Mechatronics Engineering

Monty Holden

First Edition, 2012

ISBN 978-81-323-3083-7

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

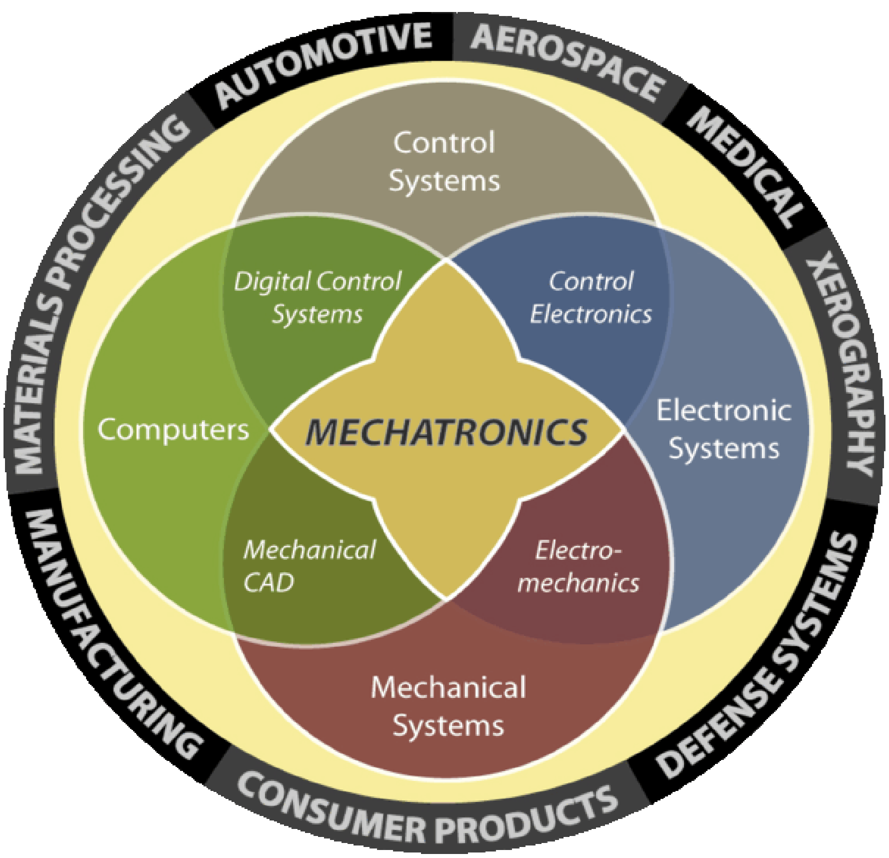
Table of Contents

Introduction

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

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

Description



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

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

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

Science Fiction

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

Course structure

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

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

Application

- Machine vision

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

Physical implementations

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

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

Variant of the field

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

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

Education

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

Chapter 1

Automation

KUKA Industrial Robots being used at a bakery for food production

Automation is the use of control systems and information technologies to reduce the need for human work in the production of goods and services. In the scope of industrialization, automation is a step beyond mechanization. Whereas mechanization provided human operators with machinery to assist them with the muscular requirements of work, automation greatly decreases the need for human sensory and mental requirements as well. Automation plays an increasingly important role in the world economy and in daily experience.

Automation has had a notable impact in a wide range of industries beyond manufacturing (where it began). Once-ubiquitous telephone operators have been replaced largely by automated telephone switchboards and answering machines. Medical processes such as primary screening in electrocardiography or radiography and laboratory analysis of human genes, sera, cells, and tissues are carried out at much greater speed and accuracy by automated systems. Automated teller machines have reduced the need for bank visits to obtain cash and carry out transactions. In general, automation has been responsible for the shift in the world economy from industrial jobs to service jobs in the 20th and 21st centuries.

Advantages and disadvantages

The main advantages of automation are:

- Replacing human operators in tasks that involve hard physical or monotonous work.
- Replacing humans in tasks done in dangerous environments (i.e. fire, space, volcanoes, nuclear facilities, underwater, etc.)
- Performing tasks that are beyond human capabilities of size, weight, speed, endurance, etc.
- Economy improvement. Automation may improve in economy of enterprises, society or most of humanity. For example, when an enterprise invests in automation, technology recovers its investment; or when a state or country increases its income due to automation like Germany or Japan in the 20th Century.

The main disadvantages of automation are:

- Technology limits. Current technology is unable to automate all the desired tasks.
- Unpredictable development costs. The research and development cost of automating a process may exceed the cost saved by the automation itself.
- High initial cost. The automation of a new product or plant requires a huge initial investment in comparison with the unit cost of the product, although the cost of automation is spread in many product batches.

Relationship to unemployment

Multivariate effect

Most people consider it common sense that automation has the potential to foster unemployment, because it obviates human work by transferring tasks to machines. However, the translation of that potential into observed effect has largely not happened in the two centuries during which it has been continually predicted. After many decades of automation development and dissemination, the net macroeconomic effect has been generally positive—automation has been part of a general trend of economic growth worldwide; standards of living have risen in many places; and automation has never yet been shown to have induced any widespread structural unemployment. The main explanation for this is that, so far, job losses in any one particular economic niche have always been more than offset by job gains in other niches. As the lowered unit cost of goods and services (which the automation made possible) gave consumers more purchasing power to devote to other goods and services, new jobs sprang up in the production of those goods and services. Thus each time that automation has freed up human resources, those resources have been redeployed by market forces (although it did not always happen without turbulence in the lives of individual workers).

One of the earliest promises of automation was to allow more free time, without any threat of income reduction. This effect has been seen in many individual facets of life (for example, the automatic washing machine has made laundry less time-consuming; engine control units have reduced the amount of automotive downtime; the automatic dishwasher has made dishwashing less time-consuming), but the net outcome of modern life in developed economies remains a state of hurry and busyness, mostly because rising living standards have brought rising expectations in direct relation. (Each time-saving improvement has made room for a new aspiration to take its place.)

Automation also does not imply unemployment when it makes possible tasks that were unimaginable without it (such as exploring Mars with the Sojourner rover). Likewise with fields where the economy is already fully adapted to an automated technology, and the jobs were lost long enough ago that the displacement was long since absorbed by the workforce (as with the continually advancing automation of the telephone switchboard, which eliminated most telephone operator jobs and kept many more from ever existing in the first place).

Today automation is quite advanced (relative to just a few lifetimes ago), and it continues to advance with an accelerating pace throughout the world. Although it has been encroaching on ever more skilled jobs, the general well-being and quality of life of most people in the world (where political factors have not muddied the picture) have improved. Clearly a multivariate effect has been at work (something much more than just the obvious idea that automation has the *potential* to cause unemployment). In fact, the idea that automation posed an *imminent* threat to employment, first articulated in 1811 by a group of textile workers known as Luddites, has proven to be so fallacious over the ensuing two centuries that economists call the imminent-threat idea the *Luddite fallacy*.

There is some concern today that the economy's ability to continue absorbing ever-increasing automation without experiencing significant structural unemployment may be heading toward an upper limit—that is, that we are approaching a point where the Luddite premise will no longer be entirely fallacious, because the relationship of humans to machines that made it fallacious is changing. In this view, the empirical strength of the eternal-fallaciousness idea is only a reflection of the parameter values of the environment thus far. In other words, the idea is undoubtedly an excellent explanation of the past, but whether it can accurately predict the future is an independent problem. Like an investment prospectus, proponents of this view caution that "past performance is no guarantee of future results."

Timeline of concerns about automation's relationship to unemployment

Early in the Industrial Revolution

Historical concerns about the effects of automation date back to the very beginning of the Industrial Revolution, when a social movement of English textile machine operators in the early 19th century known as the Luddites protested against Jacquard's automated weaving looms. The Luddites destroyed a number of these machines, which they felt threatened their jobs.

Later in the Industrial Revolution

The development of the American system of manufacturing disgusted many skilled machinists at a time when the very definition of being a machinist included a core element of skilled toolmaking and fitting on a craft basis. Innovations of this system included increasing reliance on jigs and gauges and on machine tools that built more of a process into the tool's movements (such as turret lathes and screw machines). These innovations continually turned skilled work into semi-skilled or unskilled, contributing to vast migrations of laborers across borders and oceans. However, despite this transformation, there were always other economic niches for skilled workers to go to, given enough searching. Recessions interfered with employment, but no foundational aspects of structural unemployment were caused by automation itself.

During the Machine Age

As in the preceding century, the period of 1880 to 1940 saw no underlying automation-induced structural lack of new economic opportunities for skilled workers to go to, given enough searching, although the Great Depression caused a tremendous disruption to employment. The foundational *potential* for full employment had not been lost, as would later be shown by the post–World War II economic expansion and other economic miracles.

During the 1950s through 1990s

The postwar development of new automation technologies using electronics, servomechanisms, and digital computers stoked a new wave of fears similar to the old Luddite ones. Among the working class and labor unions, there was stiff resistance to loss of employment through automation, including contract clauses won in hard-fought contract negotiations that mandated alternate employment for any workers whose positions were eliminated by automation. These clauses seemed a great victory for union workers at large corporations in developed nations, but because they had no effect at smaller, nonunionized companies or in developing nations, those corporations faced withering competition that shrank their market shares until their workers' gains eventually undermined their own success. However, the salvation for employment rates damaged in the industrial sector (secondary sector of the economy) came from the service sector (tertiary sector), which absorbed all of the workers that automation displaced elsewhere.

For example, many manufacturing jobs left the United States during the 1990s but were offset by a one-time massive increase in IT jobs at the same time. And in some cases the freeing up of the labor force allowed more people to enter higher skilled managerial jobs and technically specialized jobs, which are typically higher paying. Therefore, fears of unemployment due to automation were generally dismissed as just another instance of the Luddite premise, which had proven fallacious time and again over many decades. Given this obvious empirical contradiction of the premise, people who nevertheless returned to it were usually viewed by the mainstream as cranks misled by quixotic leftist political bias. For example, works by scholars including David F. Noble and Jeremy Rifkin were often respected but discounted. At worst, they were mocked with the disparaging label "neo-Luddite". Noble even wrote a later book titled *Progress Without People: In Defence of Luddism* to try to further explain why the Luddite premise should not be laughed out of academia.

Post-market musings

Rifkin's *End of Work*, published in 1995 and written by a non-engineer, predicted automation-induced unemployment despite having a rather hazy idea of how IT would evolve over the next decade. (The book mentioned the Internet once in passing and the World Wide Web not at all. Its IT focus was mostly on robotics.) Also hazy was Rifkin's explanation of any solution to the problem. The book's subtitle called the solution a "post-market economy", but its concluding chapters did not clearly lay out how such an economy could be engineered, leaving readers to conclude that a non-market solution involving a planned economy was implied between the lines.

In terms of political economy implications, there was no clear differentiation at the time between the ideas of authors like Noble or Rifkin (on the one hand) and traditional leftist agitation (on the other hand). To the extent that readers could ask "What point is this guy getting to?" and answer the question with "socialism" or "a welfare state", they dismissed these authors.

During the 2000s and 2010s

Since the 1990s, the possibility has been raised again in even an apolitical, technocratic way that the Luddite premise (that automation creates unemployment) was only fallacious in the absence of highly advanced and ubiquitous automation, which until recently was mostly out of reach technologically. This would explain why it has always been fallacious until now, but also why it might not always remain so. For example, Marshall Brain, Martin Ford, and others have suggested that exponentially accelerating information technology (IT) may ultimately result in widespread structural unemployment, because an implicit assumption underlying the "eternally fallacious" idea (that lots of regular humans will always find ways to do service work that machines can't do) will itself be fallacious as IT advances. They suggest that, unlike in the 20th century, when the tertiary sector absorbed all of the workers that the automation of the secondary sector expelled, the tertiary sector now also faces depopulation via automation; its employment will shrink, not grow, and this time there is no other sector to backstop the process by absorbing the displaced workers. The high unemployment rates of the late-

2000s recession have brought the idea of structural unemployment back into mainstream attention, as observations are made about positions that require extensive specialized skill and experience standing long vacant even while general unemployment rates above 9% (and horror stories of fruitless job searches) would seem to suggest that such vacancies ought to be scarcer. The idea that automation has finally advanced to the point that the Luddite premise is no longer entirely fallacious is one of the components of some theoretical explanations for the string of jobless recoveries in developed economies in recent decades. Expectations that the (already eroding) fallaciousness will fall off sharply in coming decades underlie the fear of structural shift.

Writers such as Rifkin, Brain, and Ford often suggest that the structure of the economy will have to shift to a basic income because its present structural foundation (trading labor for income) will no longer be an available option on a full employment basis. It would perhaps be available to only 90% of workers in the next decade, perhaps 75% of workers a decade after that, and so on. Often included in the basic income idea is an element of civic obligation, such that able people must somehow contribute civically in order to receive the basic income. The labor-market economy (trading labor for income) already achieves that outcome today (because working for income generally produces civic value in various ways, directly and indirectly), but the argument is that advanced automation will decouple the linkage that makes that possible. Thus the same result (trading civic value for income) would have to be driven by different forces—either non-market ones, or via a new kind of market. The non-market idea seems infeasible given the generally abysmal performance record of planned economies. But the idea of engineered new markets leaves room for the disciplining and motivating powers that make capitalist markets capable of positively shaping human behavior where government alone is usually unable.

New-market engineering

Brain and Ford's books, in stark contrast to Rifkin's, came later and were written by engineers with extensive under-the-hood knowledge of modern production methods, computer hardware and software, and the Internet. They explicitly reject non-market solutions as unworkable and instead suggest new kinds of markets. Rather than being "post-market" proponents, such authors could be called "new-market" proponents. They vigorously distance themselves from socialism or welfare states—generally seeking to keep a market economy with private enterprise, which they believe cannot be preserved *unless* its foundation is modified from its current structure. Thus, quite contrary to being anti-market agents (as critics might suppose them to be), they believe themselves to be *salvaging* markets from destruction. They envision creating consumer purchasing power by some other mechanism than the traditional labor market as we have known it so far, in order that free markets may continue to provide the invisible hand component of production-possibilities decisions. In other words, they believe that market forces are necessary to generate allocative efficiency, and they believe that without a structural modification that (at least partially) decouples purchasing power (and consumer confidence) from employment determined by the traditional labor market, there will be a systemic market failure, which they seek to avoid.

Just as new-market advocates are pro-market and pro-private-property, they are also very much non-Luddite (in fact, exactly opposite of Luddite) in the respect that they *like* technology—they don't *hate* it. They want it to continue advancing as robustly as ever. They simply feel that income and purchasing power must be decoupled from human participation in production. (The decoupling does not have to happen all at once; it could start small and gradually increase.) If that happens, then they essentially do not have any problem with technology or automation, per se. In contrast to old-style welfare, they do not feel that income should be unconditional, or equal, or "free" (given out "for nothing"). They believe that people should have to work for it (in a new sense of the word "work"), in the respect that they are given incentives to do positive things, like take classes, read books, conserve (or remediate) environmental resources, and so on. People would be paid to do civically valuable things, and if they chose not to do those things, they would not be paid. In this way, new-market advocates align themselves with human nature, which generally requires selfish motivations and incentives to shape behavior, and with the market's invisible hand, which is needed to make the right production-possibilities decisions (because the idea that individual human managers, or groups of them, are capable of making those decisions correctly with zero invisible-hand assistance has been empirically discredited).

Discussions of new-market ideas usually lead to the topic of post-scarcity economic paradigms. But new-market engineers argue that without clear-eyed, realist engineering of the intermediate steps, there is an abyss of dysfunction and hardship between today's economy of scarcity and the starry-eyed goal of any fully developed economy of abundance.

Wage-recapture market variant

Ford's main new-market mechanism would be to create a tax that recaptures most (not all) of the value that firms and their customers gain from eliminating wages, then use the tax revenue to pay people for doing civically valuable actions—that is, pursuing activities, such as higher education or environmental preservation, that have positive externalities. The main reason for paying these "wages" need not be their altruistic or environmentalist components; the main reason is simply to prevent the market economy from collapsing due to noncirculation of value (that is, the lack of adequate trade which would occur if lack of consumer purchasing power and confidence left no way for an adequate mass market to exist). Ford points out that the tax could not take *all* of the gains away from the corporations and their customers, because this would destroy the natural incentive to innovate that a market economy needs to be sustainable. The value would be split between the innovators, their customers, and the rest of the population, because leaving out any of that trio would wreck the sustainability of the model. (The leaving out of the third leg is what is causing today's economic pathologies and promising tomorrow's, in the view of new-market engineers.) Ford's idea is an earnest market effort because it preserves the invisible hand as the maker of production-possibilities decisions for goods and services. However, it does rely on human planning (via a technocratic government agency in each country) to make the production-possibilities decisions for civic actions. The latter is viewed as unfortunate but necessary due to the lack of an alternative.

Mirror-image market variant

Another idea for a new-market mainspring which solves the aforementioned "lack of an alternative" problem is a "mirror image" idea, which has an even more private-sector approach in which the invisible hand helps make even the civic-actions production-possibilities decisions. In this model, the government does not collect a wage-recapture tax at all. Instead of enforcing tax payment, it only enforces payment of a new-style "wage" directly from corporations to consumers that looks to us today like something we might label "mandatory philanthropy", but which would actually be a true market wage of a new type. In today's old market, money flows from consumers, through (partially automated) companies, past the eyes of the government enforcement sentry (but not through its hands) as wages, into the hands of workers (who are also the consumers, thus completing the cycle of value recirculation). In the new market (mirror-image variant), money would flow from consumers, through (highly automated) companies, past the eyes of the government enforcement sentry (but not through its hands) as [new-style] "wages", and into the hands of [new-style] "wage" earners, who are paid the "wage" for civically valuable actions. (They are also the consumers, thus completing the cycle of value recirculation).

In this model, the decisions about what the civic actions are can be made by the invisible hand, because each "mandated philanthropist" gets a large degree of authority in what actions their "philanthropy" (which is actually [functionally] a new-style "payroll") will or won't pay for. Many such paymasters functioning simultaneously could constitute the "buyers" in a market for civically valuable actions (with mass-market "workers" as the "sellers"). There would still be *some* regulation involved, because, for example, it would be illegal to base the "payroll" decisions on race, color, religion, creed, gender, sexual orientation, ethnicity, disability, marital or veteran status, and so forth. To decide which "workers" were on a given "payroll", there might be a clearinghouse to randomly match the two, rotating assignments every several years. Or perhaps the businesses that run the "payroll" could even "hire" the "workers" themselves, in which case workers would compete for "jobs" by showing off how "productive" they could be in doing the civic actions (another level of invisible hand yet again). The "mirror image" name comes from the idea that this variant of new market is a very free market where the invisible hand remains just as powerful as it was in the 1945-2008 economy, but with many mirror-image aspects (which are visualized above by the amount of quotation marks that are necessary to signify mirror-image senses for words that were always [up till now] widely known only in their non-mirror-image senses).

The axis of reflection in the mirroring seems to be, at root, a "polarity shift" from where human individuals can add value only by *doing production* (from within production systems) to where they can also add value by *avoiding hurting production systems* (from outside). The hurt-avoidance comes from such civic actions as providing goods-and-services demand via consumption (which the system requires in order to stay running) instead of failing to consume (because of lack of income); ensuring the sustainable supply of energy and environmental resources to the production systems (by avoiding *over*consuming those); and ensuring the supply of people educated enough to provide the few humans that the production systems will need in the future, by pursuing education and cognitively enriching pastimes. The humans that the systems need will be few, but those few will need to be highly intelligent, talented, and educated, constituting a human resource that might be endangered if the general population does not act as a "farm team system" for it by valuing education and self-education as a civic action. An analogy is provided by sports' relationship to general life. Few humans are talented and practiced enough to play professional sports, but the professional teams rely on a system that filters such scarce people out of the general population via little league/pee wee programs, high school play, college play, farm-team play, etc. People in the general population are not considered inferior human beings (versus the pro players) because of their lack of pro talent. They are valued as the fans and ticket-buyers that make the pro system economically viable. And a small fraction of them grow up to become pros themselves.

In today's old market, governments enforce the payment of wages by having outlawed their nonpayment (i.e., slavery); by levying tariffs on cheap competition from countries that kept their nonpayment (slavery) (that outlawing has now been global for many decades); and by attempting to minimize their underpayment (i.e., wage slavery, a sharply cheapened value of work [with elites and their customers keeping the money]). In the new market (mirror-image variant), governments enforce the payment of "wages" by outlawing their nonpayment (i.e., evading the "payroll"); by levying tariffs on cheap competition from countries that kept their nonpayment (non-participating countries); and by attempting to minimize their underpayment (i.e., "wage" slavery, a sharply cheapened value of civic actions [with elites and their customers keeping the money]).

One of the inherent challenges of the mirror-image variant is that various forms of dressing up corporations' financial self-interest in a specious cloak of civic virtue would inevitably arise. This would be a "washing" form of marketing and operations that included greenwashing and analogous washing in other domains of life (e.g., education, infrastructure). It seems unlikely that this can be entirely negated; instead, it would have to be perennially pruned by social censure and regulatory oversight. However, no other system is without its chronic weaknesses, either. For example, the classical variants of capitalism (implemented thus far) have scored poorly on various tests, such as environmental sustainability and (potentially) the employability of the average human (as that was traditionally defined) as automation grows pervasive. Twentieth-century variants of communism fared even worse in environmental sustainability, and also failed economically in average standard of living and politically in individual freedom. The wage-recapture new-market variant, with its technocratic decisions on how to spend the revenue, holds promise to minimize the corporate "washing" problem, yet it also holds risks of failing on allocative efficiency and market-driven innovation, which the mirror-image variant mitigates. As elsewhere in reality, each choice has pros and cons, rather than any choice being perfect. The "washing" problem may be the mirror-image analog of classical capitalism's tendency to exaggerate needs (for example, a maker of antibacterial soaps encouraging the populace to fear microbes to an irrational degree). Both are forms of conflict of interest that cause "chronic irritation" to a socioeconomic system but need not be "fatal" to it if given adequate "medical management". The washing problem may be less systemically injurious than the allocative inefficiency problem, just as the "exaggerated needs" problem of classical capitalism was less systemically injurious than the allocative inefficiency problem of central economic planning.

In choosing the decider of production-possibilities decisions (whether of goods, services, or civic actions), the invisible hand is generally preferred to committees of humans because it has proven to be superior at the decision making (except for regulatory issues such as race-color-religion-etc and the "washing" discussed above). In the future it will also be necessary to ask what role artificial intelligence might possibly have in making those decisions, and whether humans would allow it. Perhaps artificial intelligence, like human intelligence, will share the role with the invisible hand but be barred from usurping the entirety of it.

Implementations

Regarding the chances of any new-market ideas being implemented, there are both significant barriers and significant drivers, with a net potential of perhaps "even chances". Ford discusses many of these barriers and drivers. The barrier side includes (a) natural cultural conservatism that powerfully resists systemic changes; (b) the powerful influence of laissez-faire ideals, which would resist any engineered systemic change to markets (especially *anything* involving a tax); (c) the fact that early implementation by individual countries faces an immediate threat from the export and offshoring competition of countries that *haven't* yet implemented; and (relatedly) (d) the all-or-nothing problem, which may occur if a new system would work well but only if the switch from old to new was an off-on switching rather than a gradual evolution. However, on the driver side there are powerful forces that may answer all of the barriers. Foremost would be a dawning realization by economic elites that they have a choice between a new market with prosperity, or the old market spiraling into near-total failure. Globalization so far has not threatened the wallets of economic elites (only those of average workers), and has in fact enriched the elites thus far; but the changing parameter values of the economic system as automation advances would alter that runtime environment and transform it into a new one, where even the elites' wealth would be threatened by a market failure that killed their businesses and reduced asset values throughout the economy. Realizing these options, elites might actually switch from opposing new markets to actively supporting their implementation (including addressing the competition between countries whose policies differed). The all-or-nothing problem does not have to occur if an implementation is engineered such that extremely profitable, extremely automated industries began piloting new markets while other industries continued with an old-market status quo for quite some time. In this model, the early adopters voluntarily become leaders, and the pilot projects would act simply as economic stimulus on the broader economy (although a type of stimulus much more effective than old-style stimulus, whose efficacy seems to be eroding because it relies on the Luddite premise being a total fallacy as opposed to shifting by degrees out of total fallaciousness). The overall transition in this model could actually be quite painless, as a generally prosperous economy changed gradually over decades from mostly-old-with-some-new to mostly-new-with-some-old. The selfish motivation of the early-adopter leaders would be the aforementioned choice faced by economic elites. They would choose to stimulate the broader economy because that result would ensure their own continuing strong sales and growth by preserving a runtime environment of general prosperity for them to operate within, without which depression or malaise would occur.

