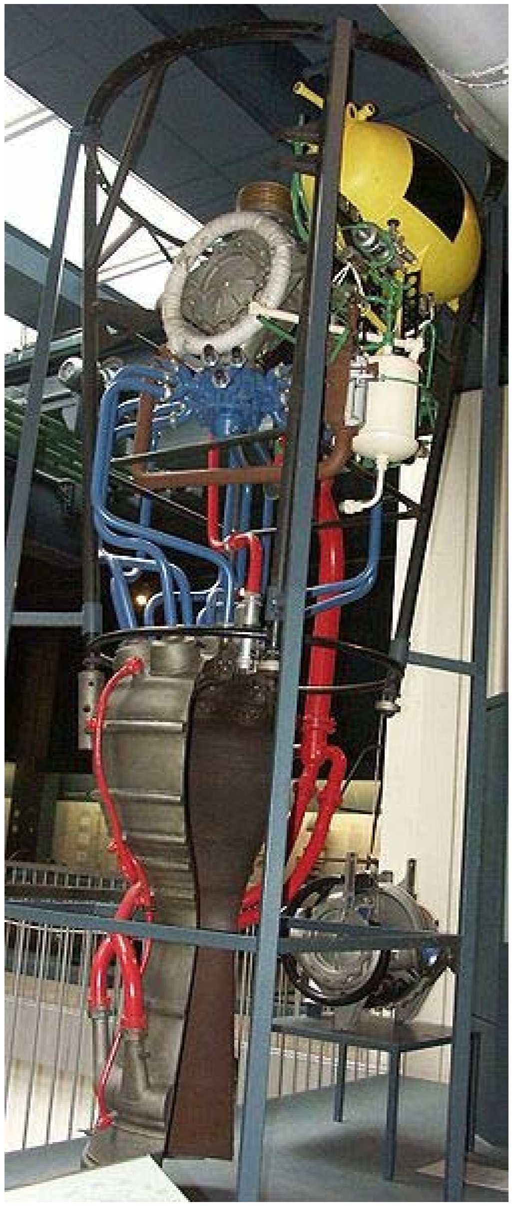
An axial turbopump designed and built for the M-1 rocket engine

A turbopump can comprise one of two types of pumps: centrifugal pump, where the pumping is done by throwing fluid outward at high speed; or axial flow pump, where alternating rotating and static blades progressively raise the pressure of a fluid.

Axial flow pumps have small diameters, but give relatively modest pressure increases, and multiple compression stages are needed, but work well with low density fluids. Centrifugal pumps are far more powerful for high density fluids, but require physically large diameters for low density fluids.

Turbopumps operate in much the same way as turbo units for vehicles. Higher fuel pressures allow fuel to be supplied to higher-pressure combustion chambers for higher performance engines.

History



The V-2 rocket used a circular turbopump to pressurize the propellant.

Early development

High-pressure pumps for larger missiles had been discussed by rocket pioneers such as Hermann Oberth. In mid-1935 Wernher von Braun initiated a fuel pump project at the southwest German firm *Klein, Schanzlin & Becker* that was experienced in building large fire-fighting pumps.[:80] The V-2 rocket design used hydrogen peroxide decomposed through a Walther steam generator to power the uncontrolled turbopump[:81] produced at the Heinkel plant at Jenbach, so V-2 turbopumps and combustion chamber were tested and matched to prevent the pump from overpressurizing the chamber.[:172] The first engine fired successfully in September, and on August 16, 1942, a trial rocket stopped in mid-air and crashed due to a failure in the turbopump. The first successful V-2 launch was on October 3, 1942.

Development from 1947 to 1949

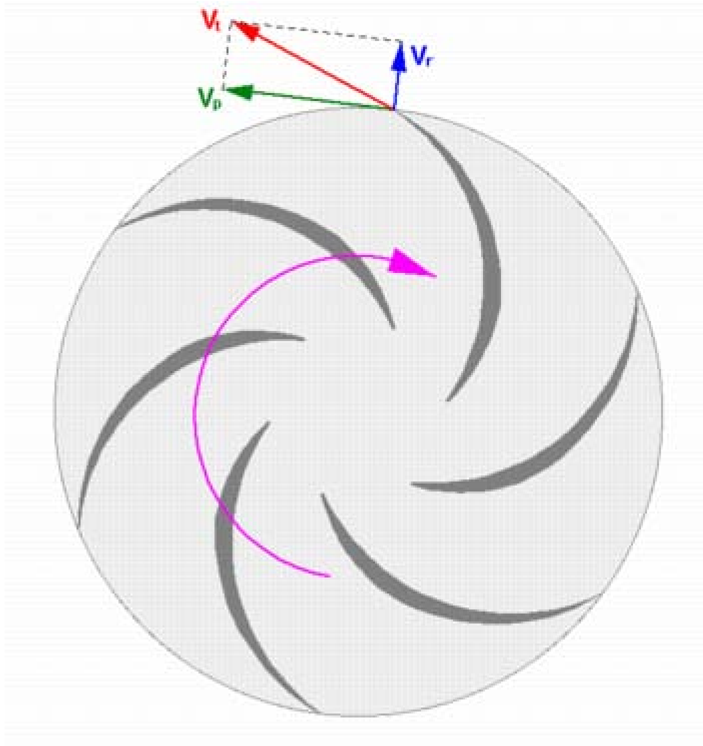
The principal engineer for turbopump development at Aerojet was George Bosco. During the second half of 1947, Bosco and his group learned about the pump work of others and made preliminary design studies. Aerojet representatives visited Ohio State University where Florant was working on hydrogen pumps, and consulted Dietrich Singelmann, a German pump expert at Wright Field. Bosco subsequently used Singelmann's data in designing Aerojet's first hydrogen pump.