Given the aforementioned choice faced by economic elites, those in the private sector might even choose to pursue the new market without government involvement. But the private sector faces two hurdles that would make it difficult: the natural competition between firms (which is necessary and thus must be protected by competition law), and legal obligation to maximize shareholder value. The traditional definitions of shareholder value evolved in an earlier era whose commercial environment had different parameter values due to lack of advanced automation. Those traditional definitions would bar new-style payrolls. But in the face of market failure without them, perhaps a case would emerge for an updated definition. Competition is the other hurdle. Companies are barred by competition law from agreeing to limit competition, and even if they weren't, individual companies generally cannot make the first move of increasing expense without being killed by competition from rivals who don't. This is why "a level playing field" would have to be created by policy, or to use a different analogy, "a high tide that lifted all boats equally". This is directly analogous to existing minimum wage laws. Individual companies generally could not survive in the market if they volunteered to self-enforce minimum limits on wages (in the absence of any laws requiring them). There *is* breathing room for above-market wages (e.g., to attract superior talent) at some companies in some industries who enjoy a relatively high level of imperfect competition; but most companies in most industries face competition too close to pure to survive the attempt. In this sense, the mandated value recirculation (whatever anyone calls it, from "wage recapture" to "new-style wages") is as unremarkable and non-novel an idea as any legislative or regulatory mechanism in commerce. For goals that make long-term systemic balance possible but cannot be pursued by the self-interest of individual market players, these mechanisms provide a path by forcing all competitors to play the game by the same rules. Existing examples include employment standards (e.g., child labor laws, minimum wage laws), environmental protection, and financial regulation (to prevent bubbles and thus crashes). These exist in perennial tension with the forces of pure capitalism; thus the extremes perform checks and balances on each other. Businesses usually fight for inadequate regulation; government usually fights for excessive regulation; and a sustainable balance results. Over decades, systemic pathologies gradually push the balance point out of the sustainable range; periodic breakdowns then yield correction by counteractive forces (e.g., trust-busting [leftward correction], the Reagan revolution [rightward correction]).

Laissez-faire ideals reigned supreme worldwide for about three decades (roughly centered on the fall of the Soviet Bloc, which vindicated capitalism over central planning in many ways). In this environment, where the lesson commonly extrapolated was that pure capitalism will always be better than any mixed-economy alternatives, the prevailing theory has been that higher corporate taxes can only harm economic prosperity. The reasoning is partly that countries can simply compete to undercut each other's corporate tax rates (which is true), but also, more importantly, that only the invisible hand is capable of recirculating capital back toward the base of the economy in a successful manner (which is not to be dismissed lightly, and may in fact be true). The disparaging label for such ideas is "trickle-down economics", but many intelligent people have earnestly believed in these ideals; and the fact that their discounting has often been facile and done by imperfect opponents has only encouraged believers to stay faithful. Widespread fervor for trickle-down beliefs (in both the public and private sectors) poses a formidable barrier to the wage-recapture-tax new-market variant. But these conventional beliefs rely on the assumption that the Luddite premise is entirely and eternally fallacious. Unfortunately, there has already been a decade of empirical evidence that low taxes, new business investment, and economic growth no longer have a sure-fire correlation to strong, "good-jobs" employment in developed economies. If the Luddite premise has been starting to shift into partial accuracy, then no amount of continued low taxes and deregulation will ever be able to produce enough trickling down to create broad-based prosperity. In that case, mandating the payment of new-style wages could recirculate value back to the base of the mass market. The promise of the mirror-image variant would be that humans need not turn to taxes for the value recovery (at the top) nor to central planning for the distribution details (at the bottom). As long as the "minimum wage" (referring to the new-style wages) and other employment standards are being enforced, then government's role ends there.

Conceivably, both the Luddite premise and the lump of labor premise could change states in continuous-graph fashion, from completely false to partially true, depending on parameter values in the commercial environment, most specifically, the available modes of value recirculation. In this view, for two centuries they were completely false or very nearly so, because the traditional labor market provided sufficient means of value recirculation. As that fails because of advanced automation, they could enter partial influence. But if new forms of broad-based personal income came into being (for example, basic income, guaranteed minimum income, or new-style wages), they could revert back to a state of complete or near-complete falsehood again. The difference would be that enough value was circulating broadly enough through a mass market of consumers and corporations that services which today could not possibly garner middle-class wages and benefits (for example, full-time jobs reading stories to hospitalized children, painting murals, or attending university) would become viable at that wage level.

Other goals of automation (beyond productivity gains and cost reduction)

In manufacturing, the purpose of automation has shifted to issues broader than productivity and costs.

Reliability and precision

The old focus on using automation simply to increase productivity and reduce costs was seen to be short-sighted, because it is also necessary to provide a skilled workforce who can make repairs and manage the machinery. Moreover, the initial costs of automation were high and often could not be recovered by the time entirely new manufacturing processes replaced the old. (Japan's "robot junkyards" were once world famous in the manufacturing industry.)

Automation is now often applied primarily to increase quality in the manufacturing process, where automation can increase quality substantially. For example, automobile and truck pistons used to be installed into engines manually. This is rapidly being transitioned to automated machine installation, because the error rate for manual installment was around 1-1.5%, but has been reduced to 0.00001% with automation.

Health and environment

The costs of automation to the environment are different depending on the technology, product or engine automated. There are automated engines that consume more energy resources from the Earth in comparison with previous engines and those that do the opposite too. Hazardous operations, such as oil refining, the manufacturing of industrial chemicals, and all forms of metal working, were always early contenders for automation.

Convertibility and turnaround time

Another major shift in automation is the increased demand for flexibility and convertibility in manufacturing processes. Manufacturers are increasingly demanding the ability to easily switch from manufacturing Product A to manufacturing Product B without having to completely rebuild the production lines. Flexibility and distributed processes have led to the introduction of Automated Guided Vehicles with Natural Features Navigation.

Digital electronics helped too. Former analogue-based instrumentation was replaced by digital equivalents which can be more accurate and flexible, and offer greater scope for more sophisticated configuration, parametrization and operation. This was accompanied by the fieldbus revolution which provided a networked (i.e. a single cable) means of communicating between control systems and field level instrumentation, eliminating hard-wiring.

Discrete manufacturing plants adopted these technologies fast. The more conservative process industries with their longer plant life cycles have been slower to adopt and analogue-based measurement and control still dominates. The growing use of Industrial Ethernet on the factory floor is pushing these trends still further, enabling manufacturing plants to be integrated more tightly within the enterprise, via the internet if necessary. Global competition has also increased demand for Reconfigurable Manufacturing Systems.

Automation tools

Engineers now can have numerical control over automated devices. The result has been a rapidly expanding range of applications and human activities. Computer-aided technologies (or CAx) now serve the basis for mathematical and organizational tools used to create complex systems. Notable examples of CAx include Computer-aided design (CAD software) and Computer-aided manufacturing (CAM software). The improved design, analysis, and manufacture of products enabled by CAx has been beneficial for industry.

Information technology, together with industrial machinery and processes, can assist in the design, implementation, and monitoring of control systems. One example of an industrial control system is a programmable logic controller (PLC). PLCs are specialized hardened computers which are frequently used to synchronize the flow of inputs from (physical) sensors and events with the flow of outputs to actuators and events.

Human-machine interfaces (HMI) or computer human interfaces (CHI), formerly known as *man-machine interfaces*, are usually employed to communicate with PLCs and other computers. Service personnel who monitor and control through HMIs can be called by different names. In industrial process and manufacturing environments, they are called operators or something similar. In boiler houses and central utilities departments they are called stationary engineers.

Different types of automation tools exist:

- ANN - Artificial neural network
- DCS - Distributed Control System
- HMI - Human Machine Interface
- SCADA - Supervisory Control and Data Acquisition
- PLC - Programmable Logic Controller
- PAC - Programmable automation controller
- Instrumentation
- Motion control
- Robotics

Current limits

Many roles for humans in industrial processes presently lie beyond the scope of automation. Human-level pattern recognition, language recognition, and language production ability are well beyond the capabilities of modern mechanical and computer systems. Tasks requiring subjective assessment or synthesis of complex sensory data, such as scents and sounds, as well as high-level tasks such as strategic planning, currently require human expertise. In many cases, the use of humans is more cost-effective than mechanical approaches even where automation of industrial tasks is possible.

Applications of Automation

- **Automated Video surveillance:**

The Defense Advanced Research Projects Agency (DARPA) started the research and development of automated Visual surveillance and Monitoring (VSAM) program 1997-99 and airborne Video Surveillance (AVS) program 1998-2002. Currently there is a major effort underway in the vision community to develop a fully automated tracking surveillance system. Automated video surveillance monitors people and vehicle in real time within a busy environment. Existing automated surveillance systems are based on the environment they are primarily designed to observe, i.e., indoor, outdoor or airborne, the amount of sensors that the automated system can handle and the mobility of sensor, i.e., stationary camera vs. mobile camera. The purpose of a surveillance system is to record properties and trajectories of objects in a given area, generate warnings or notify designated authority in case of occurrence of particular events.

- **Automated Highway Systems:**

As demands for safety and mobility have grown and technological possibilities have multiplied, interest in automation have grown. Seeking to accelerate the development and introduction of fully automated vehicles and highways, The United States Congress authorized more than $650 million over 6 years for intelligent transport systems (ITS) and demonstration projects in the 1991 Intermodal Surface Transportation Efficiency Act (ISTEA). Congress legislated in ISTEA that "The secretary [of transportation] shall develop an automated highway and vehicle prototype from which future fully automated intelligent vehicle-highway systems can be developed. Such development shall include research in human factors to ensure the success of the man-machine relationship. The goal of this program is to have the first fully automated highway roadway or an automated test track in operation by 1997. This system shall accommodate installation of equipment in new and existing motor vehicles." [ISTEA 1991, part B, Section 6054(b)].

Full automation commonly defined as requiring no control or very limited control by the driver; such automation would be accomplished through a combination of sensor, computer, and communications systems in vehicles and along the roadway. Fully automated driving would, in theory, allow closer vehicle spacing and higher speeds, which could enhance traffic capacity in places where additional road building is physically impossible, politically unacceptable, or prohibitively expensive. Automated controls also might enhance road safety by reducing the opportunity for driver error, which causes a large share of motor vehicle crashes. Other potential benefits include improved air quality (as a result of more-efficient traffic flows), increased fuel economy, and spin-off technologies generated during research and development related to automated highway systems.

- **Automated manufacturing:**

Automated manufacturing refers to the application of automation to produce things in the factory way. Most of the advantages of the automation technology has its influence in the manufacture processes.

The main advantage of the automated manufacturing are: higher consistency and quality, reduce the lead times, simplification of production, reduce handling, improve work flow and increase the morale of workers when a good implementation of the automation is made.

- **Home Automation**

Home automation (also called domotics) designates an emerging practice of increased automation of household appliances and features in residential dwellings, particularly through electronic means that allow for things impracticable, overly expensive or simply not possible in recent past decades.

Chapter 2

Mechanical Engineering

Mechanical engineers design and build engines and power plants

structures and vehicles of all sizes.

Mechanical engineering is a discipline of engineering that applies the principles of physics and materials science for analysis, design, manufacturing, and maintenance of mechanical systems. It is the branch of engineering that involves the production and usage of heat and mechanical power for the design, production, and operation of machines and tools. It is one of the oldest and broadest engineering disciplines.

The engineering field requires a vast understanding of core concepts including mechanics, kinematics, thermodynamics, materials science, and structural analysis. Mechanical engineers use these core principles along with tools like computer-aided engineering and product lifecycle management to design and analyze manufacturing plants, industrial equipment and machinery, heating and cooling systems, motorized vehicles, aircraft, watercraft, robotics, medical devices and more.

Mechanical engineering emerged as a field during the industrial revolution in Europe in the 19th century; however, its development can be traced back several thousand years around the world. The field has continually evolved to incorporate advancements in technology, and mechanical engineers today are pursuing developments in such fields as composites, mechatronics, and nanotechnology. Mechanical engineering overlaps with aerospace engineering, civil engineering, electrical engineering, and petroleum engineering to varying amounts.

Development

Applications of mechanical engineering are found in the records of many ancient and medieval societies throughout the globe. In ancient Greece, the works of Archimedes (287 BC–212 BC) deeply influenced mechanics in the Western tradition and Heron of Alexandria (c. 10–70 AD) created the first steam engine. In China, Zhang Heng (78–139 AD) improved a water clock and invented a seismometer, and Ma Jun (200–265 AD) invented a chariot with differential gears. The medieval Chinese horologist and engineer Su Song (1020–1101 AD) incorporated an escapement mechanism into his astronomical clock tower two centuries before any escapement can be found in clocks of medieval Europe, as well as the world's first known endless power-transmitting chain drive.

During the years from 7th to 15th century, the era called the Islamic Golden Age, there were remarkable contributions from Muslim inventors in the field of mechanical technology. Al-Jazari, who was one of them, wrote his famous *Book of Knowledge of Ingenious Mechanical Devices* in 1206, and presented many mechanical designs. He is also considered to be the inventor of such mechanical devices which now form the very basic of mechanisms, such as the crankshaft and camshaft.

Important breakthroughs in the foundations of mechanical engineering occurred in England during the 17th century when Sir Isaac Newton both formulated the three Newton's Laws of Motion and developed calculus. Newton was reluctant to publish his methods and laws for years, but he was finally persuaded to do so by his colleagues, such as Sir Edmund Halley, much to the benefit of all mankind.

During the early 19th century in England, Germany and Scotland, the development of machine tools led mechanical engineering to develop as a separate field within engineering, providing manufacturing machines and the engines to power them. The first British professional society of mechanical engineers was formed in 1847 Institution of Mechanical Engineers, thirty years after the civil engineers formed the first such professional society Institution of Civil Engineers. On the European continent, Johann Von Zimmermann (1820–1901) founded the first factory for grinding machines in Chemnitz (Germany) in 1848.

In the United States, the American Society of Mechanical Engineers (ASME) was formed in 1880, becoming the third such professional engineering society, after the American Society of Civil Engineers (1852) and the American Institute of Mining Engineers (1871). The first schools in the United States to offer an engineering education were the United States Military Academy in 1817, an institution now known as Norwich University in 1819, and Rensselaer Polytechnic Institute in 1825. Education in mechanical engineering has historically been based on a strong foundation in mathematics and science.

Education

Degrees in mechanical engineering are offered at universities worldwide. In Bangladesh, Brazil, China, England, India, Israel, Nepal, North America, and Pakistan, mechanical engineering programs typically take four to five years of study and result in a Bachelor of Science (B.Sc), Bachelor of Technology (B.Tech), Bachelor of Engineering (B.Eng), or Bachelor of Applied Science (B.A.Sc) degree, in or with emphasis in mechanical engineering. In Spain, Portugal and most of South America, where neither BSc nor BTech programs have been adopted, the formal name for the degree is "Mechanical Engineer", and the course work is based on five or six years of training. In Italy the course work is based on five years of training; but in order to qualify as an Engineer you have to pass a state exam at the end of the course.

In Australia, mechanical engineering degrees are awarded as Bachelor of Engineering (Mechanical). The degree takes four years of full time study to achieve. To ensure quality in engineering degrees, the Australian Institution of Engineers accredits engineering degrees awarded by Australian universities. Before the degree can be awarded, the student must complete at least 3 months of on the job work experience in an engineering firm.

In the United States, most undergraduate mechanical engineering programs are accredited by the Accreditation Board for Engineering and Technology (ABET) to ensure similar course requirements and standards among universities. The ABET web site lists 276 accredited mechanical engineering programs as of June 19, 2006. Mechanical engineering programs in Canada are accredited by the Canadian Engineering Accreditation Board (CEAB), and most other countries offering engineering degrees have similar accreditation societies.

Some mechanical engineers go on to pursue a postgraduate degree such as a Master of Engineering, Master of Technology, Master of Science, Master of Engineering Management (MEng.Mgt or MEM), a Doctor of Philosophy in engineering (EngD, PhD) or an engineer's degree. The master's and engineer's degrees may or may not include research. The Doctor of Philosophy includes a significant research component and is often viewed as the entry point to academia. The Engineer's degree exists at a few institutions at an intermediate level between the master's degree and the doctorate.

Coursework

Standards set by each country's accreditation society are intended to provide uniformity in fundamental subject material, promote competence among graduating engineers, and to maintain confidence in the engineering profession as a whole. Engineering programs in the U.S., for example, are required by ABET to show that their students can "work professionally in both thermal and mechanical systems areas." The specific courses required to graduate, however, may differ from program to program. Universities and Institutes of technology will often combine multiple subjects into a single class or split a subject into multiple classes, depending on the faculty available and the university's major area(s) of research.

The fundamental subjects of mechanical engineering usually include:

- Statics and dynamics
- Strength of materials and solid mechanics
- Instrumentation and measurement
- Electrotechnology
- Thermodynamics, heat transfer, energy conversion, and HVAC
- Fluid mechanics and fluid dynamics
- Mechanism design (including kinematics and dynamics)
- Manufacturing engineering, technology, or processes
- Hydraulics and pneumatics
- Mathematics - in particular, calculus, differential equations, and linear algebra.
- Engineering design
- Mechatronics and control theory
- Material Engineering

- Design engineering, Drafting, computer-aided design (CAD) (including solid modeling), and computer-aided manufacturing (CAM)

Mechanical engineers are also expected to understand and be able to apply basic concepts from chemistry, physics, chemical engineering, civil engineering, and electrical engineering. Most mechanical engineering programs include multiple semesters of calculus, as well as advanced mathematical concepts including differential equations, partial differential equations, linear algebra, abstract algebra, and differential geometry, among others.

In addition to the core mechanical engineering curriculum, many mechanical engineering programs offer more specialized programs and classes, such as robotics, transport and logistics, cryogenics, fuel technology, automotive engineering, biomechanics, vibration, optics and others, if a separate department does not exist for these subjects.

Most mechanical engineering programs also require varying amounts of research or community projects to gain practical problem-solving experience. In the United States it is common for mechanical engineering students to complete one or more internships while studying, though this is not typically mandated by the university. Cooperative education is another option.

License

Engineers may seek license by a state, provincial, or national government. The purpose of this process is to ensure that engineers possess the necessary technical knowledge, real-world experience, and knowledge of the local legal system to practice engineering at a professional level. Once certified, the engineer is given the title of Professional Engineer (in the United States, Canada, Japan, South Korea, Bangladesh and South Africa), Chartered Engineer (in the United Kingdom, Ireland, India and Zimbabwe), *Chartered Professional Engineer* (in Australia and New Zealand) or *European Engineer* (much of the European Union). Not all mechanical engineers choose to become licensed; those that do can be distinguished as Chartered or Professional Engineers by the post-nominal title P.E., P.Eng., or C.Eng., as in: Mike Thompson, P.Eng.

In the U.S., to become a licensed Professional Engineer, an engineer must pass the comprehensive FE (Fundamentals of Engineering) exam, work a given number of years as an *Engineering Intern (EI)* or *Engineer-in-Training (EIT)*, and finally pass the "Principles and Practice" or PE (Practicing Engineer or Professional Engineer) exams.

In the United States, the requirements and steps of this process are set forth by the National Council of Examiners for Engineering and Surveying (NCEES), a national non-profit representing all states. In the UK, current graduates require a BEng plus an appropriate masters degree or an integrated MEng degree, a minimum of 4 years post graduate on the job competency development, and a peer reviewed project report in the candidates specialty area in order to become chartered through the Institution of Mechanical Engineers.

In most modern countries, certain engineering tasks, such as the design of bridges, electric power plants, and chemical plants, must be approved by a Professional Engineer or a Chartered Engineer. "Only a licensed engineer, for instance, may prepare, sign, seal and submit engineering plans and drawings to a public authority for approval, or to seal engineering work for public and private clients." This requirement can be written into state and provincial legislation, such as in the Canadian provinces, for example the Ontario or Quebec's Engineer Act.

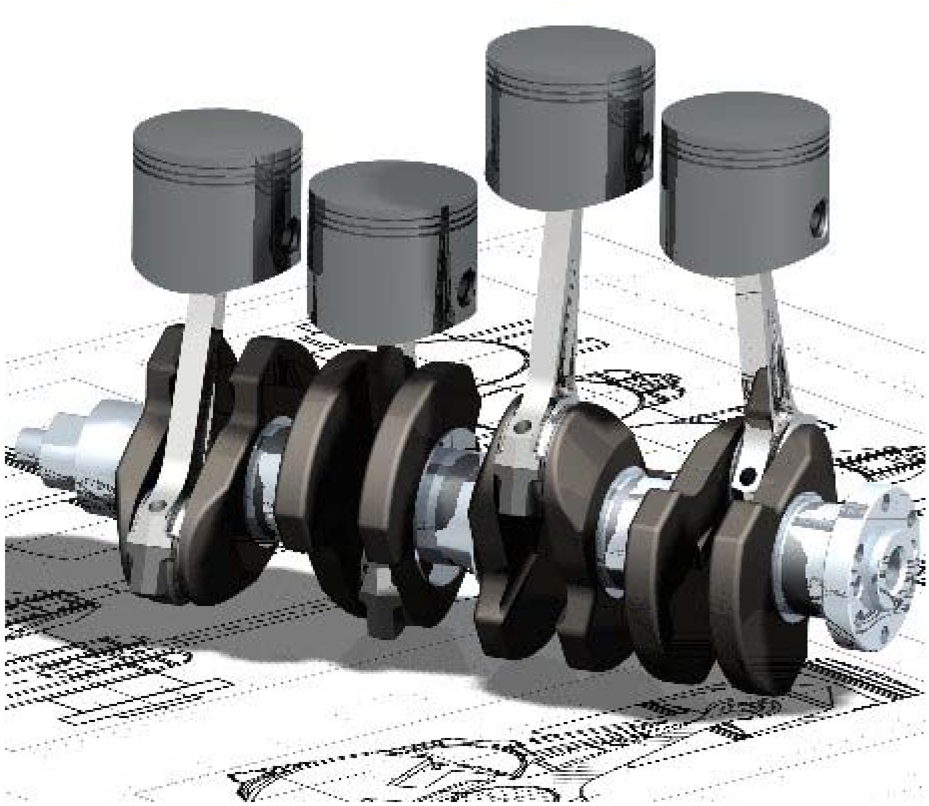
In other countries, such as Australia, no such legislation exists; however, practically all certifying bodies maintain a code of ethics independent of legislation that they expect all members to abide by or risk expulsion.

Salaries and workforce statistics

The total number of engineers employed in the U.S. in 2009 was roughly 1.6 million. Of these, 239,000 were mechanical engineers (14.9%), the second largest discipline by size behind civil (278,000). The total number of mechanical engineering jobs in 2009 was projected to grow 6% over the next decade, with average starting salaries being $58,800 with a bachelor's degree. The median annual income of mechanical engineers in the U.S. workforce was roughly $74,900. This number was highest when working for the government ($86,250), and lowest in education ($63,050).

In 2007, Canadian engineers made an average of CAD$29.83 per hour with 4% unemployed. The average for all occupations was $18.07 per hour with 7% unemployed. Twelve percent of these engineers were self-employed, and since 1997 the proportion of female engineers had risen to 6%.

Modern tools



An oblique view of a four-cylinder inline crankshaft with pistons

Many mechanical engineering companies, especially those in industrialized nations, have begun to incorporate computer-aided engineering (CAE) programs into their existing design and analysis processes, including 2D and 3D solid modeling computer-aided design (CAD). This method has many benefits, including easier and more exhaustive visualization of products, the ability to create virtual assemblies of parts, and the ease of use in designing mating interfaces and tolerances.

Other CAE programs commonly used by mechanical engineers include product lifecycle management (PLM) tools and analysis tools used to perform complex simulations. Analysis tools may be used to predict product response to expected loads, including fatigue life and manufacturability. These tools include finite element analysis (FEA), computational fluid dynamics (CFD), and computer-aided manufacturing (CAM).

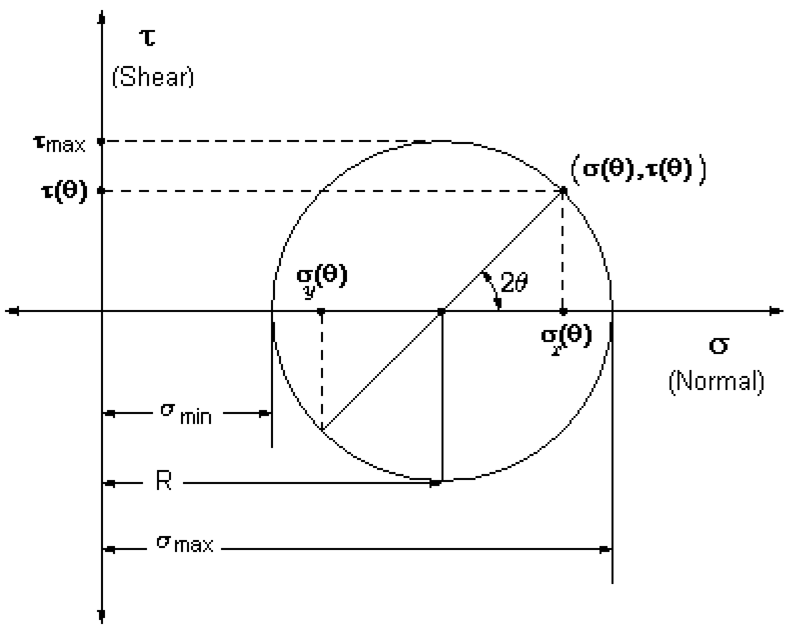
Using CAE programs, a mechanical design team can quickly and cheaply iterate the design process to develop a product that better meets cost, performance, and other constraints. No physical prototype need be created until the design nears completion, allowing hundreds or thousands of designs to be evaluated, instead of a relative few. In addition, CAE analysis programs can model complicated physical phenomena which cannot be solved by hand, such as viscoelasticity, complex contact between mating parts, or non-Newtonian flows

As mechanical engineering begins to merge with other disciplines, as seen in mechatronics, multidisciplinary design optimization (MDO) is being used with other CAE programs to automate and improve the iterative design process. MDO tools wrap around existing CAE processes, allowing product evaluation to continue even after the analyst goes home for the day. They also utilize sophisticated optimization algorithms to more intelligently explore possible designs, often finding better, innovative solutions to difficult multidisciplinary design problems.

Subdisciplines

The field of mechanical engineering can be thought of as a collection of many mechanical disciplines. Several of these subdisciplines which are typically taught at the undergraduate level are listed below, with a brief explanation and the most common application of each. Some of these subdisciplines are unique to mechanical engineering, while others are a combination of mechanical engineering and one or more other disciplines. Most work that a mechanical engineer does uses skills and techniques from several of these subdisciplines, as well as specialized subdisciplines.

Mechanics



Mohr's circle, a common tool to study stresses in a mechanical element

Mechanics is, in the most general sense, the study of forces and their effect upon matter. Typically, engineering mechanics is used to analyze and predict the acceleration and deformation (both elastic and plastic) of objects under known forces (also called loads) or stresses. Subdisciplines of mechanics include

- Statics, the study of non-moving bodies under known loads, how forces affect static bodies
- Dynamics (or kinetics), the study of how forces affect moving bodies
- Mechanics of materials, the study of how different materials deform under various types of stress
- Fluid mechanics, the study of how fluids react to forces
- Continuum mechanics, a method of applying mechanics that assumes that objects are continuous (rather than discrete)

Mechanical engineers typically use mechanics in the design or analysis phases of engineering. If the engineering project were the design of a vehicle, statics might be employed to design the frame of the vehicle, in order to evaluate where the stresses will be most intense. Dynamics might be used when designing the car's engine, to evaluate the forces in the pistons and cams as the engine cycles. Mechanics of materials might be used to choose appropriate materials for the frame and engine. Fluid mechanics might be used to design a ventilation system for the vehicle, or to design the intake system for the engine.

Kinematics

Kinematics is the study of the motion of bodies (objects) and systems (groups of objects), while ignoring the forces that cause the motion. The movement of a crane and the oscillations of a piston in an engine are both simple kinematic systems. The crane is a type of open kinematic chain, while the piston is part of a closed four-bar linkage.

Mechanical engineers typically use kinematics in the design and analysis of mechanisms. Kinematics can be used to find the possible range of motion for a given mechanism, or, working in reverse, can be used to design a mechanism that has a desired range of motion.

Mechatronics and robotics

Training FMS with learning robot SCORBOT-ER 4u, workbench CNC Mill and CNC Lathe

Mechatronics is an interdisciplinary branch of mechanical engineering, electrical engineering and software engineering that is concerned with integrating electrical and mechanical engineering to create hybrid systems. In this way, machines can be automated through the use of electric motors, servo-mechanisms, and other electrical systems in conjunction with special software. A common example of a mechatronics system is a CD-ROM drive. Mechanical systems open and close the drive, spin the CD and move the laser, while an optical system reads the data on the CD and converts it to bits. Integrated software controls the process and communicates the contents of the CD to the computer.

Robotics is the application of mechatronics to create robots, which are often used in industry to perform tasks that are dangerous, unpleasant, or repetitive. These robots may be of any shape and size, but all are preprogrammed and interact physically with the world. To create a robot, an engineer typically employs kinematics (to determine the robot's range of motion) and mechanics (to determine the stresses within the robot).

Robots are used extensively in industrial engineering. They allow businesses to save money on labor, perform tasks that are either too dangerous or too precise for humans to perform them economically, and to insure better quality. Many companies employ assembly lines of robots,especially in Automotive Industries and some factories are so robotized that they can run by themselves. Outside the factory, robots have been employed in bomb disposal, space exploration, and many other fields. Robots are also sold for various residential applications.

Structural analysis

Structural analysis is the branch of mechanical engineering (and also civil engineering) devoted to examining why and how objects fail and to fix the objects and their performance. Structural failures occur in two general modes: static failure, and fatigue failure. *Static structural failure* occurs when, upon being loaded (having a force applied) the object being analyzed either breaks or is deformed plastically, depending on the criterion for failure. *Fatigue failure* occurs when an object fails after a number of repeated loading and unloading cycles. Fatigue failure occurs because of imperfections in the object: a microscopic crack on the surface of the object, for instance, will grow slightly with each cycle (propagation) until the crack is large enough to cause ultimate failure.

Failure is not simply defined as when a part breaks, however; it is defined as when a part does not operate as intended. Some systems, such as the perforated top sections of some plastic bags, are designed to break. If these systems do not break, failure analysis might be employed to determine the cause.

Structural analysis is often used by mechanical engineers after a failure has occurred, or when designing to prevent failure. Engineers often use online documents and books such as those published by ASM to aid them in determining the type of failure and possible causes.

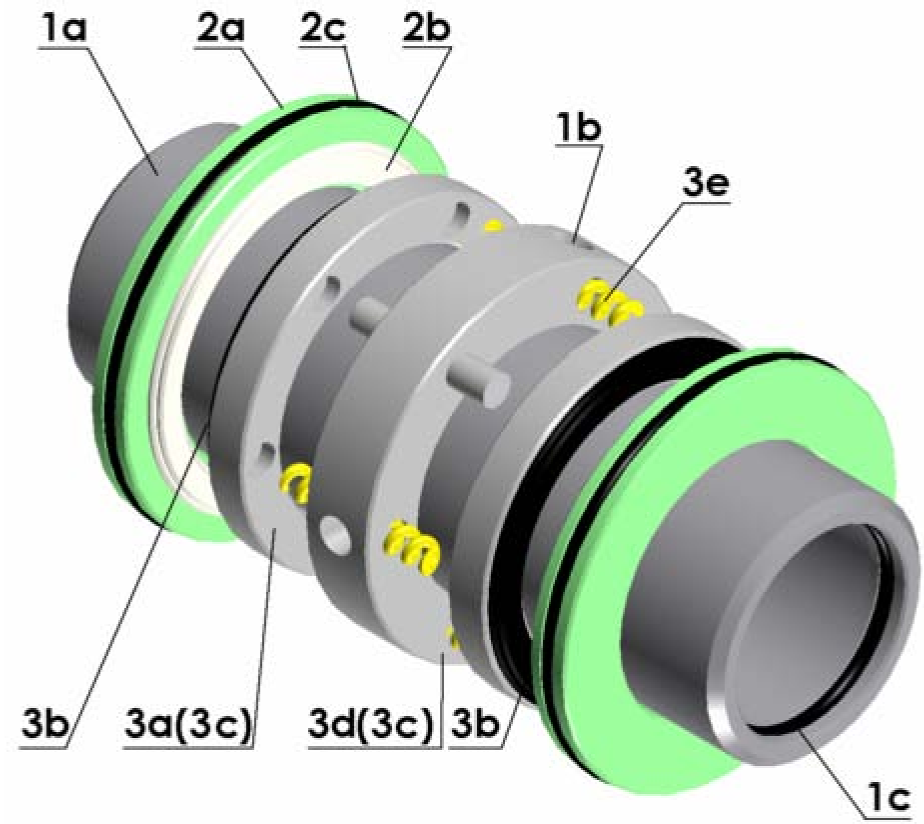
Structural analysis may be used in the office when designing parts, in the field to analyze failed parts, or in laboratories where parts might undergo controlled failure tests.

Thermodynamics and thermo-science

Thermodynamics is an applied science used in several branches of engineering, including mechanical and chemical engineering. At its simplest, thermodynamics is the study of energy, its use and transformation through a system. Typically, engineering thermodynamics is concerned with changing energy from one form to another. As an example, automotive engines convert chemical energy (enthalpy) from the fuel into heat, and then into mechanical work that eventually turns the wheels.

Thermodynamics principles are used by mechanical engineers in the fields of heat transfer, thermofluids, and energy conversion. Mechanical engineers use thermo-science to design engines and power plants, heating, ventilation, and air-conditioning (HVAC) systems, heat exchangers, heat sinks, radiators, refrigeration, insulation, and others.

Drafting



A CAD model of a mechanical double seal

Drafting or technical drawing is the means by which mechanical engineers create instructions for manufacturing parts. A technical drawing can be a computer model or hand-drawn schematic showing all the dimensions necessary to manufacture a part, as well as assembly notes, a list of required materials, and other pertinent information. A U.S. mechanical engineer or skilled worker who creates technical drawings may be referred to as a drafter or draftsman. Drafting has historically been a two-dimensional process, but computer-aided design (CAD) programs now allow the designer to create in three dimensions.

Instructions for manufacturing a part must be fed to the necessary machinery, either manually, through programmed instructions, or through the use of a computer-aided manufacturing (CAM) or combined CAD/CAM program. Optionally, an engineer may also manually manufacture a part using the technical drawings, but this is becoming an increasing rarity, with the advent of computer numerically controlled (CNC) manufacturing. Engineers primarily manually manufacture parts in the areas of applied spray coatings, finishes, and other processes that cannot economically or practically be done by a machine.

Drafting is used in nearly every subdiscipline of mechanical engineering, and by many other branches of engineering and architecture. Three-dimensional models created using

CAD software are also commonly used in finite element analysis (FEA) and computational fluid dynamics (CFD).

Frontiers of research

Mechanical engineers are constantly pushing the boundaries of what is physically possible in order to produce safer, cheaper, and more efficient machines and mechanical systems. Some technologies at the cutting edge of mechanical engineering are listed below.

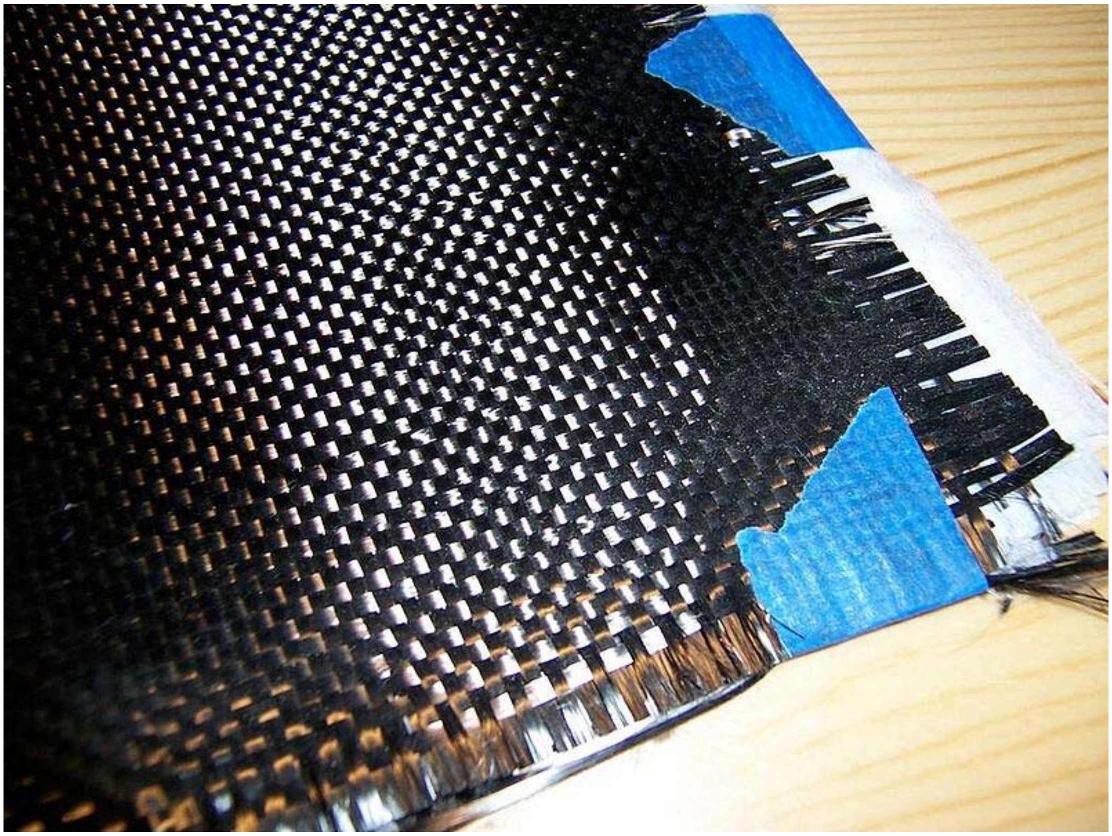
Micro electro-mechanical systems (MEMS)

Micron-scale mechanical components such as springs, gears, fluidic and heat transfer devices are fabricated from a variety of substrate materials such as silicon, glass and polymers like SU8. Examples of MEMS components will be the accelerometers that are used as car airbag sensors, modern cell phones, gyroscopes for precise positioning and microfluidic devices used in biomedical applications.

Friction stir welding (FSW)

Friction stir welding, a new type of welding, was discovered in 1991 by The Welding Institute (TWI). This innovative steady state (non-fusion) welding technique joins materials previously un-weldable, including several aluminum alloys. It may play an important role in the future construction of airplanes, potentially replacing rivets. Current uses of this technology to date include welding the seams of the aluminum main Space Shuttle external tank, Orion Crew Vehicle test article, Boeing Delta II and Delta IV Expendable Launch Vehicles and the SpaceX Falcon 1 rocket, armor plating for amphibious assault ships, and welding the wings and fuselage panels of the new Eclipse 500 aircraft from Eclipse Aviation among an increasingly growing pool of uses.

Composites



Composite cloth consisting of woven carbon fiber.

Composites or composite materials are a combination of materials which provide different physical characteristics than either material separately. Composite material research within mechanical engineering typically focuses on designing (and, subsequently, finding applications for) stronger or more rigid materials while attempting to reduce weight, susceptibility to corrosion, and other undesirable factors. Carbon fiber reinforced composites, for instance, have been used in such diverse applications as spacecraft and fishing rods.

Mechatronics

Mechatronics is the synergistic combination of mechanical engineering, Electronic Engineering, and software engineering. The purpose of this interdisciplinary engineering field is the study of automation from an engineering perspective and serves the purposes of controlling advanced hybrid systems.

Nanotechnology

At the smallest scales, mechanical engineering becomes nanotechnology —one speculative goal of which is to create a molecular assembler to build molecules and materials via mechanosynthesis. For now that goal remains within exploratory engineering.

Finite element analysis

This field is not new, as the basis of Finite Element Analysis (FEA) or Finite Element Method (FEM) dates back to 1941. But evolution of computers has made FEM a viable option for analysis of structural problems. Many commercial codes such as ANSYS, Nastran and ABAQUS are widely used in industry for research and design of components.

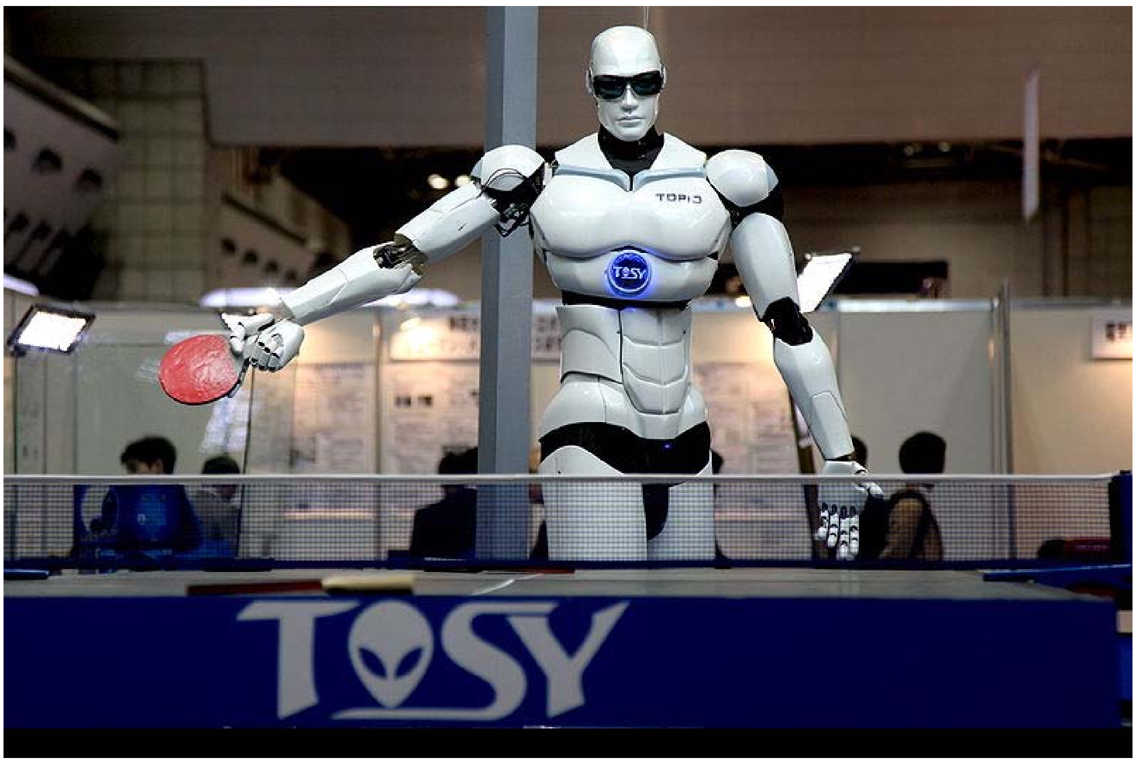
Other techniques such as finite difference method (FDM) and finite-volume method (FVM) are employed to solve problems relating heat and mass transfer, fluid flows, fluid surface interaction etc.

Related fields

Manufacturing engineering and Aerospace Engineering are sometimes grouped with mechanical engineering. A bachelor's degree in these areas will typically have a difference of a few specialized classes.

Chapter 3

Robotics



TOPIO, a humanoid robot, played ping pong at Tokyo International Robot Exhibition (IREX) 2009.

The Shadow robot hand system

A Pick and Place robot in a factory

Robotics is the branch of technology that deals with the design, construction, operation, structural disposition, manufacture and application of robots. Robotics is related to the sciences of electronics, engineering, mechanics, and software. The word "robot" was introduced to the public by Czech writer Karel Čapek in his play R.U.R. (Rossum's Universal Robots), published in 1920. The term "robotics" was coined by Isaac Asimov in his 1941 science fiction short-story "Liar!"

History

Stories of artificial helpers and companions and attempts to create them have a long history.

In 1921, Czech writer Karel Čapek introduced the word "robot" in his play *R.U.R. (Rossum's Universal Robots)*. The word "robot" comes from the word "robota", meaning, in Czech, "forced labour, drudgery".

In 1927, the Maschinenmensch ("machine-human"), a gynoid humanoid robot, also called "Parody", "Futura", "Robotrix", or the "Maria impersonator" (played by German actress Brigitte Helm), the first and perhaps the most memorable depiction of a robot ever to appear on film, was depicted in Fritz Lang's film Metropolis.

In 1942, the science fiction writer Isaac Asimov formulated his Three Laws of Robotics, and in the process of doing so, coined the word "robotics".

In 1948, Norbert Wiener formulated the principles of cybernetics, the basis of practical robotics.

Fully autonomous robots only appeared in the second half of the 20th century. The first digitally operated and programmable robot, the Unimate, was installed in 1961 to lift hot pieces of metal from a die casting machine and stack them. Today, commercial and industrial robots are in widespread use performing jobs more cheaply or more accurately and reliably than humans. They are also employed in jobs which are too dirty, dangerous, or dull to be suitable for humans. Robots are widely used in manufacturing, assembly, and packing; transport; earth and space exploration; surgery; weaponry; laboratory research; safety; and mass production of consumer and industrial goods.

Date	Significance	Robot Name	Inventor
First century A.D. and earlier	Descriptions of more than 100 machines and automata, including a fire engine, a wind organ, a coin-operated machine, and a steam-powered engine, in *Pneumatica* and *Automata* by Heron of Alexandria		Ctesibius, Philo of Byzantium, Heron of Alexandria, and others
1206	Created early humanoid automata, programmable automaton band	Robot band, hand-washing automaton, automated moving peacocks	Al-Jazari
1495	Designs for a humanoid robot	Mechanical knight	Leonardo da Vinci
1738	Mechanical duck that was able to eat, flap its wings, and excrete	Digesting Duck	Jacques de Vaucanson
1898	Nikola Tesla demonstrates first radio-controlled vessel.	Teleautomaton	Nikola Tesla
1921	First fictional automatons called "robots" appear in the play *R.U.R.*	Rossum's Universal Robots	Karel Čapek
1930s	Humanoid robot exhibited at the 1939 and 1940 World's Fairs	Elektro	Westinghouse Electric Corporation
1948	Simple robots exhibiting biological behaviors	Elsie and Elmer	William Grey Walter
1956	First commercial robot, from the Unimation company founded by George Devol and Joseph Engelberger, based on Devol's patents	Unimate	George Devol
1961	First installed industrial robot.	Unimate	George Devol
1963	First palletizing robot	Palletizer	Fuji Yusoki

			Kogyo
1973	First industrial robot with six electromechanically driven axes	Famulus	KUKA Robot Group
1975	Programmable universal manipulation arm, a Unimation product	PUMA	Victor Scheinman

Etymology

According to the *Oxford English Dictionary*, the word *robotics* was first used in print by Isaac Asimov, in his science fiction short story "Liar!", published in May 1941 in *Astounding Science Fiction*. Asimov was unaware that he was coining the term; since the science and technology of electrical devices is *electronics*, he assumed *robotics* already referred to the science and technology of robots. However, in some of Asimov's other works, he states that the first use of the word *robotics* was in his short story *Runaround* (Astounding Science Fiction, March 1942). The word *robotics* was derived from the word *robot*, which was introduced to the public by Czech writer Karel Čapek in his play *R.U.R. (Rossum's Universal Robots)*, which premiered in 1921.

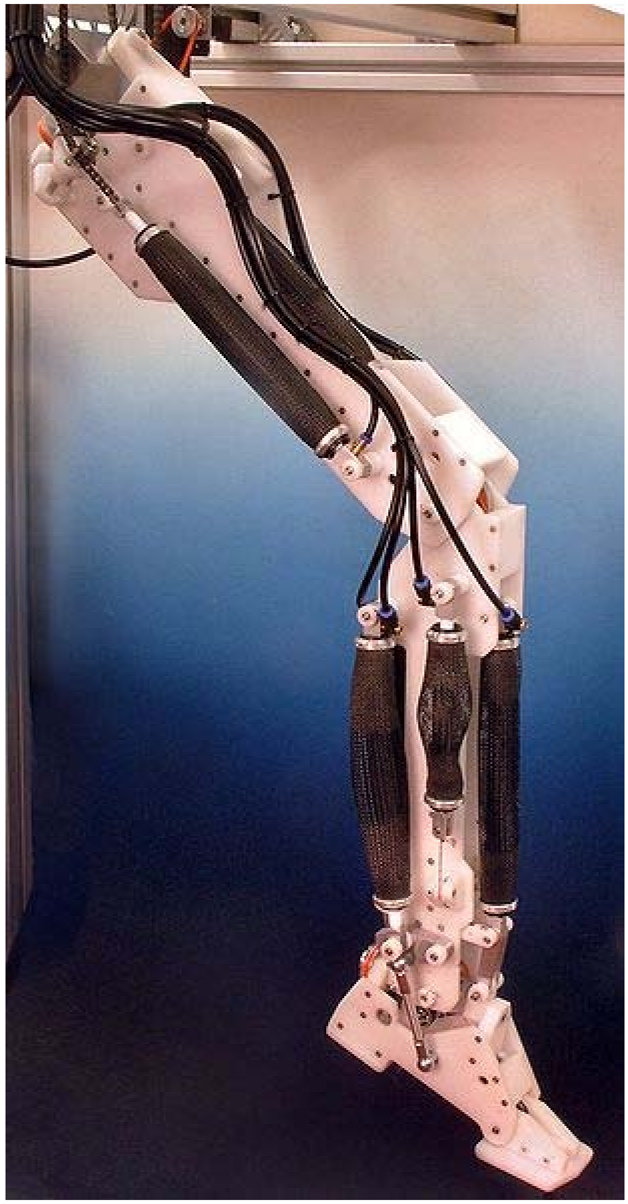
Components

Power source

At present; mostly (lead-acid) batteries are used, but potential power sources could be:

- pneumatic (compressed gases)
- hydraulics (compressed liquids)
- flywheel energy storage
- organic garbage (through anaerobic digestion)
- faeces (human, animal); may be interesting in a military context as faeces of small combat groups may be reused for the energy requirements of the robot assistant
- still unproven energy sources: for example Nuclear fusion, as yet not used in nuclear reactors whereas Nuclear fission is proven (although there are not many robots using it as a power source apart from the Chinese rover tests.).
- radioactive source (such as with the proposed Ford car of the '50s); to those proposed in movies such as *Red Planet*

Actuation



A robotic leg powered by Air Muscles

Actuators are like the "muscles" of a robot, the parts which convert stored energy into movement. By far the most popular actuators are electric motors that spin a wheel or gear, and linear actuators that control industrial robots in factories. But there are some recent advances in alternative types of actuators, powered by electricity, chemicals, or compressed air:

- Electric motors: The vast majority of robots use electric motors, often brushed and brushless DC motors in portable robots or AC motors in industrial robots and CNC machines.