By mid-1948, Aerojet had selected centrifugal pumps for both liquid hydrogen and liquid oxygen. They obtained some German radial-vane pumps from the Navy and tested them during the second half of the year.

By the end of 1948, Aerojet had designed, built, and tested a liquid hydrogen pump (15 cm diameter). Initially, it used ball bearings that were run clean and dry, because the low temperature made conventional lubrication impractical. The pump was first operated at low speeds to allow its parts to cool down to operating temperature. When temperature gauges showed that liquid hydrogen had reached the pump, an attempt was made to accelerate from 5000 to 35 000 revolutions per minute. The pump failed and examination of the pieces pointed to a failure of the bearing, as well as the impeller. After some testing, super-precision bearings, lubricated by oil that was atomized and directed by a stream of gaseous nitrogen, were used. On the next run, the bearings worked satisfactorily but the stresses were too great for the brazed impeller and it flew apart. A new one was made by milling from a solid block of aluminum. Time was running out, as the contract had less than six months to go. The next two runs with the new pump were a great disappointment; the instruments showed no significant flow or pressure rise. The problem was traced to the exit diffuser of the pump, which was too small and insufficiently cooled during the cool-down cycle so that it limited the flow. This was corrected by adding vent holes in the pump housing; the vents were opened during cool down and closed when the pump was cold. With this fix, two additional runs were made in March 1949 and both were successful. Flow rate and pressure were found to be in approximate agreement with theoretical predictions. The maximum pressure was 26 atmospheres and the flow was 0.25 kilogram per second.

Today the Space Shuttle Main Engine's turbopumps spin at over 30,000 rpm, delivering 150 lb (38 kg) of liquid hydrogen and 896 lb (406 kg) of liquid oxygen to the engine per second.

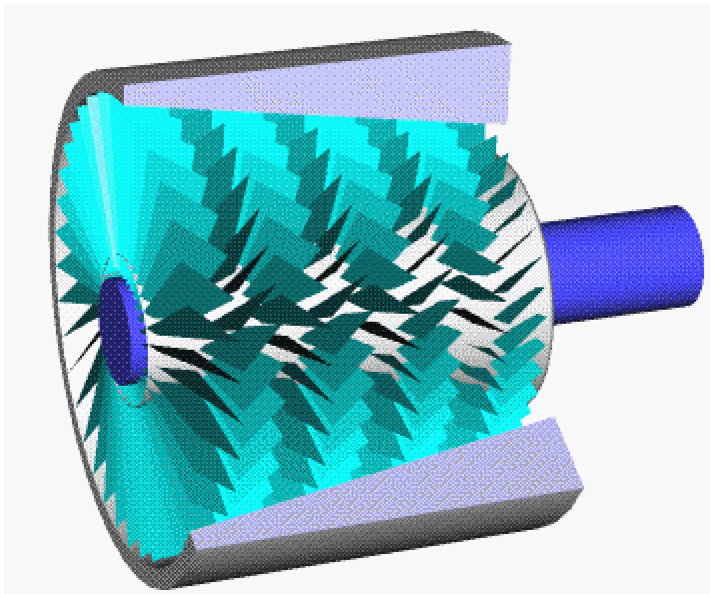
Centrifugal turbopumps



In centrifugal turbopumps a rotating disk throws the fluid to the rim

Most turbopumps are centrifugal - the fluid enters the pump near the axis and the rotor accelerates the fluid to high speed. The fluid then passes through a diffuser which is a progressively enlarging pipe, which permits recovery of the dynamic pressure. The diffuser turns the high kinetic energy into high pressures (hundreds of bar is not uncommon), and if the outlet backpressure is not too high, high flow rates can be achieved.

Axial turbopumps



Axial compressors

Axial turbopumps also exist - in this case the axle essentially has propellers attached to the shaft and the fluid is forced by these parallel with the main axis of the pump. Generally, axial pumps tend to give much lower pressures than centrifugal pumps, and a few bar is not uncommon. They are however still useful - axial pumps are commonly used as 'inducers' for centrifugal pumps, which raise the inlet pressure of the centrifugal pump enough to prevent excessive cavitation from occurring therein.

Complexities of centrifugal turbopumps

Turbopumps have a reputation for being extremely hard to design to get optimum performance. Whereas a well-engineered and debugged pump can manage 70-90% efficiency, figures less than half that are not uncommon. Low efficiency may be acceptable in some applications, but in rocketry this is a severe problem. Turbopumps in rockets are important and problematic enough that launch vehicles using one have been caustically described as a 'turbopump with a rocket attached'- up to 55% of the total cost has been ascribed to this area.

Common problems include:

1. excessive flow from the high pressure rim back to the low pressure inlet along the gap between the casing of the pump and the rotor
2. excessive recirculation of the fluid at inlet
3. excessive vortexing of the fluid as it leaves the casing of the pump
4. damaging cavitation to impeller blade surfaces in low (fluid) pressure zones

In addition, the precise shape of the rotor itself is critical.

Driving Turbopumps

Steam turbine powered turbopumps do exist and are employed when there is a source of steam, e.g. the boilers of steam ships. Now gas turbines are usually used when electricity or steam is not available and place or weight restrictions permit the use of more-efficient sources of mechanical energy.

One of such cases are rocket engines which need to pump fuel and oxidizer into their combustion chamber. This is necessary for large liquid rockets, since forcing the fluids or gases to flow by simple pressurizing of the tanks is often not feasible: The high pressure needed for the required flow rates would need strong and heavy tanks.

Ramjet motors are also usually fitted with turbopumps, the turbine being driven either directly by external freestream ram air or internally by airflow diverted from combustor entry. In both cases the turbine exhaust stream is dumped overboard.