- Linear Actuators: Various types of linear actuators move in and out instead of by spinning, particularly when very large forces are needed such as with industrial robotics. They are typically powered by compressed air (pneumatic actuator) or an oil (hydraulic actuator).
- Series Elastic Actuators: A spring can be designed as part of the motor actuator, to allow improved force control. It has been used in various robots, particularly walking humanoid robots.
- Air muscles: (Also known as Pneumatic Artificial Muscles) are special tubes that contract (typically up to 40%) when air is forced inside it. They have been used for some robot applications.
- Muscle wire: (Also known as Shape Memory Alloy, Nitinol or Flexinol Wire) is a material that contracts slightly (typically under 5%) when electricity runs through it. They have been used for some small robot applications.
- Electroactive Polymers: (EAPs or EPAMs) are a new plastic material that can contract substantially (up to 400%) from electricity, and have been used in facial muscles and arms of humanoid robots, and to allow new robots to float, fly, swim or walk.
- Piezo motor: A recent alternative to DC motors are piezo motors or ultrasonic motors. These work on a fundamentally different principle, whereby tiny piezoceramic elements, vibrating many thousands of times per second, cause linear or rotary motion. There are different mechanisms of operation; one type uses the vibration of the piezo elements to walk the motor in a circle or a straight line. Another type uses the piezo elements to cause a nut to vibrate and drive a screw. The advantages of these motors are nanometer resolution, speed, and available force for their size. These motors are already available commercially, and being used on some robots.
- Elastic nanotubes: These are a promising artificial muscle technology in early-stage experimental development. The absence of defects incarbon nanotubes enables these filaments to deform elastically by several percent, with energy storage levels of perhaps 10 J/cm^3 for metal nanotubes. Human biceps could be replaced with an 8 mm diameter wire of this material. Such compact "muscle" might allow future robots to outrun and outjump humans.

Sensing

Touch

Current robotic and prosthetic hands receive far less tactile information than the human hand. Recent research has developed a tactile sensor array that mimics the mechanical properties and touch receptors of human fingertips. The sensor array is constructed as a rigid core surrounded by conductive fluid contained by an elastomeric skin. Electrodes are mounted on the surface of the rigid core and are connected to an impedance-measuring device within the core. When the artificial skin touches an object the fluid path around the electrodes is deformed, producing impedance changes that map the forces received from the object. The researchers expect that an important function of such artificial fingertips will be adjusting robotic grip on held objects.

Scientists from several European countries and Israel developed a prosthetic hand in 2009, called SmartHand, which functions like a real one—allowing patients to write with it, type on a keyboard, play piano and perform other fine movements. The prosthesis has sensors which enable the patient to sense real feeling in its fingertips.

Vision

Computer vision is the science and technology of machines that see. As a scientific discipline, computer vision is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences and views from cameras.

In most practical computer vision applications, the computers are pre-programmed to solve a particular task, but methods based on learning are now becoming increasingly common.

Computer vision systems rely on image sensors which detect electromagnetic radiation which is typically in the form of either visible light or infra-red light. The sensors are designed using solid-state physics. The process by which light propagates and reflects off surfaces is explained using optics. Sophisticated image sensors even require quantum mechanics to provide a complete understanding of the image formation process.

There is a subfield within computer vision where artificial systems are designed to mimic the processing and behavior of biological systems, at different levels of complexity. Also, some of the learning-based methods developed within computer vision have their background in biology.

Manipulation

Robots which must work in the real world require some way to manipulate objects; pick up, modify, destroy, or otherwise have an effect. Thus the "hands" of a robot are often referred to as *end effectors*, while the "arm" is referred to as a *manipulator*. Most robot arms have replaceable effectors, each allowing them to perform some small range of tasks. Some have a fixed manipulator which cannot be replaced, while a few have one very general purpose manipulator, for example a humanoid hand.

- Mechanical Grippers: One of the most common effectors is the gripper. In its simplest manifestation it consists of just two fingers which can open and close to pick up and let go of a range of small objects. Fingers can for example be made of a chain with a metal wire run through it.
- Vacuum Grippers: Pick and place robots for electronic components and for large objects like car windscreens, will often use very simple vacuum grippers. These are very simple astrictive devices, but can hold very large loads provided the prehension surface is smooth enough to ensure suction.
- General purpose effectors: Some advanced robots are beginning to use fully humanoid hands, like the Shadow Hand, MANUS, and the Schunk hand. These

highly dexterous manipulators, with as many as 20 degrees of freedom and hundreds of tactile sensors.

For the definitive guide to all forms of robot end-effectors, their design, and usage consult the book "Robot Grippers".

Locomotion

Rolling robots

Segway in the Robot museum in Nagoya.

For simplicity most mobile robots have four wheels or a number of continuous tracks. Some researchers have tried to create more complex wheeled robots with only one or two wheels. These can have certain advantages such as greater efficiency and reduced parts, as well as allowing a robot to navigate in confined places that a four wheeled robot would not be able to.

- Two-wheeled balancing: Balancing robots generally use a gyroscope to detect how much a robot is falling and then drive the wheels proportionally in the opposite direction, to counter-balance the fall at hundreds of times per second, based on the dynamics of an inverted pendulum. Many different balancing robots have been designed. While the Segway is not commonly thought of as a robot, it can be thought of as a component of a robot, such as NASA's Robonaut that has been mounted on a Segway.
- One-wheeled balancing: A one-wheeled balancing robot is an extension of a two-wheeled balancing robot so that it can move in any 2D direction using a round ball as its only wheel. Several one-wheeled balancing robots have been designed recently, such as Carnegie Mellon University's "Ballbot" that is the approximate height and width of a person, and Tohoku Gakuin University's "BallIP". Because of the long, thin shape and ability to maneuver in tight spaces, they have the potential to function better than other robots in environments with people.
- Spherical orb robots: Several attempts have been made in robots that are completely inside a spherical ball, either by spinning a weight inside the ball, or by rotating the outer shells of the sphere. These have also been referred to as an orb bot or a ball bot
- Six-wheeled robots: Using six wheels instead of four wheels can give better traction or grip in outdoor terrain such as on rocky dirt or grass.
- Tracked robots: Tank tracks provide even more traction than a six-wheeled robot. Tracked wheels behave as if they were made of hundreds of wheels, therefore are very common for outdoor and military robots, where the robot must drive on very rough terrain. However, they are difficult to use indoors such as on carpets and smooth floors. Examples include NASA's Urban Robot "Urbie".

Walking robots

iCub robot, designed by the RobotCub Consortium

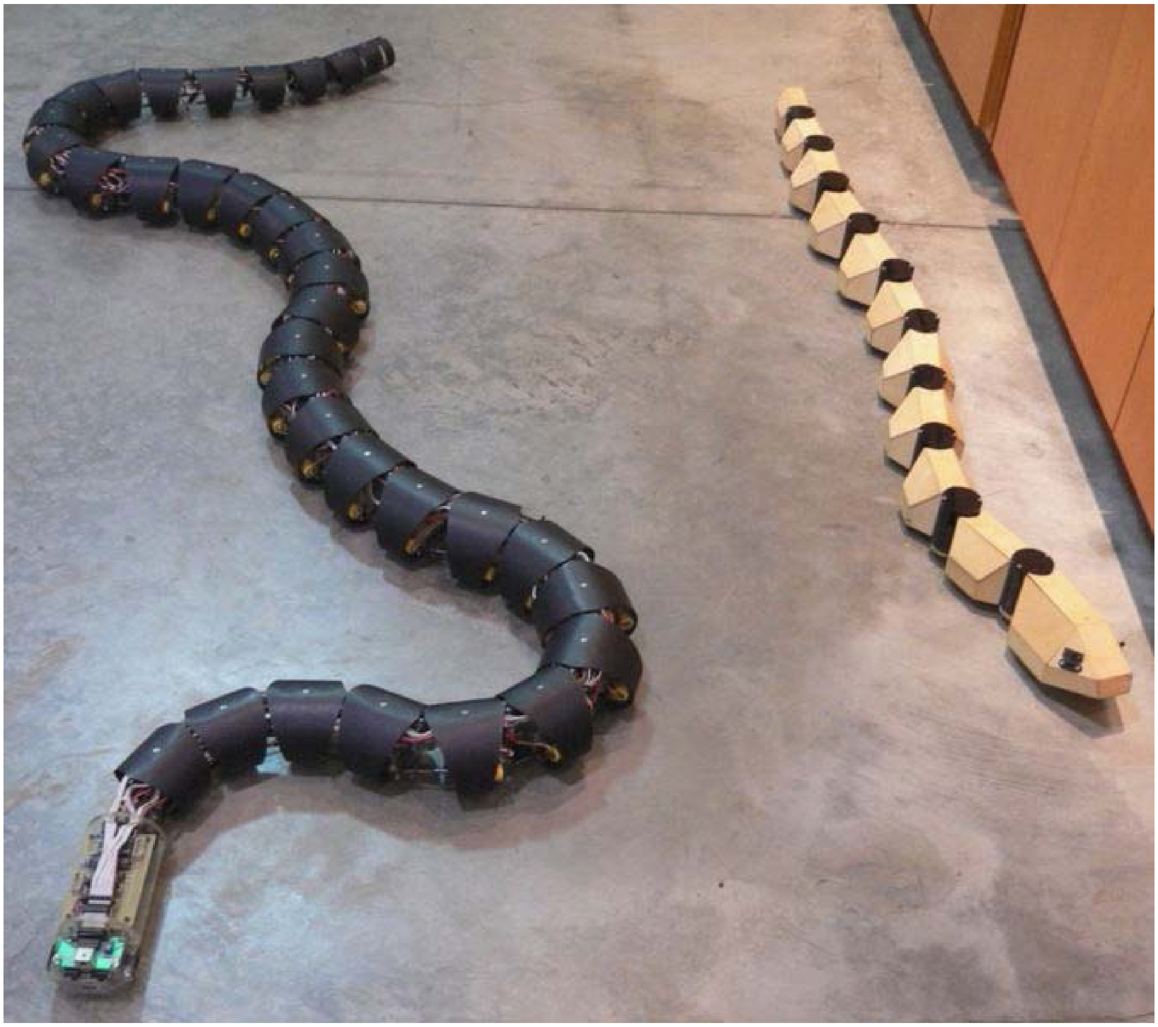
Walking is a difficult and dynamic problem to solve. Several robots have been made which can walk reliably on two legs, however none have yet been made which are as robust as a human. Many other robots have been built that walk on more than two legs, due to these robots being significantly easier to construct. Hybrids too have been proposed in movies such as I, Robot, where they walk on 2 legs and switch to 4 (arms+legs) when going to a sprint. Typically, robots on 2 legs can walk well on flat floors and can occasionally walk up stairs. None can walk over rocky, uneven terrain. Some of the methods which have been tried are:

- ZMP Technique: The Zero Moment Point (ZMP) is the algorithm used by robots such as Honda's ASIMO. The robot's onboard computer tries to keep the total inertial forces (the combination of earth's gravity and the acceleration and deceleration of walking), exactly opposed by the floor reaction force (the force of the floor pushing back on the robot's foot). In this way, the two forces cancel out, leaving no moment (force causing the robot to rotate and fall over). However, this is not exactly how a human walks, and the difference is obvious to human observers, some of whom have pointed out that ASIMO walks as if it needs the lavatory. ASIMO's walking algorithm is not static, and some dynamic balancing is used (see below). However, it still requires a smooth surface to walk on.
- Hopping: Several robots, built in the 1980s by Marc Raibert at the MIT Leg Laboratory, successfully demonstrated very dynamic walking. Initially, a robot with only one leg, and a very small foot, could stay upright simply by hopping. The movement is the same as that of a person on a pogo stick. As the robot falls to one side, it would jump slightly in that direction, in order to catch itself. Soon, the algorithm was generalised to two and four legs. A bipedal robot was demonstrated running and even performing somersaults. A quadruped was also demonstrated which could trot, run, pace, and bound.
- Dynamic Balancing or controlled falling: A more advanced way for a robot to walk is by using a dynamic balancing algorithm, which is potentially more robust than the Zero Moment Point technique, as it constantly monitors the robot's motion, and places the feet in order to maintain stability. This technique was recently demonstrated by Anybots' Dexter Robot, which is so stable, it can even jump. Another example is the TU Delft Flame.
- Passive Dynamics: Perhaps the most promising approach utilizes passive dynamics where the momentum of swinging limbs is used for greater efficiency. It has been shown that totally unpowered humanoid mechanisms can walk down a gentle slope, using only gravity to propel themselves. Using this technique, a robot need only supply a small amount of motor power to walk along a flat surface or a little more to walk up a hill. This technique promises to make walking robots at least ten times more efficient than ZMP walkers, like ASIMO.

Other methods of locomotion

RQ-4 Global Hawk unmanned aerial vehicle

- Flying: A modern passenger airliner is essentially a flying robot, with two humans to manage it. The autopilot can control the plane for each stage of the journey, including takeoff, normal flight, and even landing. Other flying robots are uninhabited, and are known as unmanned aerial vehicles (UAVs). They can be smaller and lighter without a human pilot onboard, and fly into dangerous territory for military surveillance missions. Some can even fire on targets under command. UAVs are also being developed which can fire on targets automatically, without the need for a command from a human. Other flying robots include cruise missiles, the Entomopter, and the Epson micro helicopter robot. Robots such as the Air Penguin, Air Ray, and Air Jelly have lighter-than-air bodies, propelled by paddles, and guided by sonar.



Two robot snakes. Left one has 64 motors (with 2 degrees of freedom per segment), the right one 10.

- Snaking: Several snake robots have been successfully developed. Mimicking the way real snakes move, these robots can navigate very confined spaces, meaning they may one day be used to search for people trapped in collapsed buildings. The Japanese ACM-R5 snake robot can even navigate both on land and in water.
- Skating: A small number of skating robots have been developed, one of which is a multi-mode walking and skating device. It has four legs, with unpowered wheels, which can either step or roll. Another robot, Plen, can use a miniature skateboard or rollerskates, and skate across a desktop.
- Climbing: Several different approaches have been used to develop robots that have the ability to climb vertical surfaces. One approach mimicks the movements of a human climber on a wall with protrusions; adjusting the center of mass and moving each limb in turn to gain leverage. An example of this is Capuchin, built by Stanford University, California. Another approach uses the specialised toe pad method of wall-climbing geckoes, which can run on smooth surfaces such as vertical glass. Examples of this approach include Wallbot and Stickybot. China's

"Technology Daily" November 15, 2008 reported New Concept Aircraft (ZHUHAI) Co., Ltd. Dr. Li Hiu Yeung and his research group have recently successfully developed the bionic gecko robot "Speedy Freelander".According to Dr. Li introduction, this gecko robot can rapidly climbing up and down in a variety of building walls, ground and vertical wall fissure or walking upside down on the ceiling, it is able to adapt on smooth glass, rough or sticky dust walls as well as the various surface of metallic materials and also can automatically identify obstacles, circumvent the bypass and flexible and realistic movements. Its flexibility and speed are comparable to the natural gecko. A third approach is to mimick the motion of a snake climbing a pole.

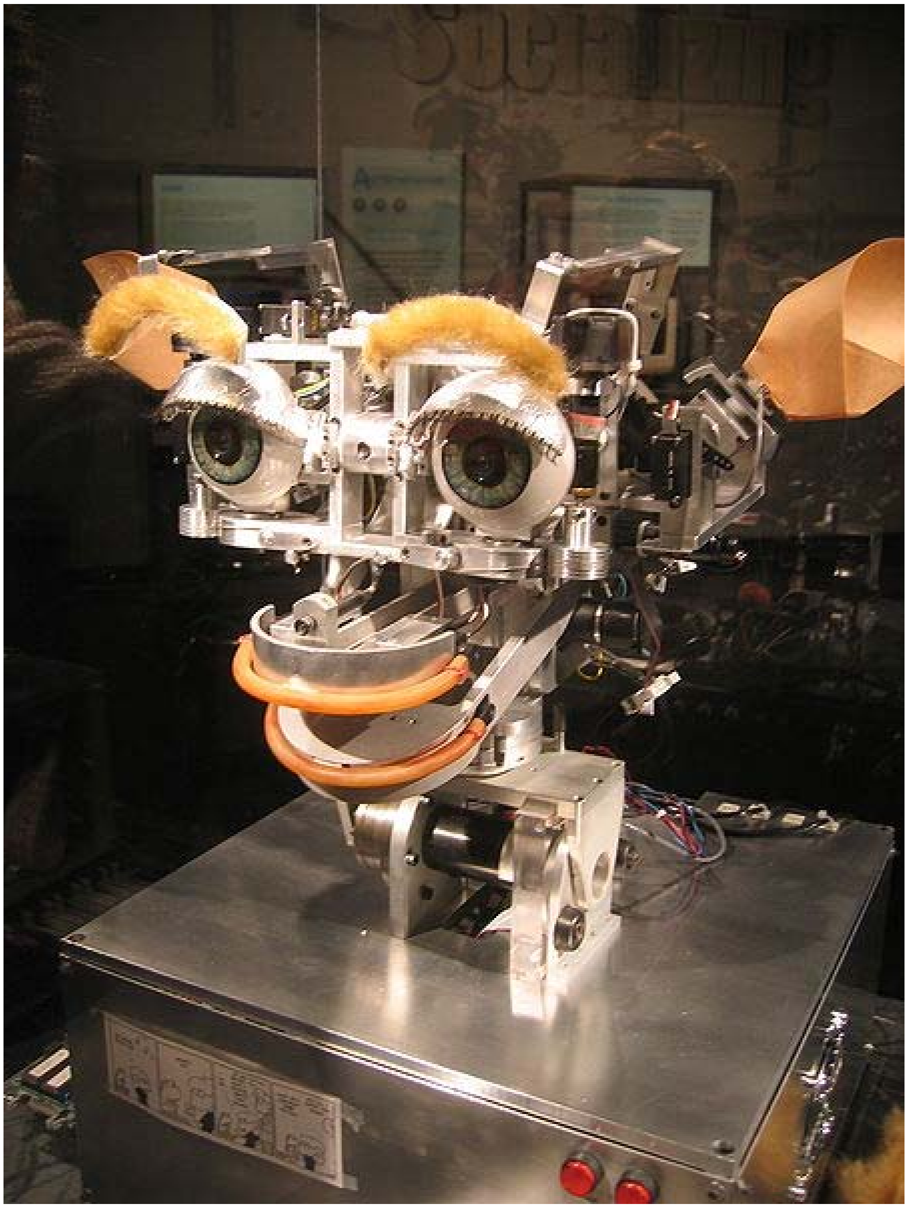
- Swimming: It is calculated that when swimming some fish can achieve a propulsive efficiency greater than 90%. Furthermore, they can accelerate and maneuver far better than any man-made boat or submarine, and produce less noise and water disturbance. Therefore, many researchers studying underwater robots would like to copy this type of locomotion. Notable examples are the Essex University Computer Science Robotic Fish, and the Robot Tuna built by the Institute of Field Robotics, to analyze and mathematically model thunniform motion. The Aqua Penguin, designed and built by Festo of Germany, copies the streamlined shape and propulsion by front "flippers" of penguins. Festo have also built the Aqua Ray and Aqua Jelly, which emulate the locomotion of manta ray, and jellyfish, respectively.

Environmental interaction and navigation

RADAR, GPS, LIDAR, ... are all combined to provide proper navigation and obstacle avoidance

Though a significant percentage of robots in commission today are either human controlled, or operate in a static environment, there is an increasing interest in robots that can operate autonomously in a dynamic environment. These robots require some combination of navigation hardware and software in order to traverse their environment. In particular unforeseen events (e.g. people and other obstacles that are not stationary) can cause problems or collisions. Some highly advanced robots as ASIMO, EveR-1, Meinü robot have particularly good robot navigation hardware and software. Also, self-controlled cars, Ernst Dickmanns' driverless car, and the entries in the DARPA Grand Challenge, are capable of sensing the environment well and subsequently making navigational decisions based on this information. Most of these robots employ a GPS navigation device with waypoints, along with radar, sometimes combined with other sensory data such as LIDAR, video cameras, and inertial guidance systems for better navigation between waypoints.

Human-robot interaction



Kismet can produce a range of facial expressions.

If robots are to work effectively in homes and other non-industrial environments, the way they are instructed to perform their jobs, and especially how they will be told to stop will be of critical importance. The people who interact with them may have little or no training in robotics, and so any interface will need to be extremely intuitive. Science fiction authors also typically assume that robots will eventually be capable of communicating with humans through speech, gestures, and facial expressions, rather than a command-line interface. Although speech would be the most natural way for the human to communicate, it is unnatural for the robot. It will probably be a long time before robots interact as naturally as the fictional C-3PO.

- Speech recognition: Interpreting the continuous flow of sounds coming from a human (speech recognition), in real time, is a difficult task for a computer, mostly because of the great variability of speech. The same word, spoken by the same person may sound different depending on local acoustics, volume, the previous word, whether or not the speaker has a cold, etc.. It becomes even harder when the speaker has a different accent. Nevertheless, great strides have been made in the field since Davis, Biddulph, and Balashek designed the first "voice input system" which recognized "ten digits spoken by a single user with 100% accuracy" in 1952. Currently, the best systems can recognize continuous, natural speech, up to 160 words per minute, with an accuracy of 95%.
- Robotic voice: other hurdles exist when allowing the robot to use voice for interacting with humans. For social reasons, synthetic voice proves suboptimal as a communication medium, making it necessary to develop the emotional component of robotic voice through various techniques.
- Gestures: One can imagine, in the future, explaining to a robot chef how to make a pastry, or asking directions from a robot police officer. In both of these cases, making hand gestures would aid the verbal descriptions. In the first case, the robot would be recognizing gestures made by the human, and perhaps repeating them for confirmation. In the second case, the robot police officer would gesture to indicate "down the road, then turn right". It is likely that gestures will make up a part of the interaction between humans and robots. A great many systems have been developed to recognize human hand gestures.
- Facial expression: Facial expressions can provide rapid feedback on the progress of a dialog between two humans, and soon it may be able to do the same for humans and robots. Robotic faces have been constructed by Hanson Robotics using their elastic polymer called Frubber, allowing a great amount of facial expressions due to the elasticity of the rubber facial coating and imbedded subsurface motors (servos) to produce the facial expressions. The coating and servos are built on a metal skull. A robot should know how to approach a human, judging by their facial expression and body language. Whether the person is happy, frightened, or crazy-looking affects the type of interaction expected of the robot. Likewise, robots like Kismet and the more recent addition, Nexi can produce a range of facial expressions, allowing it to have meaningful social exchanges with humans.
- Artificial emotions: Artificial emotions can also be imbedded and are composed of a sequence of facial expressions and/or gestures. As can be seen from the movie Final Fantasy: The Spirits Within, the programming of these artificial emotions is complex and requires a great amount of human observation. To simplify this programming in the movie, presets were created together with a special software program. This decreased the amount of time needed to make the film. These presets could possibly be transferred for use in real-life robots.
- Personality: Many of the robots of science fiction have a personality, something which may or may not be desirable in the commercial robots of the future. Nevertheless, researchers are trying to create robots which appear to have a

personality: i.e. they use sounds, facial expressions, and body language to try to convey an internal state, which may be joy, sadness, or fear. One commercial example is Pleo, a toy robot dinosaur, which can exhibit several apparent emotions.

Control

A robot-manipulated marionette, with complex control systems

The mechanical structure of a robot must be controlled to perform tasks. The control of a robot involves three distinct phases - perception, processing, and action (robotic paradigms). Sensors give information about the environment or the robot itself (e.g. the position of its joints or its end effector). This information is then processed to calculate the appropriate signals to the actuators (motors) which move the mechanical.

The processing phase can range in complexity. At a reactive level, it may translate raw sensor information directly into actuator commands. Sensor fusion may first be used to estimate parameters of interest (e.g. the position of the robot's gripper) from noisy sensor data. An immediate task (such as moving the gripper in a certain direction) is inferred from these estimates. Techniques from control theory convert the task into commands that drive the actuators.

At longer time scales or with more sophisticated tasks, the robot may need to build and reason with a "cognitive" model. Cognitive models try to represent the robot, the world, and how they interact. Pattern recognition and computer vision can be used to track objects. Mapping techniques can be used to build maps of the world. Finally, motion planning and other artificial intelligence techniques may be used to figure out how to act. For example, a planner may figure out how to achieve a task without hitting obstacles, falling over, etc.

Autonomy levels

Control systems may also have varying levels of autonomy.

1. Direct interaction is used for haptic or tele-operated devices, and the human has nearly complete control over the robot's motion.
2. Operator-assist modes have the operator commanding medium-to-high-level tasks, with the robot automatically figuring out how to achieve them.
3. An autonomous robot may go for extended periods of time without human interaction. Higher levels of autonomy do not necessarily require more complex cognitive capabilities. For example, robots in assembly plants are completely autonomous, but operate in a fixed pattern.

Another classification takes into account the interaction between human control and the machine motions.

1. Teleoperation. A human controls each movement, each machine actuator change is specified by the operator.
2. Supervisory. A human specifies general moves or position changes and the machine decides specific movements of its actuators.
3. Task-level autonomy. The operator specifies only the task and the robot manages itself to complete it.
4. Full autonomy. The machine will create and complete all its tasks without human interaction.

Dynamics and kinematics

The study of motion can be divided into kinematics and dynamics. Direct kinematics refers to the calculation of end effector position, orientation, velocity, and acceleration when the corresponding joint values are known. Inverse kinematics refers to the opposite case in which required joint values are calculated for given end effector values, as done in path planning. Some special aspects of kinematics include handling of redundancy (different possibilities of performing the same movement), collision avoidance, and singularity avoidance. Once all relevant positions, velocities, and accelerations have been calculated using kinematics, methods from the field of dynamics are used to study the effect of forces upon these movements. Direct dynamics refers to the calculation of accelerations in the robot once the applied forces are known. Direct dynamics is used in computer simulations of the robot. Inverse dynamics refers to the calculation of the actuator forces necessary to create a prescribed end effector acceleration. This information can be used to improve the control algorithms of a robot.

In each area mentioned above, researchers strive to develop new concepts and strategies, improve existing ones, and improve the interaction between these areas. To do this, criteria for "optimal" performance and ways to optimize design, structure, and control of robots must be developed and implemented.

Robot research

Much of the research in robotics focuses not on specific industrial tasks, but on investigations into new types of robots, alternative ways to think about or design robots, and new ways to manufacture them but other investigations, such as MIT's cyberflora project, are almost wholly academic.

A first particular new innovation in robot design is the opensourcing of robot-projects. To describe the level of advancement of a robot, the term "Generation Robots" can be used. This term is coined by Professor Hans Moravec, Principal Research Scientist at the Carnegie Mellon University Robotics Institute in describing the near future evolution of robot technology. *First generation* robots, Moravec predicted in 1997, should have an intellectual capacity comparable to perhaps a lizard and should become available by 2010. Because the *first generation* robot would be incapable of learning, however, Moravec predicts that the *second generation* robot would be an improvement over the *first* and become available by 2020, with an intelligence maybe comparable to that of a mouse. The *third generation* robot should have an intelligence comparable to that of a monkey. Though *fourth generation* robots, robots with human intelligence, professor Moravec predicts, would become possible, he does not predict this happening before around 2040 or 2050.

The second is Evolutionary Robots. This is a methodology that uses evolutionary computation to help design robots, especially the body form, or motion and behavior controllers. In a similar way to natural evolution, a large population of robots is allowed to compete in some way, or their ability to perform a task is measured using a fitness function. Those that perform worst are removed from the population, and replaced by a new set, which have new behaviors based on those of the winners. Over time the population improves, and eventually a satisfactory robot may appear. This happens without any direct programming of the robots by the researchers. Researchers use this method both to create better robots, and to explore the nature of evolution. Because the process often requires many generations of robots to be simulated, this technique may be run entirely or mostly in simulation, then tested on real robots once the evolved algorithms are good enough. Currently, there are about 1 million industrial robots toiling around the world, and Japan is the top country having high density of utilizing robots in its manufacturing industry.

Education and training

The SCORBOT-ER 4u - educational robot.

Robots recently became a popular tool in raising interests in computing for middle and high school students. First year computer science courses at several universities were developed which involves the programming of a robot instead of the traditional software engineering based coursework.

Career training

Universities offer Bachelors, Masters and Doctoral degrees in the field of robotics. Select Private Career Colleges and vocational schools offer robotics training to train individuals towards being job ready and employable in the emerging robotics industry.

Certification

The Robotics Certification Standards Alliance (RCSA) is an international robotics certification authority who confers various industry and educational related robotics certifications.

Employment in robotics

A robot technician builds small all-terrain robots.

Robotics is an essential component in any modern manufacturing environment. As factories increase their use of robots, the number of robotics related jobs grow and have been observed to be on a steady rise.

Relationship to unemployment

Some analysts, such as Martin Ford, argue that robots and other forms of automation will ultimately result in significant unemployment as machines begin to match and exceed the capability of workers to perform most jobs. At present the negative impact is only on menial and repetitive jobs, and there is actually a positive impact on the number of jobs for highly skilled technicians, engineers, and specialists. However, these highly skilled jobs are not sufficient in number to offset the greater decrease in employment among the general population, causing structural unemployment in which overall (net) unemployment rises.

As robotics and artificial intelligence develop further, some worry even many skilled jobs may be threatened. In conventional economic theory this should merely cause an increase in the productivity of the involved industries, resulting in higher demand for other goods, and hence higher labour demand in these sectors, off-setting whatever negatives are caused. Conventional theory describes the past well but may not describe the future due to shifts in the parameter values that shape the context.

Chapter 4

Automotive Engineering

Modern **automotive engineering** is a branch of vehicle engineering, incorporating elements of mechanical, electrical, electronic, software and safety engineering as applied to the design, manufacture and operation of motorcycles, automobiles, buses and trucks and their respective engineering subsystems.

Product Engineering

Some of the engineering attributes/disciplines that are of importance to the automotive engineer:

Safety Engineering: Safety Engineering is the assessment of various crash scenarios and their impact on the vehicle occupants. These are tested against very stringent governmental regulations. Some of these requirements include: Seat belt and air bag functionality. Front and side crash worthiness. Resistance to rollover. Assessments are done with various methods and tools: Computer crash simulation, crash test dummies, partial system sled and full vehicle crashes.

Fuel Economy/Emissions: Fuel economy is the measured fuel efficiency of the vehicle in miles per gallon or litres per 100 kilometers. Emissions testing the measurement of the vehicles emissions: hydrocarbons, nitrogen oxides (NOx), carbon monoxide (CO), carbon dioxide (CO2), and evaporative emissions.

Vehicle Dynamics: Vehicle dynamics is the vehicle's response of the following attributes: ride, handling, steering, braking, and traction. Design of the chassis systems of suspension, steering, braking, structure (frame), wheels and tires, and traction control are highly leveraged by the Vehicle Dynamics engineer to deliver the Vehicle Dynamics qualities desired.

NVH Engineering (Noise, Vibration, and Harshness): NVH is the customer's impression both tactile (feel) and audible (hear) feedback from the vehicle. While sound can be interpreted as a rattle, squeal, or hoot, a tactile response can be seat vibration, or a buzz in the steering wheel. This feedback is generated by components either rubbing, vibrating or rotating. NVH response can be classified in various ways: powertrain NVH, road noise, wind noise, component noise, and squeak and rattle. Note, there are both good and bad NVH qualities. The NVH engineer works to either eliminate bad NVH, or change the "bad NVH" to good (i.e., exhaust tones).

Performance: Performance is a measurable and testable value of a vehicles ability to perform in various conditions. Performance can be considered in a wide variety of tasks, but it's generally associated with how quickly a car can accelerate (i.e. 0-60 mph, 1/4 mile, trap speed, top speed, etc), how short and quickly a car can come to a complete stop from a set distance (i.e. 70-0 mph), how many g-forces a car can generate without losing grip, figure 8, recorded trap lap times, cornering speed, brake fade, etc. Performance can also reflect the amount of control in inclement weather (snow, ice, rain).

Shift Quality: Shift Quality is the driver's perception of the vehicle to an automatic transmission banana event. This is influenced by the powertrain (engine, transmission), and the vehicle (driveline, suspension, etc). Shift feel is both a tactile (feel) and audible (hear) response of the vehicle. Shift Quality is experienced as various events: Transmission shifts are felt as an upshift at acceleration (1-2), or a downshift maneuver in passing (4-2). Shift engagements of the vehicle are also evaluated, as in Park to Reverse, etc.

Durability / Corrosion engineering: Durability and Corrosion engineering is the evaluation testing of a vehicle for its useful life. This includes mileage accumulation, severe driving conditions, and corrosive salt baths.

Package / Ergonomics Engineering: Package Engineering is a discipline that designs/analyzes the occupant accommodations (seat roominess), ingress/egress to the vehicle, and the driver's field of vision (gauges and windows). The Package Engineer is also responsible for other areas of the vehicle like the engine compartment, and the component to component placement. Ergonomics is the discipline that assesses the occupant's access to the steering wheel, pedals, and other driver/passenger controls.

Climate Control: Climate Control is the customer's impression of the cabin environment and level of comfort related to the temperature and humidity. From the windshield defrosting, to the heating and cooling capacity, all vehicle seating positions are evaluated to a certain level of comfort.

Drivability: Drivability is the vehicle's response to general driving conditions. Cold starts and stalls, rpm dips, idle response, launch hesitations and stumbles, and performance levels.

Cost: The cost of a vehicle program is typically split into the effect on the variable cost of the vehicle, and the up-front tooling and fixed costs associated with developing the vehicle. There are also costs associated with warranty reductions, and marketing.

Program timing: To some extent programs are timed with respect to the market, and also to the production schedules of the assembly plants. Any new part in the design must support the development and manufacturing schedule of the model.

Assembly Feasibility: It is easy to design a module that is hard to assemble, either resulting in damaged units, or poor tolerances. The skilled product development engineer works with the assembly/manufacturing engineers so that the resulting design is easy and cheap to make and assemble, as well as delivering appropriate functionality and appearance.

Development Engineer

A Development Engineer is a job function within Automotive Engineering, in which the development engineer has the responsibility for coordinating delivery of the engineering attributes of a complete automobile (bus, car, truck, van, SUV, etc.) as dictated by the automobile manufacturer, governmental regulations, and the customer who buys the product.

Much like the Systems Engineer, the Development Engineer is concerned with the interactions of all systems in the complete automobile. While there are multiple components and systems in an automobile that have to function as designed, they must also work in harmony with the complete automobile. As an example, the brake system's main function is to provide braking functionality to the automobile. Along with this, it must also provide an acceptable level of: pedal feel (spongy, stiff), brake system “noise” (squeal, shudder, etc), and interaction with the ABS (anti-lock braking system)

Another aspect of the development engineer's job is a trade-off process required to deliver all the automobile attributes at a certain acceptable level. An example of this is the trade-off between engine performance and fuel economy. While some customers are looking for maximum power from their engine, the automobile is still required to deliver an acceptable level of fuel economy. From the engine's perspective, these are opposing requirements. Engine performance is looking for maximum displacement (bigger, more power), while fuel economy is looking for a smaller displacement engine (ex: 1.4 L vs. 5.4 L). The engine size, though is not the only contributing factor to fuel economy and automobile performance. Other attributes include: automobile weight, aerodynamic drag, transmission gearing, emission control devices, and tires.

The Development Engineer is also responsible for organising automobile level testing, validation, and certification. Components and systems are designed and tested individually by the Product Engineer. The final evaluation though, has to be conducted at the automobile level to evaluate system to system interactions. As an example, the audio system (radio) needs to be evaluated at the automobile level. Interaction with other electronic components can cause interference. Heat dissipation of the system and ergonomic placement of the controls need to be evaluated. Sound quality in all seating positions needs to be provided at acceptable levels.

Other automotive engineering roles

There are also other automotive engineers:

- The aerodynamics engineers will often give guidance to the styling studio so that the shapes they design are aerodynamic, as well as attractive.
- Body engineers will also let the studio know if it is feasible to make the panels for their designs.

Chapter 5

Servomechanism and Electromechanics

Servomechanism

Iraapptyrubn

The grey/green cylinder is the brush-type DC motor. The black section at the bottom contains the planetary reduction gear, and the black object atop the motor is the optical rotary encoder for position feedback. This is the steering actuator of a large robot vehicle.

A **servomechanism**, or **servo**, is an automatic device that uses error-sensing negative feedback to correct the performance of a mechanism. The term correctly applies only to systems where the feedback or error-correction signals help control mechanical position or other parameters. For example, an automotive power window control is not a servomechanism, as there is no automatic feedback that controls position—the operator does this by observation. By contrast the car's cruise control uses closed loop feedback, which classifies it as a servomechanism.

A servomechanism may or may not use a servomotor. For example, a household furnace controlled by a thermostat is a servomechanism, yet there is no motor being controlled directly by the servomechanism.

A common type of servo provides *position control*. Servos are commonly electrical or partially electronic in nature, using an electric motor as the primary means of creating mechanical force. Other types of servos use hydraulics, pneumatics, or magnetic principles. Servos operate on the principle of negative feedback, where the control input is compared to the actual position of the mechanical system as measured by some sort of transducer at the output. Any difference between the actual and wanted values (an "error signal") is amplified and used to drive the system in the direction necessary to reduce or eliminate the error. This procedure is one widely used application of control theory.

Speed control via a governor is another type of servomechanism. The steam engine uses mechanical governors; another early application was to govern the speed of water wheels. Prior to World War II the constant speed propeller was developed to control engine speed for maneuvering aircraft. Fuel controls for gas turbine engines employ either hydromechanical or electronic governing.

Positioning servomechanisms were first used in military fire-control and marine navigation equipment. Today servomechanisms are used in automatic machine tools, satellite-tracking antennas, remote control airplanes, automatic navigation systems on boats and planes, and antiaircraft-gun control systems. Other examples are fly-by-wire systems in aircraft which use servos to actuate the aircraft's control surfaces, and radio-controlled models which use RC servos for the same purpose. Many autofocus cameras also use a servomechanism to accurately move the lens, and thus adjust the focus. A modern hard disk drive has a magnetic servo system with sub-micrometre positioning accuracy.

Typical servos give a rotary (angular) output. Linear types are common as well, using a leadscrew or a linear motor to give linear motion.

Another device commonly referred to as a servo is used in automobiles to amplify the steering or braking force applied by the driver. However, these devices are not true servos, but rather mechanical amplifiers.

In industrial machines, servos are used to perform complex motion.

History

James Watt's steam engine governor is generally considered the first powered feedback system. The windmill fantail is an earlier example of automatic control, but since it does not have an amplifier or gain, it is not usually considered a servomechanism.

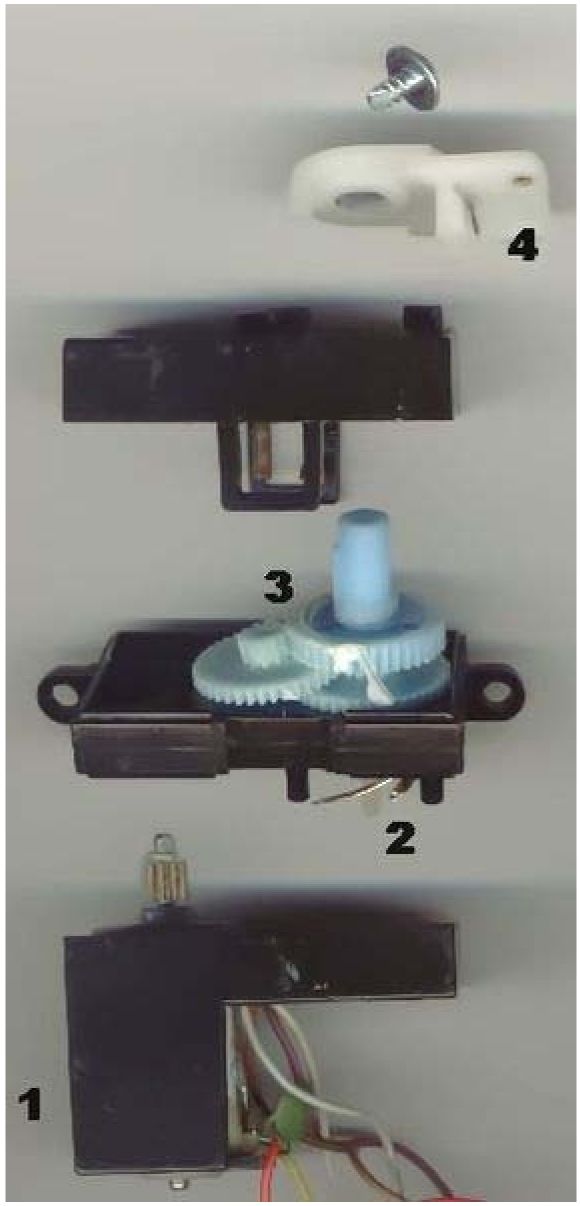
The first feedback position control device was the ship steering engine, used to position the rudder of large ships based on the position of the ship's wheel. This technology was first used on the SS Great Eastern in 1866. Steam steering engines had the characteristics of a modern servomechanism: an input, an output, an error signal, and a means for amplifying the error signal used for negative feedback to drive the error towards zero. The Ragonnet power reverse mechanism was a general purpose air or steam-powered servo amplifier for linear motion patented in 1909.

Electrical servomechanisms require a power amplifier. World War II saw the development of electrical fire-control servomechanisms, using an amplidyne as the power amplifier. Vacuum tube amplifiers were used in the UNISERVO tape drive for the UNIVAC I computer. The Royal Navy began experimenting with Remote Power Control (RPC) on HMS Champion in 1928 and began using RPC to control searchlights in the early 1930s. During WW2 RPC was used to control gun mounts and gun directors.

Modern servomechanisms use solid state power amplifiers, usually built from MOSFET or thyristor devices. Small servos may use power transistors.

The origin of the word is believed to come from the French “Le Servomoteur” or the slavemotor, first used by J. J. L. Farcot in 1868 to describe hydraulic and steam engines for use in ship steering.

RC servos



Small R/C servo mechanism
1. electric motor
2. position feedback potentiometer
3. reduction gear
4. actuator arm

RC servos are hobbyist remote control devices servos typically employed in radio-controlled models, where they are used to provide actuation for various mechanical systems such as the steering of a car, the control surfaces on a plane, or the rudder of a boat.

RC servos are composed of an electric motor mechanically linked to a potentiometer. A standard RC receiver sends Pulse-width modulation (PWM) signals to the servo. The electronics inside the servo translate the width of the pulse into a position. When the servo is commanded to rotate, the motor is powered until the potentiometer reaches the value corresponding to the commanded position.

Due to their affordability, reliability, and simplicity of control by microprocessors, RC servos are often used in small-scale robotics applications.

The servo is usually controlled by three wires: ground, power, and control. The servo will move based on the pulses sent over the control wire, which set the angle of the actuator arm. The servo expects a pulse every 20 ms in order to gain correct information about the angle. The width of the servo pulse dictates the range of the servo's angular motion.

A servo pulse of 1.5 ms width will typically set the servo to its "neutral" position or 45°, a pulse of 1.25 ms could set it to 0° and a pulse of 1.75 ms to 90°. The physical limits and timings of the servo hardware varies between brands and models, but a general servo's angular motion will travel somewhere in the range of 90° - 120° and the neutral position is almost always at 1.5 ms. This is the "standard pulse servo mode" used by all hobby analog servos.

A hobby digital servo is controlled by the same "standard pulse servo mode" pulses as an analog servo. Some hobby digital servos can be set to another mode that allows a robot controller to read back the actual position of the servo shaft. Some hobby digital servos can optionally be set to another mode and "programmed", so it has the desired PID controller characteristics when it is later driven by a standard RC receiver.

RC servos are usually powered by the receiver which in turn is powered by battery packs or an Electronic speed controller (ESC) with an integrated or a separate Battery eliminator circuit (BEC). Common battery packs are either NiCd, NiMH or lithium-ion polymer battery (LiPo) type. Voltage ratings vary, but most receivers are operated at 5 V or 6 V.

Electromechanics

In engineering, **electromechanics** combines the sciences of electromagnetism of electrical engineering and mechanics. Mechanical engineering in this context refers to the larger discipline which includes chemical engineering, and other related disciplines. Electrical engineering in this context also encompasses software engineering, computer engineering, and other related fields. This refers to the three major engineering disciplines of electrical engineering, mechanical engineering and civil engineering under which all other engineering disciples are classified.

History

Relays originated with telegraphy as electromechanical devices used to regenerate telegraph signals.

The Strowger switch, Panel switch and similar ones were widely used in early automated telephone exchanges. Crossbar switches were first widely installed in the middle 20th century in Sweden, the United States and Britain, and quickly spread to the rest of the world. The electromechanical television systems of the late 19th century were less successful.

Electric typewriters developed, up to the 1980s, as "power-assisted typewriters". They contained a single electrical component in them, the motor. Where the keystroke had previously moved a typebar directly, now it engaged mechanical linkages that directed mechanical power from the motor into the typebar. This was also true of the forthcoming IBM Selectric. At Bell Labs, in the 1940s, the Bell Model V computer was developed. It was an electromechanical relay-based monster with cycle times in seconds. In 1968 Garrett AiResearch was invited to produce a digital computer to compete with electromechanical systems then under development for the main flight control computer in the US Navy's new F-14 Tomcat fighter. Garrett created what is arguably the first complete microprocessor when it came up with the Central Air Data Computer for this project.

Modern practice

Common items, which in the 20th century would have used electromechanical devices for control, today use a less expensive and more effective standard integrated microcontroller circuit containing a few million transistors and a computer program to carry out the same task through logic. Such chips have replaced most electromechanical devices, are used in most simple feedback control systems, and appear in huge numbers in everything from traffic lights to washing machines, though the latter case, as others where mechanical motion is involved, requires that triacs control mechanical electric actuators.

Chapter 6

Anti-Lock Braking System

An **anti-lock braking system** (**ABS**) is a safety system that allows the wheels on a motor vehicle to continue interacting tractively with the road surface as directed by driver steering inputs while braking, preventing the wheels from locking up (that is, ceasing rotation) and therefore avoiding skidding.

An ABS generally offers improved vehicle control and decreases stopping distances on dry and slippery surfaces for many drivers; however, on loose surfaces like gravel or snow-covered pavement, an ABS can significantly increase braking distance, although still improving vehicle control.

Since initial widespread use in production cars, anti-lock braking systems have evolved considerably. Recent versions not only prevent wheel lock under braking, but also electronically control the front-to-rear brake bias. This function, depending on its specific capabilities and implementation, is known as electronic brakeforce distribution (EBD), traction control system, emergency brake assist, or electronic stability control (ESC).

History

Early systems

The ABS was first developed for aircraft use in 1929 by the French automobile and aircraft pioneer, Gabriel Voisin, as threshold braking on airplanes is nearly impossible. An early system was Dunlop's Maxaret system, which was introduced in the 1950s and is still in use on some aircraft models. These sistems use a flywheel and valve attached to a hydraulic line that feeds the brake cylinders. The flywheel is attached to a drum that runs at the same speed as the wheel. In normal braking, the drum and flywheel should spin at the same speed. However, if a wheel were to slow down, then the drum would do the same, leaving the flywheel spinning at a faster rate. This causes the valve to open, allowing a small amount of brake fluid to bypass the master cylinder into a local reservoir, lowering the pressure on the cylinder and releasing the brakes. The use of the drum and flywheel meant the valve only opened when the wheel was turning. In testing, a 30% improvement in braking performance was noted, because the pilots immediately applied full brakes instead of slowly increasing pressure in order to find the skid point. An additional benefit was the elimination of burned or burst tires.

In 1958, a Royal Enfield Super Meteor motorcycle was used by the Road Research Laboratory to test the Maxaret anti-lock brake. The experiments demonstrated that anti-lock brakes can be of great value to motorcycles, for which skidding is involved in a high proportion of accidents. Stopping distances were reduced in most of the tests compared with locked wheel braking, particularly on slippery surfaces, in which the improvement could be as much as 30 percent. Enfield's technical director at the time, Tony Wilson-Jones, saw little future in the system, however, and it was not put into production by the company.

A fully mechanical system saw limited automobile use in the 1960s in the Ferguson P99 racing car, the Jensen FF, and the experimental all wheel drive Ford Zodiac, but saw no further use; the system proved expensive and unreliable in automobile use.

Modern systems

Chrysler, together with the Bendix Corporation, introduced a computerized, three-channel, four-sensor all-wheel ABS called "Sure Brake" for its 1971 Imperial. It was available for several years thereafter, functioned as intended, and proved reliable. In 1971, General Motors introduced the "Trackmaster" rear-wheel only ABS as an option on their Rear-wheel drive Cadillac models. In the same year, Nissan offered an EAL (Electro Anti-lock System) as an option on the Nissan President, which became Japan's first electronic ABS.

ABS brakes on a BMW motorcycle

In 1988, BMW introduced the first motorcycle with an electronic-hydraulic ABS: the BMW K100. Honda followed suit in 1992 with the launch of its first motorcycle ABS on the ST1100 Pan European. In 2007, Suzuki launched its GSF1200SA (Bandit) with an ABS. In 2005, Harley-Davidson began offering ABS as an option for police bikes. In 2008, ABS became a factory-installed option on all Harley-Davidson Touring motorcycles and standard equipment on select models.

Operation

The anti-lock brake controller is also known as the CAB (Controller Anti-lock Brake).

A typical ABS includes a central electronic control unit (ECU), four wheel speed sensors, and at least two hydraulic valves within the brake hydraulics. The ECU constantly monitors the rotational speed of each wheel; if it detects a wheel rotating significantly slower than the others, a condition indicative of impending wheel lock, it actuates the valves to reduce hydraulic pressure to the brake at the affected wheel, thus reducing the braking force on that wheel; the wheel then turns faster. Conversely, if the ECU detects a wheel turning significantly faster than the others, brake hydraulic pressure to the wheel is increased so the braking force is reapplied, slowing down the wheel. This process is repeated continuously and can be detected by the driver via brake pedal pulsation. Some anti-lock system can apply or release braking pressure 16 times per second.

The ECU is programmed to disregard differences in wheel rotative speed below a critical threshold, because when the car is turning, the two wheels towards the center of the curve turn slower than the outer two. For this same reason, a differential is used in virtually all roadgoing vehicles.

If a fault develops in any part of the ABS, a warning light will usually be illuminated on the vehicle instrument panel, and the ABS will be disabled until the fault is rectified.

The modern ABS applies individual brake pressure to all four wheels through a control system of hub-mounted sensors and a dedicated micro-controller. ABS is offered or comes standard on most road vehicles produced today and is the foundation for ESC systems, which are rapidly increasing in popularity due to the vast reduction in price of vehicle electronics over the years.

Modern electronic stability control (ESC or ESP) systems are an evolution of the ABS concept. Here, a minimum of two additional sensors are added to help the system work: these are a steering wheel angle sensor, and a gyroscopic sensor. The theory of operation is simple: when the gyroscopic sensor detects that the direction taken by the car does not coincide with what the steering wheel sensor reports, the ESC software will brake the necessary individual wheel(s) (up to three with the most sophisticated systems), so that the vehicle goes the way the driver intends. The steering wheel sensor also helps in the operation of Cornering Brake Control (CBC), since this will tell the ABS that wheels on the inside of the curve should brake more than wheels on the outside, and by how much.

The ABS equipment may also be used to implement a traction control system(TCS) on acceleration of the vehicle. If, when accelerating, the tire loses traction, the ABS controller can detect the situation and take suitable action so that traction is regained. More sophisticated versions of this can also control throttle levels and brakes simultaneously.

Components

There are four main components to an ABS: speed sensors, valves, a pump, and a controller.

Speed sensors
: The anti-lock braking system needs some way of knowing when a wheel is about to lock up. The speed sensors, which are located at each wheel, or in some cases in the differential, provide this information.

Valves
: There is a valve in the brake line of each brake controlled by the ABS. On some systems, the valve has three positions:

- In position one, the valve is open; pressure from the master cylinder is passed right through to the brake.
- In position two, the valve blocks the line, isolating that brake from the master cylinder. This prevents the pressure from rising further should the driver push the brake pedal harder.
- In position three, the valve releases some of the pressure from the brake.

Pump
: Since the valve is able to release pressure from the brakes, there has to be some way to put that pressure back. That is what the pump does; when a valve reduces the pressure in a line, the pump is there to get the pressure back up.

Controller
: The controller is an ECU type unit in the car which receives information from each individual wheel speed sensor, in turn if a wheel loses traction the signal is sent to the controller, the controller will then limit the brakeforce (EBD) and activate the ABS modulator which actuates the braking valves on and off.

Use

There are many different variations and control algorithms for use in an ABS. One of the simpler systems works as follows:

1. The controller monitors the speed sensors at all times. It is looking for decelerations in the wheel that are out of the ordinary. Right before a wheel locks up, it will experience a rapid deceleration. If left unchecked, the wheel would stop much more quickly than any car could. It might take a car five seconds to stop

from 60 mph (96.6 km/h) under ideal conditions, but a wheel that locks up could stop spinning in less than a second.

2. The ABS controller knows that such a rapid deceleration is impossible, so it reduces the pressure to that brake until it sees an acceleration, then it increases the pressure until it sees the deceleration again. It can do this very quickly, before the tire can actually significantly change speed. The result is that the tire slows down at the same rate as the car, with the brakes keeping the tires very near the point at which they will start to lock up. This gives the system maximum braking power.
3. When the ABS system is in operation the driver will feel a pulsing in the brake pedal; this comes from the rapid opening and closing of the valves. This pulsing also tells the driver that the ABS has been triggered. Some ABS systems can cycle up to 16 times per second.

Brake types

Anti-lock braking systems use different schemes depending on the type of brakes in use. They can be differentiated by the number of channels: that is, how many valves that are individually controlled—and the number of speed sensors.

Four-channel, four-sensor ABS

This is the best scheme. There is a speed sensor on all four wheels and a separate valve for all four wheels. With this setup, the controller monitors each wheel individually to make sure it is achieving maximum braking force.

Three-channel, three-sensor ABS

This scheme, commonly found on pickup trucks with four-wheel ABS, has a speed sensor and a valve for each of the front wheels, with one valve and one sensor for both rear wheels. The speed sensor for the rear wheels is located in the rear axle. This system provides individual control of the front wheels, so they can both achieve maximum braking force. The rear wheels, however, are monitored together; they both have to start to lock up before the ABS will activate on the rear. With this system, it is possible that one of the rear wheels will lock during a stop, reducing brake effectiveness.

One-channel, one-sensor ABS

This system is commonly found on pickup trucks with rear-wheel ABS. It has one valve, which controls both rear wheels, and one speed sensor, located in the rear axle. This system operates the same as the rear end of a three-channel system. The rear wheels are monitored together and they both have to start to lock up before the ABS kicks in. In this system it is also possible that one of the rear wheels will lock, reducing brake effectiveness. This system is easy to identify. Usually there will be one brake line going through a T-fitting to both rear wheels.

Effectiveness

A 2003 Australian study by Monash University Accident Research Centre found that ABS:

- Reduced the risk of multiple vehicle crashes by 18 percent,
- Reduced the risk of run-off-road crashes by 35 percent.

On high-traction surfaces such as bitumen, or concrete, many (though not all) ABS-equipped cars are able to attain braking distances better (i.e. shorter) than those that would be easily possible without the benefit of ABS. In real world conditions even an alert, skilled driver without ABS would find it difficult, even through the use of techniques like threshold braking, to match or improve on the performance of a typical driver with a modern ABS-equipped vehicle. ABS reduces chances of crashing, and/or the severity of impact. The recommended technique for non-expert drivers in an ABS-equipped car, in a typical full-braking emergency, is to press the brake pedal as firmly as possible and, where appropriate, to steer around obstructions. In such situations, ABS will significantly reduce the chances of a skid and subsequent loss of control.

In gravel, sand and deep snow, ABS tends to increase braking distances. On these surfaces, locked wheels dig in and stop the vehicle more quickly. ABS prevents this from occurring. Some ABS calibrations reduce this problem by slowing the cycling time, thus letting the wheels repeatedly briefly lock and unlock. Some vehicle manufacturers provide an "off-road" button to turn ABS function off. The primary benefit of ABS on such surfaces is to increase the ability of the driver to maintain control of the car rather than go into a skid, though loss of control remains more likely on soft surfaces like gravel or slippery surfaces like snow or ice. On a very slippery surface such as sheet ice or gravel, it is possible to lock multiple wheels at once, and this can defeat ABS (which relies on comparing all four wheels, and detecting individual wheels skidding). Availability of ABS relieves most drivers from learning threshold braking.

A June 1999 National Highway Traffic Safety Administration (NHTSA) study found that ABS increased stopping distances on loose gravel by an average of 22 percent.

According to the NHTSA,

"ABS works with your regular braking system by automatically pumping them. In vehicles not equipped with ABS, the driver has to manually pump the brakes to prevent wheel lockup. In vehicles equipped with ABS, your foot should remain firmly planted on the brake pedal, while ABS pumps the brakes for you so you can concentrate on steering to safety."

When activated, some earlier ABS systems caused the brake pedal to pulse noticeably. As most drivers rarely or never brake hard enough to cause brake lock-up, and a significant number rarely bother to read the car's manual, this may not be discovered until an emergency. When drivers do encounter an emergency that causes them to brake hard, and thus encounter this pulsing for the first time, many are believed to reduce pedal pressure, and thus lengthen braking distances, contributing to a higher level of accidents than the superior emergency stopping capabilities of ABS would otherwise promise. Some manufacturers have therefore implemented a brake assist system that determines that the driver is attempting a "panic stop" (by detecting that the brake pedal was depressed very fast, unlike a normal stop where the pedal pressure would usually be gradually increased, Some systems additionally monitor the rate at the accelerator was released) and the system automatically increases braking force where not enough pressure is applied. Hard or panic braking on bumpy surfaces, because of the bumps causing the speed of the wheel(s) to become erratic may also trigger the ABS. Nevertheless, ABS significantly improves safety and control for drivers in most on-road situations.

Anti-lock brakes are the subject of some experiments centred around risk compensation theory, which asserts that drivers adapt to the safety benefit of ABS by driving more aggressively. In a Munich study, half a fleet of taxicabs was equipped with anti-lock brakes, while the other half had conventional brake systems. The crash rate was substantially the same for both types of cab, and Wilde concludes this was due to drivers of ABS-equipped cabs taking more risks, assuming that ABS would take care of them, while the non-ABS drivers drove more carefully since ABS would not be there to help in case of a dangerous situation. A similar study was carried out in Oslo, with similar results.

Chapter 7

Mechanical Filter

Figure 1. A mechanical filter made by the Kokusai Electric Company intended for selecting the narrow 2 kHz bandwidth signals in SSB radio receivers. It operates at 455 kHz, a common IF for these receivers, and is dimensioned 45×15×15 mm (1 $^3/_4$×$^7/_{12}$×$^7/_{12}$ in).

A **mechanical filter** is a signal processing filter usually used in place of an electronic filter at radio frequencies. Its purpose is the same as that of a normal electronic filter: to pass a range of signal frequencies, but to block others. The filter acts on mechanical vibrations which are the analogue of the electrical signal. At the input and output of the filter there are transducers which convert the electrical signal into, and then back from, these mechanical vibrations.

The components of a mechanical filter are all directly analogous to the various elements found in electrical circuits. The mechanical elements obey mathematical functions which are identical to their corresponding electrical elements. This makes it possible to apply electrical network analysis and filter design methods to mechanical filters. Electrical theory has developed a large library of mathematical forms that produce useful filter frequency responses and the mechanical filter designer is able to make direct use of these. It is only necessary to set the mechanical components to appropriate values to produce a filter with an identical response to the electrical counterpart.

Steel and nickel–iron alloys are common materials for mechanical filter components; nickel is sometimes used for the input and output couplings. Resonators in the filter made from these materials need to be machined to precisely adjust their resonance frequency before final assembly.

While the meaning of *mechanical filter* here is one that is used in an electromechanical role, it is fully possible to use a mechanical design to filter mechanical vibrations or sound waves (which are also essentially mechanical) directly. For example, filtering of audio frequency response in the design of loudspeaker cabinets can be achieved with mechanical components. In the electrical application, in addition to mechanical components which correspond to their electrical counterparts, transducers are needed to convert between the mechanical and electrical domains. There are a wide variety of component forms and topologies for mechanical filters, a representative selection.

The theory of mechanical filters was first applied to improving the mechanical parts of phonographs in the 1920s. By the 1950s mechanical filters were being manufactured as self-contained components for applications in radio transmitters and high-end receivers. The high "quality factor", Q, that mechanical resonators can attain, far higher than that of an all-electrical LC circuit, made possible the construction of mechanical filters with excellent selectivity. Good selectivity, being important in radio receivers, made such filters highly attractive. Contemporary researchers are working on microelectromechanical filters, the mechanical devices corresponding to electronic integrated circuits.

Elements

The elements of a passive linear electrical network consist of inductors, capacitors and resistors which have the properties of inductance, elastance (inverse capacitance) and resistance, respectively. The mechanical counterparts of these properties are, respectively, mass, stiffness and damping. In most electronic filter designs, only inductor and capacitor elements are used in the body of the filter (although the filter may be terminated with resistors at the input and output). Resistances are not present in a theoretical filter composed of ideal components and only arise in practical designs as unwanted parasitic elements. Likewise, a mechanical filter would ideally consist only of components with the properties of mass and stiffness, but in reality some damping is present as well.

The mechanical counterparts of voltage and electric current in this type of analysis are, respectively, force (F) and velocity (v) and represent the signal waveforms. From this, a mechanical impedance can be defined in terms of the imaginary angular frequency, $j\omega$, which entirely follows the electrical analogy.

Mechanical element	Formula (in one dimension)	Mechanical impedance	Electrical counterpart
Stiffness, *S*	$S = \frac{F}{x}$	$Z = \frac{S}{j\omega}$	Elastance, 1/*C*, the inverse of capacitance
Mass, *M*	$M = \frac{F}{dv/dt} = \frac{F}{a}$	$Z = j\omega M$	Inductance, *L*
Damping, *D*	$D = \frac{F}{v}$	$Z = D$	Resistance, *R*

Notes:

- The symbols *x*, *t*, and *a* represent their usual quantities; distance, time, and acceleration respectively.
- The mechanical quantity *compliance*, which is the inverse of stiffness, can be used instead of stiffness to give a more direct correspondence to capacitance, but stiffness is used in the table as the more familiar quantity.

The scheme presented in the table is known as the impedance analogy. Circuit diagrams produced using this analogy match the electrical impedance of the mechanical system seen by the electrical circuit, making it intuitive from an electrical engineering standpoint. There is also the mobility analogy, in which force corresponds to current and velocity corresponds to voltage. This has equally valid results but requires using the reciprocals of the electrical counterparts listed above. Hence, $M \rightarrow C$, $S \rightarrow 1/L$, $D \rightarrow G$ where *G* is electrical conductance, the inverse of resistance. Equivalent circuits produced by this scheme are similar, but are the dual impedance forms whereby series elements become parallel, capacitors become inductors, and so on. Circuit diagrams using the mobility analogy more closely match the mechanical arrangement of the circuit, making it more intuitive from a mechanical engineering standpoint. In addition to their application to electromechanical systems, these analogies are widely used to aid analysis in acoustics.

Any mechanical component will unavoidably possess both mass and stiffness. This translates in electrical terms to an LC circuit, that is, a circuit consisting of an inductor and a capacitor, hence mechanical components are resonators and are often used as such. It is still possible to represent inductors and capacitors as individual lumped elements in a mechanical implementation by minimising (but never quite eliminating) the unwanted property. Capacitors may be made of thin, long rods, that is, the mass is minimised and the compliance is maximised. Inductors, on the other hand, may be made of short, wide pieces which maximise the mass in comparison to the compliance of the piece.

Mechanical parts act as a transmission line for mechanical vibrations. If the wavelength is short in comparison to the part then a lumped element model as described above is no longer adequate and a distributed element model must be used instead. The mechanical distributed elements are entirely analogous to electrical distributed elements and the mechanical filter designer can use the methods of electrical distributed element filter design.

History

Harmonic telegraph

Mechanical filter design was developed by applying the discoveries made in electrical filter theory to mechanics. However, a very early example (1870s) of acoustic filtering was the "harmonic telegraph", which arose precisely because electrical resonance was poorly understood but mechanical resonance (in particular, acoustic resonance) was very familiar to engineers. This situation was not to last for long; electrical resonance had been known to science for some time before this, and it was not long before engineers started to produce all-electric designs for filters. In its time, though, the harmonic telegraph was of some importance. The idea was to combine several telegraph signals on one telegraph line by what would now be called frequency division multiplexing thus saving enormously on line installation costs. The key of each operator activated a vibrating electromechanical reed which converted this vibration into an electrical signal. Filtering at the receiving operator was achieved by a similar reed tuned to precisely the same frequency, which would only vibrate and produce a sound from transmissions by the operator with the identical tuning.

Versions of the harmonic telegraph were developed by Elisha Gray, Alexander Graham Bell, Ernest Mercadier and others. Its ability to act as a sound transducer to and from the electrical domain was to inspire the invention of the telephone.

Mechanical equivalent circuits

Once the basics of electrical network analysis began to be established, it was not long before the ideas of complex impedance and filter design theories were carried over into mechanics by analogy. Kennelly, who was also responsible for introducing complex impedance, and Webster were the first to extend the concept of impedance into mechanical systems in 1920. Mechanical admittance and the associated mobility analogy came much later and are due to Firestone in 1932.

It was not enough to just develop a mechanical analogy. This could be applied to problems that were entirely in the mechanical domain, but for mechanical filters with an electrical application it is necessary to include the transducer in the analogy as well. Poincaré in 1907 was the first to describe a transducer as a pair of linear algebraic equations relating electrical variables (voltage and current) to mechanical variables (force and velocity). These equations can be expressed as a matrix relationship in much the same way as the z-parameters of a two-port network in electrical theory, to which this is entirely analogous:

$$\begin{bmatrix} V \\ F \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I \\ v \end{bmatrix}$$

where V and I represent the voltage and current respectively on the electrical side of the transducer.

Wegel, in 1921, was the first to express these equations in terms of mechanical impedance as well as electrical impedance. The element $z_{22}$is the open circuit mechanical impedance, that is, the impedance presented by the mechanical side of the transducer when no current is entering the electrical side. The element z_{11}, conversely, is the clamped electrical impedance, that is, the impedance presented to the electrical side when the mechanical side is clamped and prevented from moving (velocity is zero). The remaining two elements, $z_{21}$and z_{12}, describe the transducer forward and reverse transfer functions respectively. Once these ideas were in place, engineers were able to extend electrical theory into the mechanical domain and analyse an electromechanical system as a unified whole.

Sound reproduction

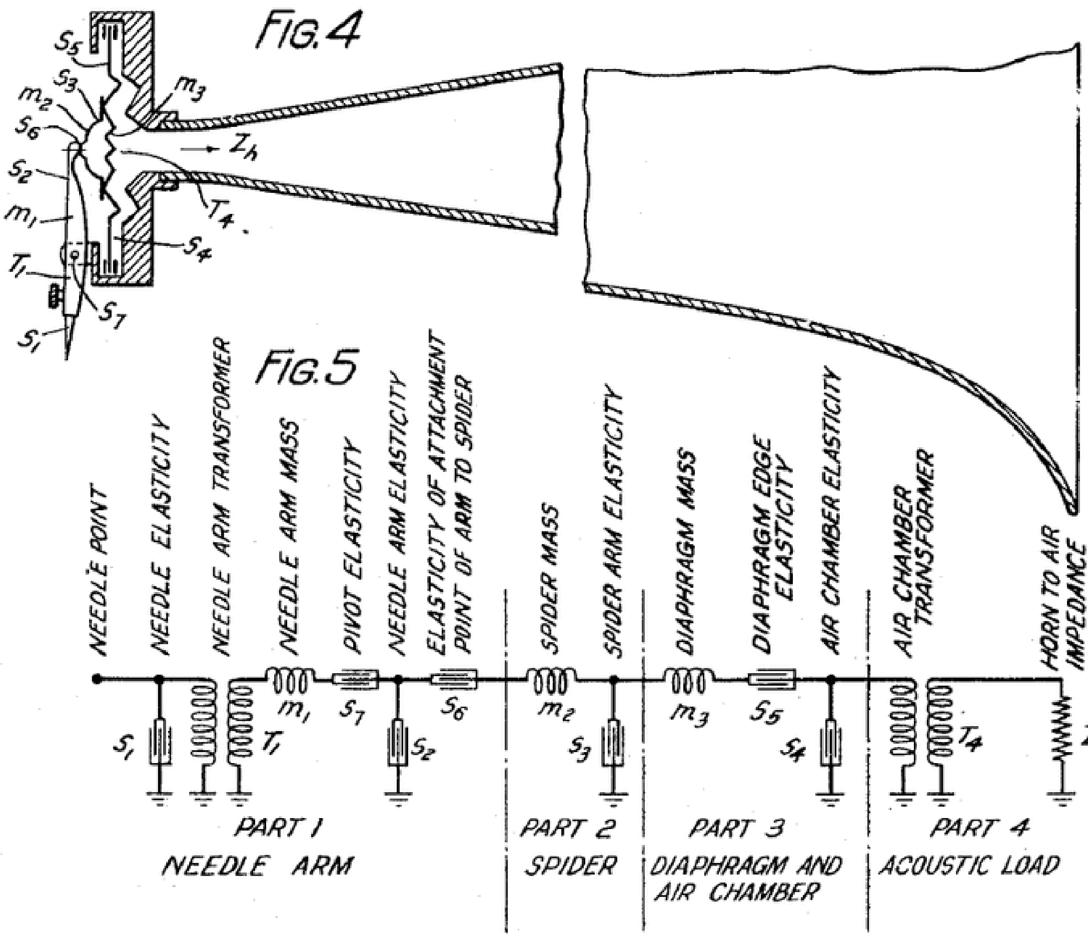


Figure 2. Harrison's phonograph mechanism and its electrical equivalent circuit.

An early application of these new theoretical tools was in phonographic sound reproduction. A recurring problem with early phonograph designs was that mechanical resonances in the pickup and sound transmission mechanism caused excessively large peaks and troughs in the frequency response, resulting in poor sound quality. In 1923, Harrison of the Western Electric Company filed a patent for a phonograph in which the mechanical design was entirely represented as an electrical circuit. The horn of the phonograph is represented as a transmission line, and is a resistive load for the rest of the circuit, while all the mechanical and acoustic parts—from the pickup needle through to the horn—are translated into lumped components according to the impedance analogy. The circuit arrived at is a ladder topology of series resonant circuits coupled by shunt capacitors. This can be viewed as a bandpass filter circuit. Harrison designed the component values of this filter to have a specific passband corresponding to the desired audio passband (in this case 100 Hz to 6 kHz) and a flat response. Translating these electrical element values back into mechanical quantities provided specifications for the mechanical components in terms of mass and stiffness, which in turn could be translated into physical dimensions for their manufacture. The resulting phonograph has a flat frequency response in its passband and is free of the resonances previously experienced. Shortly after this, Harrison filed another patent using the same methodology on telephone transmit and receive transducers.

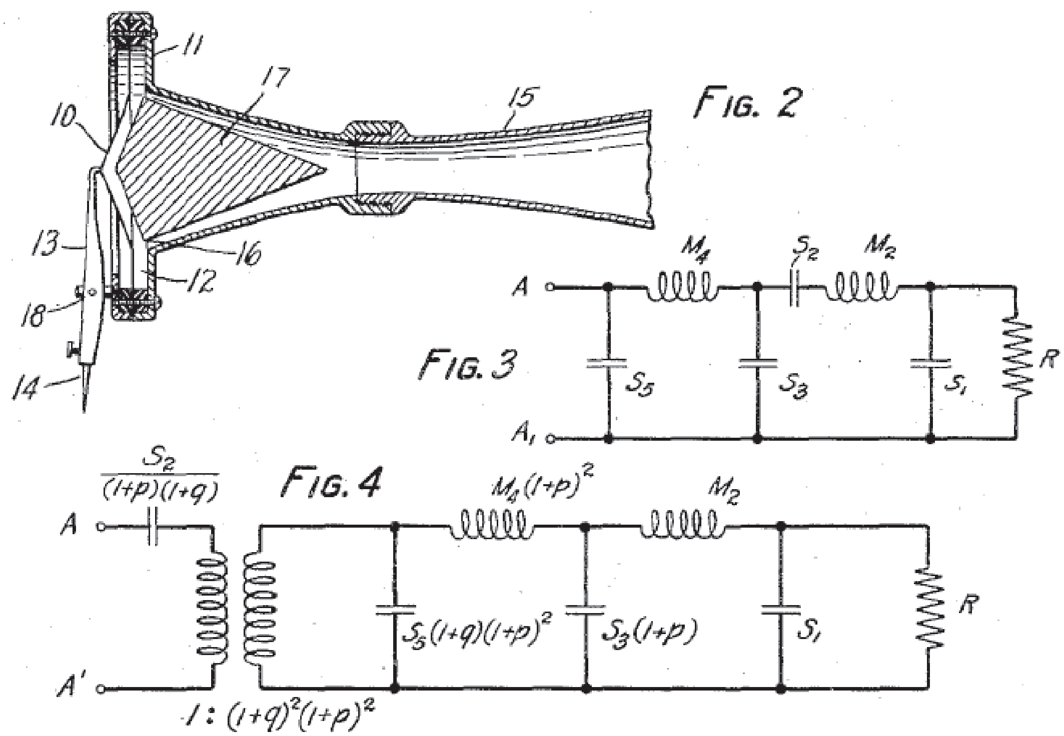


Figure 3. Norton's mechanical filter together with its electrical equivalent circuit.

Harrison used Campbell's image filter theory, which was the most advanced filter theory available at the time. In this theory, filter design is viewed essentially as an impedance matching problem. More advanced filter theory was brought to bear on this problem by Norton in 1929 at Bell Labs. Norton followed the same general approach though he later described to Darlington the filter he designed as being "maximally flat". Norton's mechanical design predates the paper by Butterworth who is usually credited as the first to describe the electronic maximally flat filter. The equations Norton gives for his filter correspond to a singly terminated Butterworth filter, that is, one driven by an ideal voltage source with no impedance, whereas the form more usually given in texts is for the doubly terminated filter with resistors at both ends, making it hard to recognise the design for what it is. Another unusual feature of Norton's filter design arises from the series capacitor, which represents the stiffness of the diaphragm. This is the only series capacitor in Norton's representation, and without it, the filter could be analysed as a low-pass prototype. Norton moves the capacitor out of the body of the filter to the input at the expense of introducing a transformer into the equivalent circuit (Norton's figure 4). Norton has used here the "turning round the L" impedance transform to achieve this.

The definitive description of the subject from this period is Maxfield and Harrison's 1926 paper. There, they describe not only how mechanical bandpass filters can be applied to sound reproduction systems, but also apply the same principles to recording systems and describe a much improved disc cutting head.

Volume production

The first volume production of mechanical filters was undertaken by Collins Radio Company starting in the 1950s. These were originally designed for telephone frequency-division multiplex applications where there is commercial advantage in using high quality filters. Precision and steepness of the transition band leads to a reduced width of guard band, which in turn leads to the ability to squeeze more telephone channels into the same cable. This same feature is useful in radio transmitters for much the same reason. Mechanical filters quickly also found popularity in VHF/UHF radio intermediate frequency (IF) stages of the high end radio sets (military, marine, amateur radio and the like) manufactured by Collins. They were favoured in the radio application because they could achieve much higher Q-factors than the equivalent *LC* filter. High *Q* allows filters to be designed which have high selectivity, important for distinguishing adjacent radio channels in receivers. They also had an advantage in stability over both *LC* filters and monolithic crystal filters. The most popular design for radio applications was torsional resonators because radio IF typically lies in the 100 to 500 kHz band.

Transducers

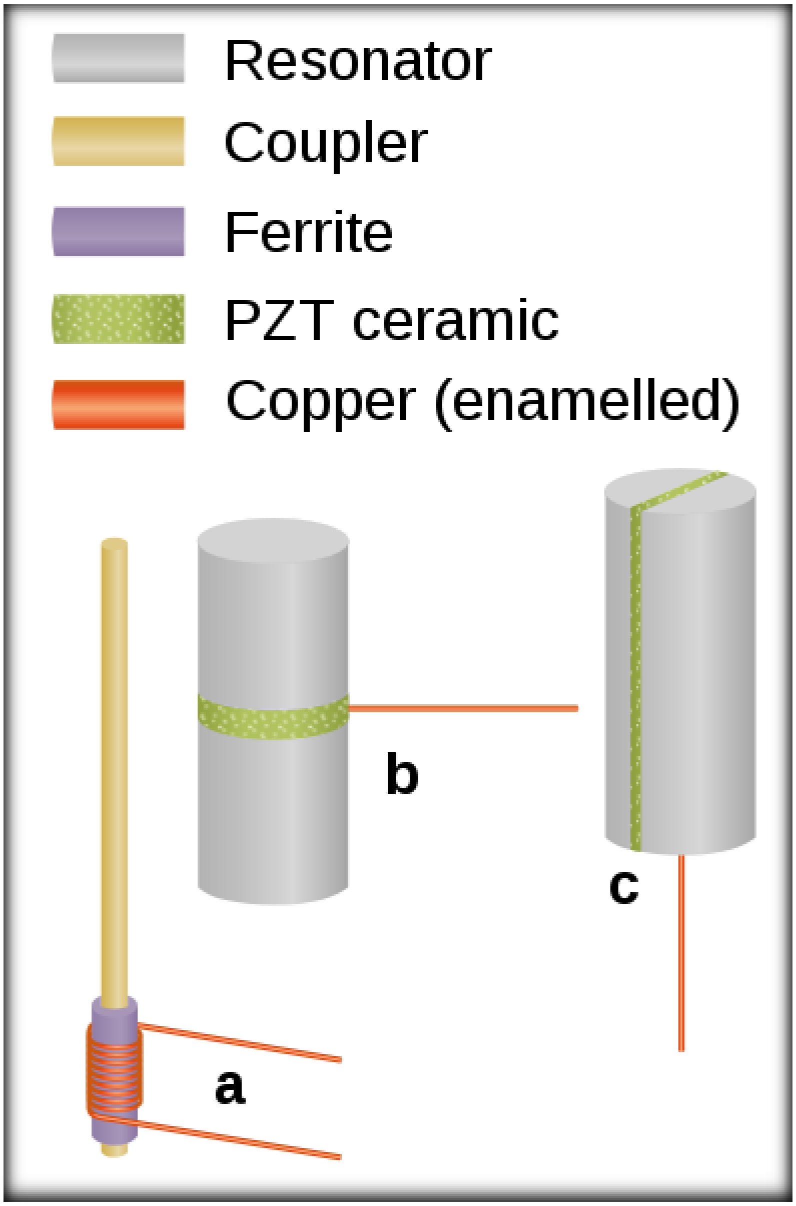


Figure 4. Mechanical filter transducers. **a** magnetostrictive transducer. **b** Langevin type piezoelectric transducer. **c** torsional piezoelectric transducer.

There are two general types of transducers used with mechanical filters: magnetostrictive and piezoelectric. Piezoelectric is favoured in more recent designs since the piezoelectric material can also be used as one of the resonators of the filter, thus reducing the number of components and thereby saving space. They also avoid the susceptibility to extraneous magnetic fields from which the magnetostrictive type suffers.

Magnetostrictive

A magnetostrictive material is one which changes shape when a magnetic field is applied. In reverse, it produces a magnetic field when distorted. The magnetostrictive transducer requires a coil of conducting wire around the magnetostrictive material. The coil either induces a magnetic field in the transducer and sets it in motion or else picks up an induced current from the motion of the transducer at the filter output. It is also usually necessary to have a small magnet to bias the magnetostrictive material into its operating range. It is possible to dispense with the magnets if the biasing is taken care of on the electronic side by providing a d.c. current superimposed on the signal, but this approach would detract from the generality of the filter design.

The usual magnetostrictive materials used for the transducer are either ferrite or compressed powdered iron. Mechanical filter designs often have the resonators coupled with steel or nickel-iron wires, but on some designs, especially older ones, nickel wire may be used for the input and output rods. This is because it is possible to wind the transducer coil directly on to a nickel coupling wire since nickel is slightly magnetostrictive. However, it is not strongly so and coupling to the electrical circuit is weak. This scheme also has the disadvantage that there are no measures taken to prevent eddy currents, a problem that is avoided if ferrites are used instead of nickel.

The coil of the transducer adds some inductance on the electrical side of the filter. It is common practice to add a capacitor in parallel with the coil so that an additional resonator is formed which can be incorporated into the filter design. While this will not improve performance to the extent that an additional mechanical resonator would, there is some benefit and the coil has to be there in any case.

Piezoelectric

A piezoelectric material is one which changes shape when an electric field is applied. In reverse, it produces an electric field when it is distorted. A piezoelectric transducer, in essence, is made simply by plating electrodes on to the piezoelectric material. Early piezoelectric materials used in transducers such as barium titanate had poor temperature stability. This precluded the transducer from functioning as one of the resonators; it had to be a separate component. This problem was solved with the introduction of lead zirconate titanate (abbreviated PZT) which is stable enough to be used as a resonator. Another common piezoelectric material is quartz, which has also been used in mechanical filters. However, ceramic materials such as PZT are preferred for their greater electromechanical coupling coefficient.

One type of piezoelectric transducer is the Langevin type, named after a transducer used by Paul Langevin in early sonar research. This is good for longitudinal modes of vibration. It can also be used on resonators with other modes of vibration if the motion can be mechanically converted into a longitudinal motion. The transducer consists of a layer of piezoelectric material sandwiched transversally into a coupling rod or resonator.

Another kind of piezoelectric transducer has the piezoelectric material sandwiched in longitudinally, usually into the resonator itself. This kind is good for torsional vibration modes and is called a torsional transducer.

Resonators

Material	Q-factor
Nickel	several 100
Steel	several 1000
Aluminium	~10,000
Nickel-iron alloy	10,000 to 25,000 depending on composition

It is possible to achieve an extremely high *Q* with mechanical resonators. Mechanical resonators typically have a *Q* of 10,000 or so, and 25,000 can be achieved in torsional resonators using a particular nickel-iron alloy. This is an unreasonably high figure to achieve with LC circuits, whose *Q* is limited by the resistance of the inductor coils.

Early designs in the 1940s and 1950s started by using steel as a resonator material. This has given way to nickel-iron alloys, primarily to maximise the *Q* since this is often the primary appeal of mechanical filters rather than price. Some of the metals that have been used for mechanical filter resonators and their *Q* are shown in the table.

Piezoelectric crystals are also sometimes used in mechanical filter designs. This is especially true for resonators that are also acting as transducers for inputs and outputs.

One advantage that mechanical filters have over LC electrical filters is that they can be made very stable. The resonance frequency can be made so stable that it varies only 1.5 parts per billion (ppb) from the specified value over the operating temperature range (−25 to 85 °C), and its average drift with time can be as low as 4 ppb per day. This stability with temperature is another reason for using nickel-iron as the resonator material. Variations with temperature in the resonance frequency (and other features of the frequency function) are directly related to variations in the Young's modulus, which is a measure of stiffness of the material. Materials are therefore sought that have a small temperature coefficient of Young's modulus. In general, Young's modulus has a negative temperature coefficient (materials become less stiff with increasing temperature) but additions of small amounts of certain other elements in the alloy can produce a material with a temperature coefficient that changes sign from negative through zero to positive with temperature. Such a material will have a zero coefficient of temperature with resonance frequency around a particular temperature. It is possible to adjust the point of zero temperature coefficient to a desired position by heat treatment of the alloy.

Resonator modes

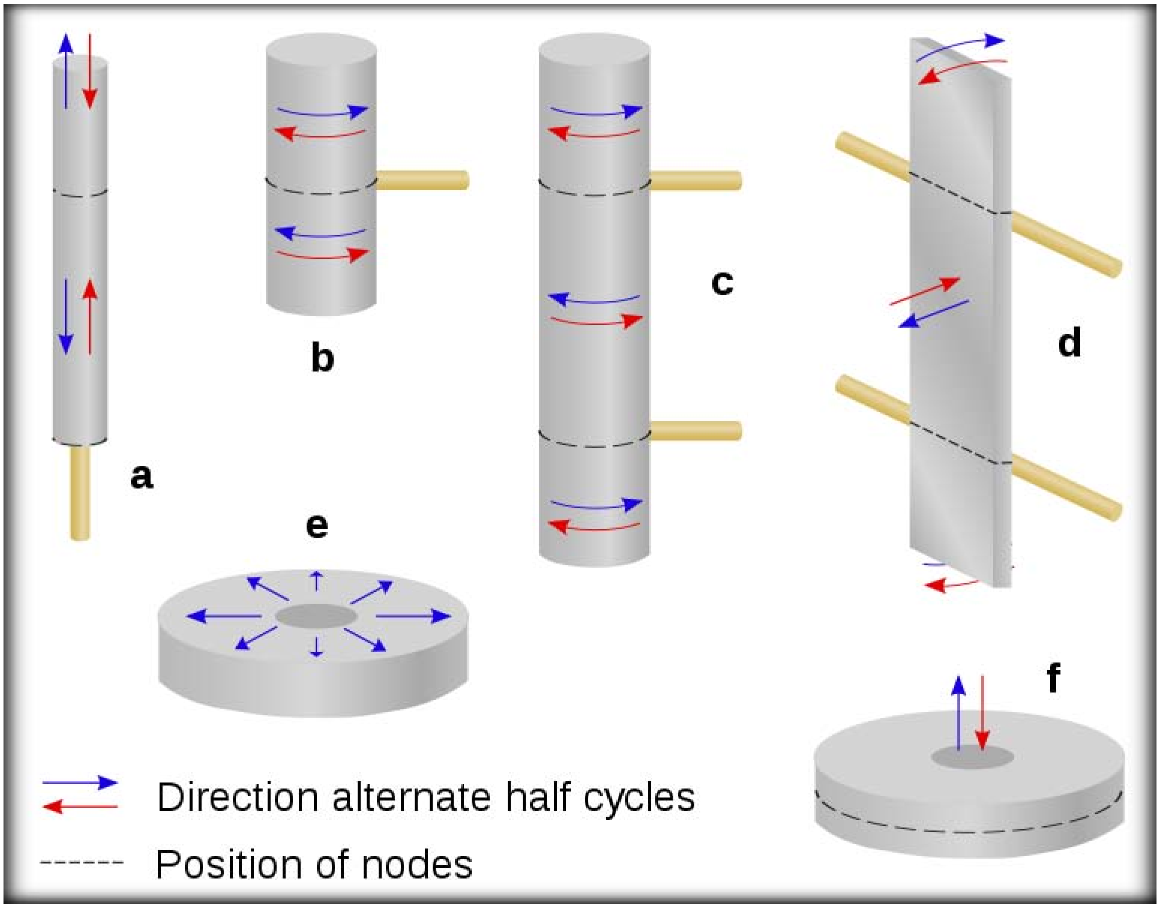


Figure 5. Some possible vibrational modes of resonators

It is usually possible for a mechanical part to vibrate in a number of different modes, however the design will be based on a particular vibrational mode and the designer will take steps to try to restrict the resonance to this mode. As well as the straightforward longitudinal mode some others which are used include flexural mode, torsional mode, radial mode and drumhead mode.

Modes are numbered according to the number of half-wavelengths in the vibration. Some modes exhibit vibrations in more than one direction (such as drumhead mode which has two) and consequently the mode number consists of more than one number. When the vibration is in one of the higher modes, there will be multiple nodes on the resonator where there is no motion. For some types of resonator, this can provide a convenient place to make a mechanical attachment for structural support. Wires attached at nodes will have no effect on the vibration of the resonator or the overall filter response. In figure 5, some possible anchor points are shown as wires attached at the nodes. The modes shown are (5a) the second longitudinal mode fixed at one end, (5b) the first torsional mode, (5c) the second torsional mode, (5d) the second flexural mode, (5e) first radial expansion mode and (5f) first radially symmetric drumhead mode.

Circuit designs

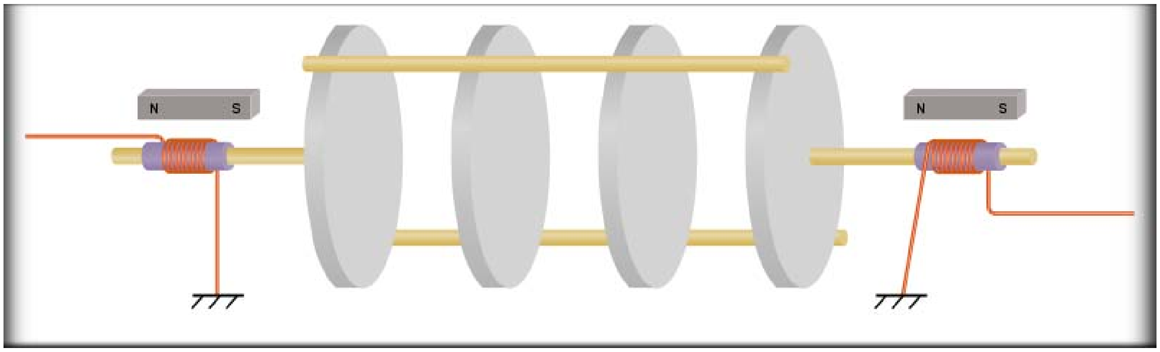


Figure 6. A mechanical filter using disc flexural resonators and magnetostrictive transducers

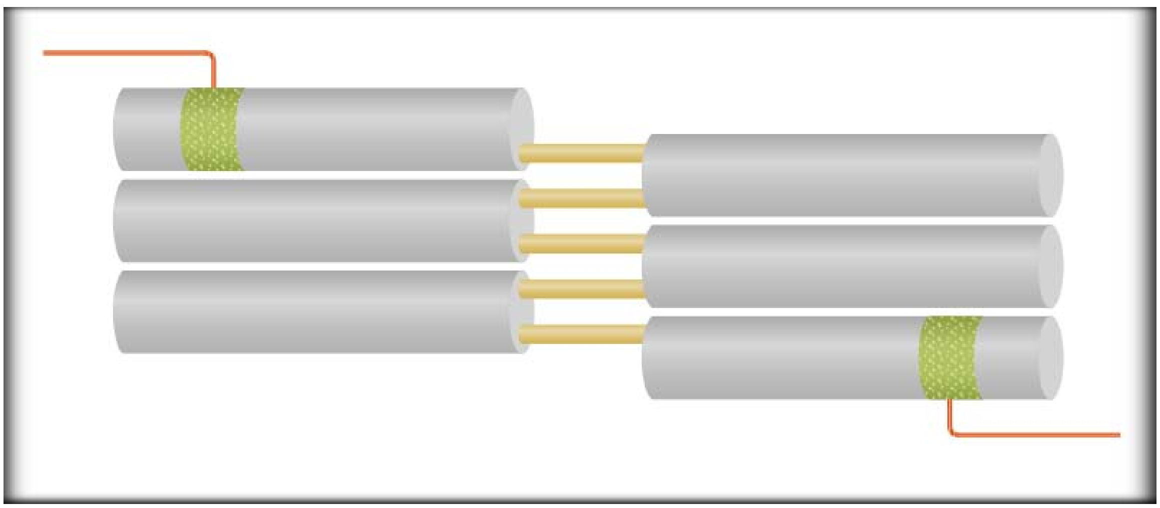


Figure 7. A filter using longitudinal resonators and Langevin type transducers

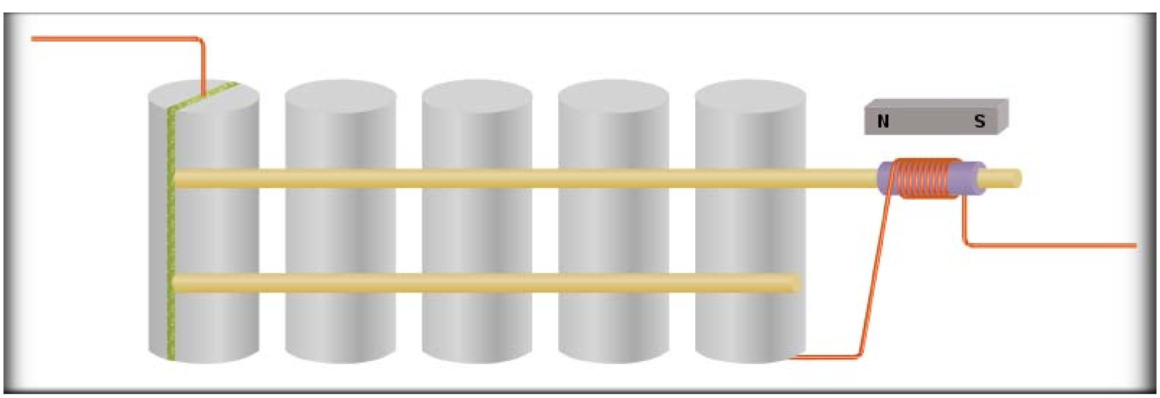


Figure 8a. A filter using torsional resonators. The input is shown with a torsional piezoelectric transducer and the output has a magnetostrictive transducer.

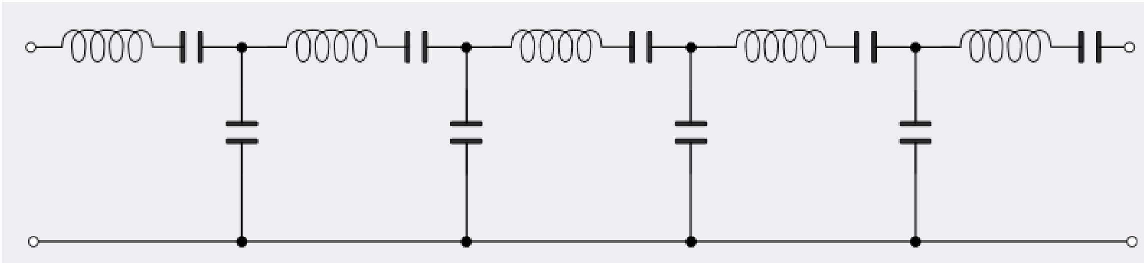


Figure 8b. Equivalent circuit of the torsional resonator circuit above

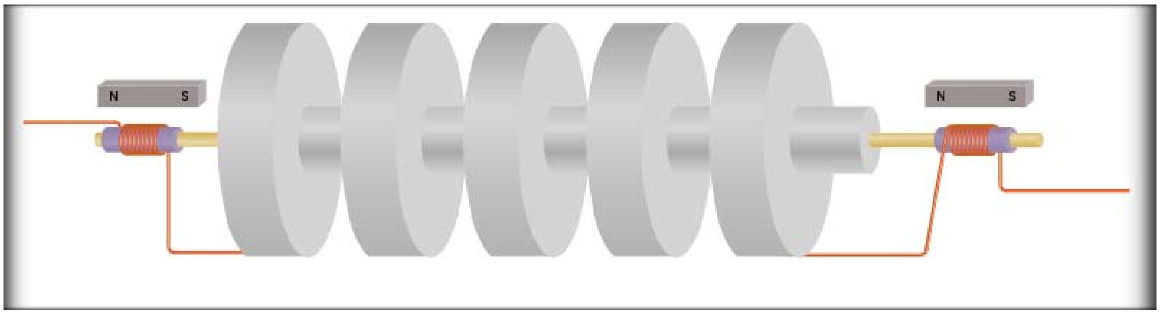


Figure 9. A filter using disc drumhead resonators

There are a great many combinations of resonators and transducers that can be used to construct a mechanical filter. A selection of some of these is shown in the diagrams. Figure 6 shows a filter using disc flexural resonators and magnetostrictive transducers. The transducer drives the centre of the first resonator, causing it to vibrate. The edges of the disc move in antiphase to the centre when the driving signal is at, or close to, resonance, and the signal is transmitted through the connecting rods to the next resonator. When the driving signal is not close to resonance, there is little movement at the edges, and the filter rejects (does not pass) the signal. Figure 7 shows a similar idea involving longitudinal resonators connected together in a chain by connecting rods. In this diagram, the filter is driven by piezoelectric transducers. It could equally well have used magnetostrictive transducers. Figure 8 shows a filter using torsional resonators. In this diagram, the input has a torsional piezoelectric transducer and the output has a magnetostrictive transducer. This would be quite unusual in a real design, as both input and output usually have the same type of transducer. The magnetostrictive transducer is only shown here to demonstrate how longitudinal vibrations may be converted to torsional vibrations and vice versa. Figure 9 shows a filter using drumhead mode resonators. The edges of the discs are fixed to the casing of the filter (not shown in the diagram) so the vibration of the disc is in the same modes as the membrane of a drum. Collins calls this type of filter a disc wire filter.

The various types of resonator are all particularly suited to different frequency bands. Overall, mechanical filters with lumped elements of all kinds can cover frequencies from about 5 to 700 kHz although mechanical filters down as low as a few kilohertz (kHz) are rare. The lower part of this range, below 100 kHz, is best covered with bar flexural resonators. The upper part is better done with torsional resonators. Drumhead disc resonators are in the middle, covering the range from around 100 to 300 kHz.

The frequency response behaviour of all mechanical filters can be expressed as an equivalent electrical circuit using the impedance analogy described above. An example of this is shown in figure 8b which is the equivalent circuit of the mechanical filter of figure 8a. Elements on the electrical side, such as the inductance of the magnetostrictive transducer, are omitted but would be taken into account in a complete design. The series resonant circuits on the circuit diagram represent the torsional resonators, and the shunt capacitors represent the coupling wires. The component values of the electrical equivalent circuit can be adjusted, more or less at will, by modifying the dimensions of the mechanical components. In this way, all the theoretical tools of electrical analysis and filter design can be brought to bear on the mechanical design. Any filter realisable in electrical theory can, in principle, also be realised as a mechanical filter. In particular, the popular finite element approximations to an ideal filter response of the Butterworth and Chebyshev filters can both readily be realised. As with the electrical counterpart, the more elements that are used, the closer the approximation approaches the ideal, however, for practical reasons the number of resonators does not normally exceed eight.

Semi-lumped designs

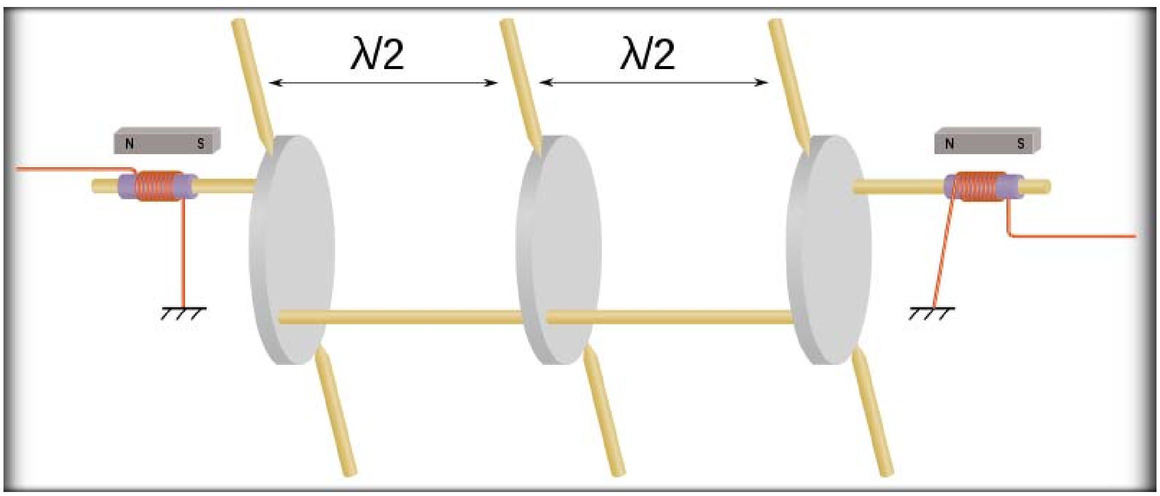


Figure 10a. A semi-lumped design using disc flexural resonators and λ/2 coupling wires

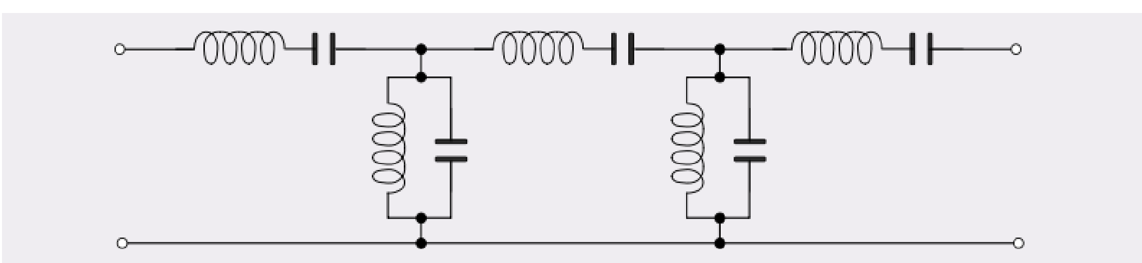


Figure 10b. Equivalent circuit of the semi-lumped circuit above

Frequencies of the order of megahertz (MHz) are above the usual range for mechanical filters. The components start to become very small, or alternatively the components are large compared to the signal wavelength. The lumped element model described above starts to break down and the components must be considered as distributed elements. The frequency at which the transition from lumped to distributed models takes place is much lower for mechanical filters than it is for their electrical counterparts. This is because mechanical vibrations travel at the speed of sound for the material the component is composed of. For solid components, this is many times (x15 for nickel-iron) the speed of sound in air (343 m/s) but still considerably less than the speed of electromagnetic waves (approx. 3×10^8 m/s in vacuum). Consequently, mechanical wavelengths are much shorter than electrical wavelengths for the same frequency. Advantage can be taken of these effects by deliberately designing components to be distributed elements, and the components and methods used in electrical distributed element filters can be brought to bear. The equivalents of stubs and impedance transformers are both achievable. Designs which use a mixture of lumped and distributed elements are referred to as semi-lumped.

An example of such a design is shown in figure 10a. The resonators are disc flexural resonators similar to those shown in figure 6, except that these are energised from an edge, leading to vibration in the fundamental flexural mode with a node in the centre, whereas the top diagram design is energised in the centre leading to vibration in the second flexural mode at resonance. The resonators are mechanically attached to the housing by pivots at right angles to the coupling wires. The pivots are to ensure free turning of the resonator and minimise losses. The resonators are treated as lumped elements; however, the coupling wires are made exactly one half-wavelength ($\lambda/2$) long and are equivalent to a $\lambda/2$ open circuit stub in the electrical equivalent circuit. For a narrow-band filter, a stub of this sort has the approximate equivalent circuit of a parallel shunt tuned circuit as shown in figure 10b. Consequently, the connecting wires are being used in this design to add additional resonators into the circuit and will have a better response than one with just the lumped resonators and short couplings. For even higher frequencies, microelectromechanical methods can be used as described below.

Bridging wires

Bridging wires are rods that couple together resonators that are not adjacent. They can be used to produce poles of attenuation in the stopband. This has the benefit of increasing the stopband rejection. When the pole is placed near the passband edge, it also has the benefit of increasing roll-off and narrowing the transition band. The typical effects of some of these on filter frequency response are shown in figure 11. Bridging across a single resonator (figure 11b) can produce a pole of attenuation in the high stopband. Bridging across two resonators (figure 11c) can produce a pole of attenuation in both the high and the low stopband. Using multiple bridges (figure 11d) will result in multiple poles of attenuation. In this way, the attenuation of the stopbands can be deepened over a broad frequency range.

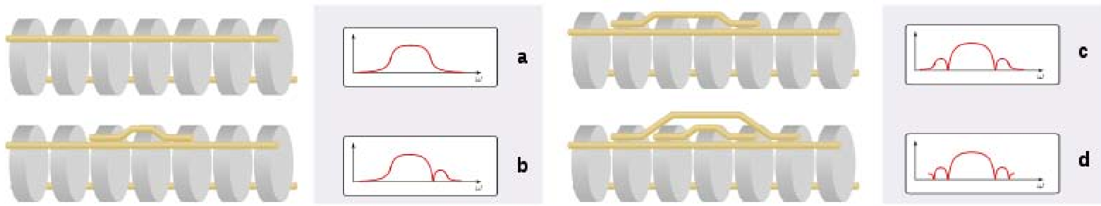


Figure 11. Schematic bridging arrangements and their effect on frequency response.

The method of coupling between non-adjacent resonators is not limited to mechanical filters. It can be applied to other filter formats. For instance, channels can be cut between cavity resonators, mutual inductance can be used with discrete component filters, and feedback paths can be used with active analogue or digital filters. Nor was the method first discovered in the field of mechanical filters; the earliest description is in a 1948 patent for filters using microwave cavity resonators. However, mechanical filter designers were the first (1960s) to develop practical filters of this kind and the method became a particular feature of mechanical filters.

Microelectromechanical filters

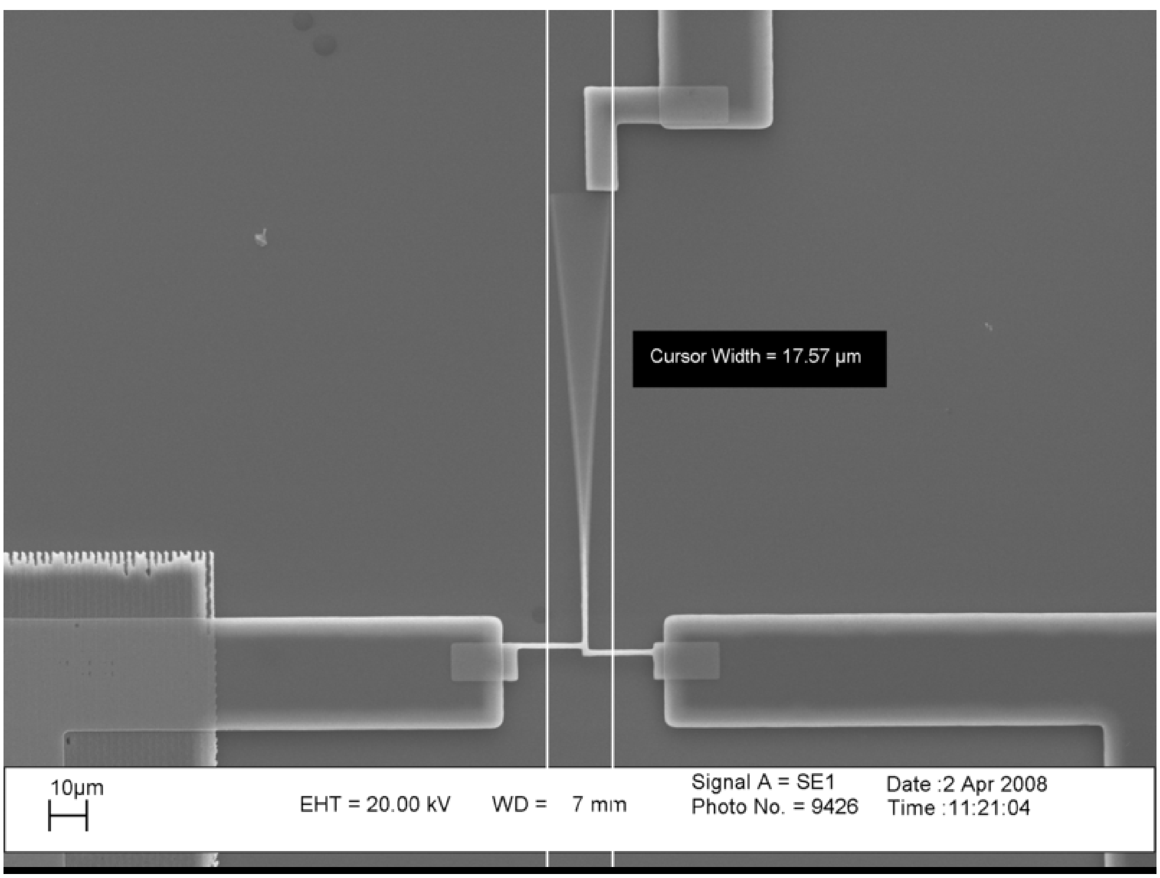


Figure 12. MEMS cantilever resonator. The device can be seen to be vibrating in this picture.

A new technology emerging in mechanical filtering is microelectromechanical systems (MEMS). MEMS are very small micromachines with component sizes measured in micrometres (μm), but not as small as nanomachines. These systems are mostly fabricated from silicon (Si), silicon nitride (Si_3N_4), or polymers. A common component used for radio frequency filtering (and MEMS applications generally), is the cantilever resonator. Cantilevers are simple mechanical components to manufacture by much the same methods used by the semiconductor industry; masking, photolithography and etching, with a final undercutting etch to separate the cantilever from the substrate. The technology has great promise since cantilevers can be produced in large numbers on a single substrate—much as large numbers of transistors are currently contained on a single silicon chip.

The resonator shown in figure 12 is around 120 μm in length. Experimental complete filters with an operating frequency of 30 GHz have been produced using cantilever varactors as the resonator elements. The size of this filter is around 4×3.5 mm. Cantilever resonators are typically applied at frequencies below 200 MHz, but other structures, such as micro-machined cavities, can be used in the microwave bands. Extremely high Q resonators can be made with this technology; flexural mode resonators with a Q in excess of 80,000 at 8 MHz are reported.

Adjustment

The precision applications in which mechanical filters are used require that the resonators are accurately adjusted to the specified resonance frequency. This is known as *trimming* and usually involves a mechanical machining process. In most filter designs, this can be difficult to do once the resonators have been assembled into the complete filter so the resonators are trimmed before assembly. Trimming is done in at least two stages; coarse and fine, with each stage bringing the resonance frequency closer to the specified value. Most trimming methods involve removing material from the resonator which will increase the resonance frequency. The target frequency for a coarse trimming stage consequently needs to be set below the final frequency since the tolerances of the process could otherwise result in a frequency higher than the following fine trimming stage could adjust for.

The coarsest method of trimming is grinding of the main resonating surface of the resonator; this process has an accuracy of around ±800 ppm. Better control can be achieved by grinding the edge of the resonator instead of the main surface. This has a less dramatic effect and consequently better accuracy. Processes that can be used for fine trimming, in order of increasing accuracy, are sandblasting, drilling, and laser ablation. Laser trimming is capable of achieving an accuracy of ±40 ppm.

Trimming by hand, rather than machine, was used on some early production components but would now normally only be encountered during product development. Methods available include sanding and filing. It is also possible to add material to the resonator by hand, thus reducing the resonance frequency. One such method is to add solder, but this is not suitable for production use since the solder will tend to reduce the high Q of the resonator.

In the case of MEMS filters, it is not possible to trim the resonators outside of the filter because of the integrated nature of the device construction. However, trimming is still a requirement in many MEMS applications. Laser ablation can be used for this but material deposition methods are available as well as material removal. These methods include laser or ion-beam induced deposition.

Chapter 8

Computer-Aided Design

Computer-aided design (**CAD**), also known as **computer-aided design and drafting** (**CADD**), is the use of computer technology for the process of design and design-documentation. Computer Aided Drafting describes the process of drafting with a computer. CADD software, or environments, provides the user with input-tools for the purpose of streamlining design processes; drafting, documentation, and manufacturing processes. CADD output is often in the form of electronic files for print or machining operations. The development of CADD-based software is in direct correlation with the processes it seeks to economize; industry-based software (construction, manufacturing, etc.) typically uses vector-based (linear) environments whereas graphic-based software utilizes raster-based (pixelated) environments.

CADD environments often involve more than just shapes. As in the manual drafting of technical and engineering drawings, the output of CAD must convey information, such as materials, processes, dimensions, and tolerances, according to application-specific conventions.

CAD may be used to design curves and figures in two-dimensional (2D) space; or curves, surfaces, and solids in three-dimensional (3D) objects.

CAD is an important industrial art extensively used in many applications, including automotive, shipbuilding, and aerospace industries, industrial and architectural design, prosthetics, and many more. CAD is also widely used to produce computer animation for special effects in movies, advertising and technical manuals. The modern ubiquity and power of computers means that even perfume bottles and shampoo dispensers are designed using techniques unheard of by engineers of the 1960s. Because of its enormous economic importance, CAD has been a major driving force for research in computational geometry, computer graphics (both hardware and software), and discrete differential geometry.

The design of geometric models for object shapes, in particular, is often called *computer-aided geometric design* (*CAGD*).

Overview

Beginning in the 1980s Computer-Aided Design programs reduced the need of draftsmen significantly, especially in small to mid-sized companies. Their affordability and ability to run on personal computers also allowed engineers to do their own drafting work, eliminating the need for entire departments. In today's world most, if not all, students in universities do not learn drafting techniques because they are not required to do so. The days of hand drawing for final drawings are almost obsolete. Universities no longer require the use of protractors and compasses to create drawings, instead there are several classes that focus on the use of CAD software such as Pro Engineer or IEAS-MS.

Current computer-aided design software packages range from 2D vector-based drafting systems to 3D solid and surface modellers. Modern CAD packages can also frequently allow rotations in three dimensions, allowing viewing of a designed object from any desired angle, even from the inside looking out. Some CAD software is capable of dynamic mathematic modeling, in which case it may be marketed as **CADD** — *computer-aided design and drafting*.

CAD is used in the design of tools and machinery and in the drafting and design of all types of buildings, from small residential types (houses) to the largest commercial and industrial structures (hospitals and factories).

CAD is mainly used for detailed engineering of 3D models and/or 2D drawings of physical components, but it is also used throughout the engineering process from conceptual design and layout of products, through strength and dynamic analysis of assemblies to definition of manufacturing methods of components. It can also be used to design objects.

CAD has become an especially important technology within the scope of computer-aided technologies, with benefits such as lower product development costs and a greatly shortened design cycle. CAD enables designers to lay out and develop work on screen, print it out and save it for future editing, saving time on their drawings.

Uses

Computer-aided design is one of the many tools used by engineers and designers and is used in many ways depending on the profession of the user and the type of software in question.

CAD is one part of the whole Digital Product Development (DPD) activity within the Product Lifecycle Management (PLM) process, and as such is used together with other tools, which are either integrated modules or stand-alone products, such as:

- Computer-aided engineering (CAE) and Finite element analysis (FEA)
- Computer-aided manufacturing (CAM) including instructions to Computer Numerical Control (CNC) machines
- Photo realistic rendering
- Document management and revision control using Product Data Management (PDM).

CAD is also used for the accurate creation of photo simulations that are often required in the preparation of Environmental Impact Reports, in which computer-aided designs of intended buildings are superimposed into photographs of existing environments to represent what that locale will be like were the proposed facilities allowed to be built. Potential blockage of view corridors and shadow studies are also frequently analyzed through the use of CAD..

Types

There are several different types of CAD. Each of these different types of CAD systems require the operator to think differently about how he or she will use them and he or she must design their virtual components in a different manner for each.

There are many producers of the lower-end 2D systems, including a number of free and open source programs. These provide an approach to the drawing process without all the fuss over scale and placement on the drawing sheet that accompanied hand drafting, since these can be adjusted as required during the creation of the final draft.

3D wireframe is basically an extension of 2D drafting. Each line has to be manually inserted into the drawing. The final product has no mass properties associated with it and cannot have features directly added to it, such as holes. The operator approaches these in a similar fashion to the 2D systems, although many 3D systems allow using the wireframe model to make the final engineering drawing views.

3D "dumb" solids are created in a way analogous to manipulations of real world objects. Basic three-dimensional geometric forms (prisms, cylinders, spheres, and so on) have solid volumes added or subtracted from them, as if assembling or cutting real-world objects. Two-dimensional projected views can easily be generated from the models. Basic 3D solids don't usually include tools to easily allow motion of components, set limits to their motion, or identify interference between components.

3D parametric solid modeling require the operator to use what is referred to as "design intent". The objects and features created are adjustable. Any future modifications will be simple, difficult, or nearly impossible, depending on how the original part was created. One must think of this as being a "perfect world" representation of the component. If a feature was intended to be located from the center of the part, the operator needs to locate it from the center of the model, not, perhaps, from a more convenient edge or an arbitrary point, as he could when using "dumb" solids. Parametric solids require the operator to consider the consequences of his actions carefully.

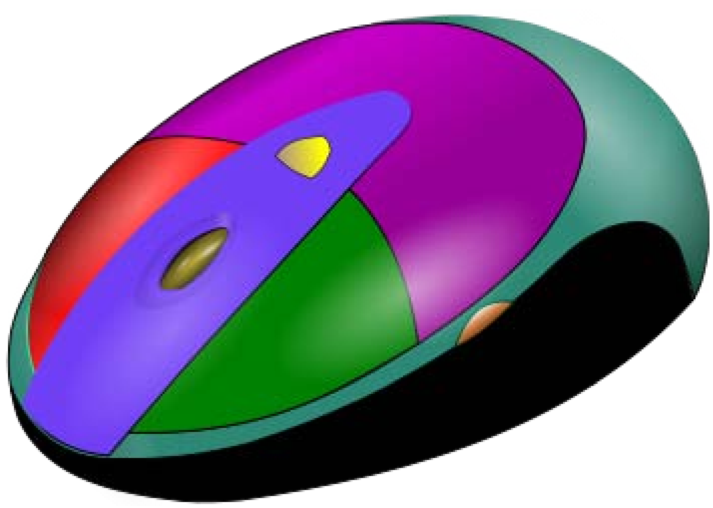
Some software packages provide the ability to edit parametric and non-parametric geometry without the need to understand or undo the design intent history of the geometry by use of direct modeling functionality. This ability may also include the additional ability to infer the correct relationships between selected geometry (e.g., tangency, concentricity) which makes the editing process less time and labor intensive while still freeing the engineer from the burden of understanding the model's design intent history. These kind of non history based systems are called Explicit Modellers or Direct CAD Modelers. The first Explicit Modeling system was introduced to the world at the end of 80's by Hewlett-Packard under the name SolidDesigner (now PTC's Creo Elements/Direct).

Draft views are able to be generated easily from the models. Assemblies usually incorporate tools to represent the motions of components, set their limits, and identify interference. The tool kits available for these systems are ever increasing; including 3D piping and injection mold designing packages.

Mid range software are integrating parametric solids more easily to the end user: integrating more intuitive functions (SketchUp), using the best of both 3D dumb solids and parametric characteristics (VectorWorks), making very real-view scenes in relative few steps (Cinema4D) or offering all-in-one (form•Z).

Top end systems offer the capabilities to incorporate more organic, aesthetics and ergonomic features into designs (Catia, GenerativeComponents). Freeform surface modelling is often combined with solids to allow the designer to create products that fit the human form and visual requirements as well as they interface with the machine.

Technology



A CAD model of a computer mouse.

Originally software for Computer-Aided Design systems was developed with computer languages such as Fortran, but with the advancement of object-oriented programming methods this has radically changed. Typical modern parametric feature based modeler and freeform surface systems are built around a number of key C modules with their own APIs. A CAD system can be seen as built up from the interaction of a graphical user interface (GUI) with NURBS geometry and/or boundary representation (B-rep) data via a geometric modeling kernel. A geometry constraint engine may also be employed to manage the associative relationships between geometry, such as wireframe geometry in a sketch or components in an assembly.

Unexpected capabilities of these associative relationships have led to a new form of prototyping called digital prototyping. In contrast to physical prototypes, which entail manufacturing time in the design.

Today, CAD systems exist for all the major platforms (Windows, Linux, UNIX and Mac OS X); some packages even support multiple platforms.

Right now, no special hardware is required for most CAD software. However, some CAD systems can do graphically and computationally expensive tasks, so a good graphics card, high speed (and possibly multiple) CPUs and large amounts of RAM are recommended.

The human-machine interface is generally via a computer mouse but can also be via a pen and digitizing graphics tablet. Manipulation of the view of the model on the screen is also sometimes done with the use of a spacemouse/SpaceBall. Some systems also support stereoscopic glasses for viewing the 3D model.

Effects

Beginning in the 1980s Computer-Aided Design programs reduced the need of draftsmen significantly, especially in small to mid-sized companies. Their affordability and ability to run on personal computers also allowed engineers to do their own drafting work, eliminating the need for entire departments. In today's world most, if not all, students in universities do not learn drafting techniques because they are not required to do so. The days of hand drawing for final drawings are almost obsolete. Universities no longer require the use of protractors and compasses to create drawings, instead there are several classes that focus on the use of CAD software such as Pro Engineer or IDEAS-MS.

Another consequence had been that since the latest advances were often quite expensive, small and even mid-size firms often could not compete against large firms who could use their computational edge for competitive purposes. Today, however, hardware and software costs have come down. Even high-end packages work on less expensive platforms and some even support multiple platforms. The costs associated with CAD implementation now are more heavily weighted to the costs of training in the use of these high level tools, the cost of integrating a CAD/CAM/CAE PLM using enterprise across multi-CAD and multi-platform environments and the costs of modifying design work flows to exploit the full advantage of CAD tools. CAD vendors have effectively lowered these training costs. These methods can be split into three categories:

1. Improved and simplified user interfaces. This includes the availability of “role” specific tailorable user interfaces through which commands are presented to users in a form appropriate to their function and expertise.
2. Enhancements to application software. One such example is improved design-in-context, through the ability to model/edit a design component from within the context of a large, even multi-CAD, active digital mockup.
3. User oriented modeling options. This includes the ability to free the user from the need to understand the design intent history of a complex intelligent model.

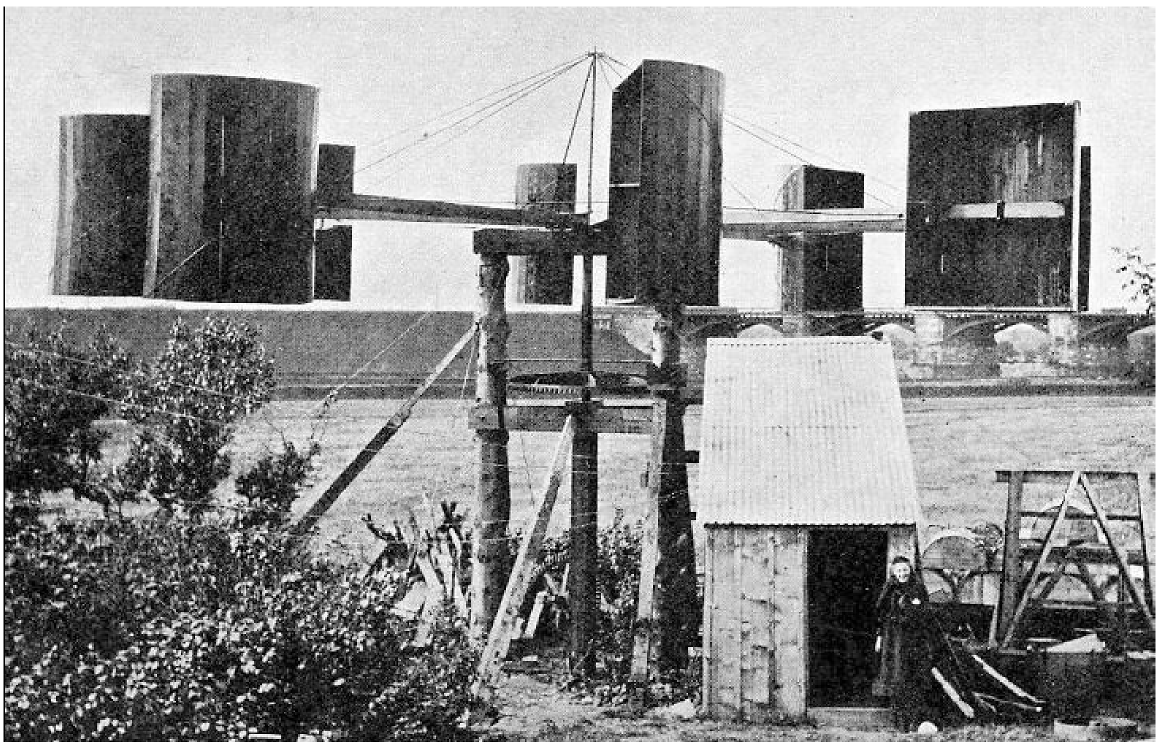
Chapter 9

Wind Turbine

Offshore wind farm using 5MW turbines REpower M5 in the North Sea off Belgium

A **wind turbine** is a device that converts kinetic energy from the wind into mechanical energy. If the mechanical energy is used to produce electricity, the device may be called a **wind generator** or **wind charger**. If the mechanical energy is used to drive machinery, such as for grinding grain or pumping water, the device is called a windmill or wind pump. Developed for over a millenium, today's wind turbines are manufactured in a range of vertical and horizontal axis types. The smallest turbines are used for applications such as battery charging or auxiliary power on sailing boats; while large grid-connected arrays of turbines are becoming an increasingly large source of commercial electric power.

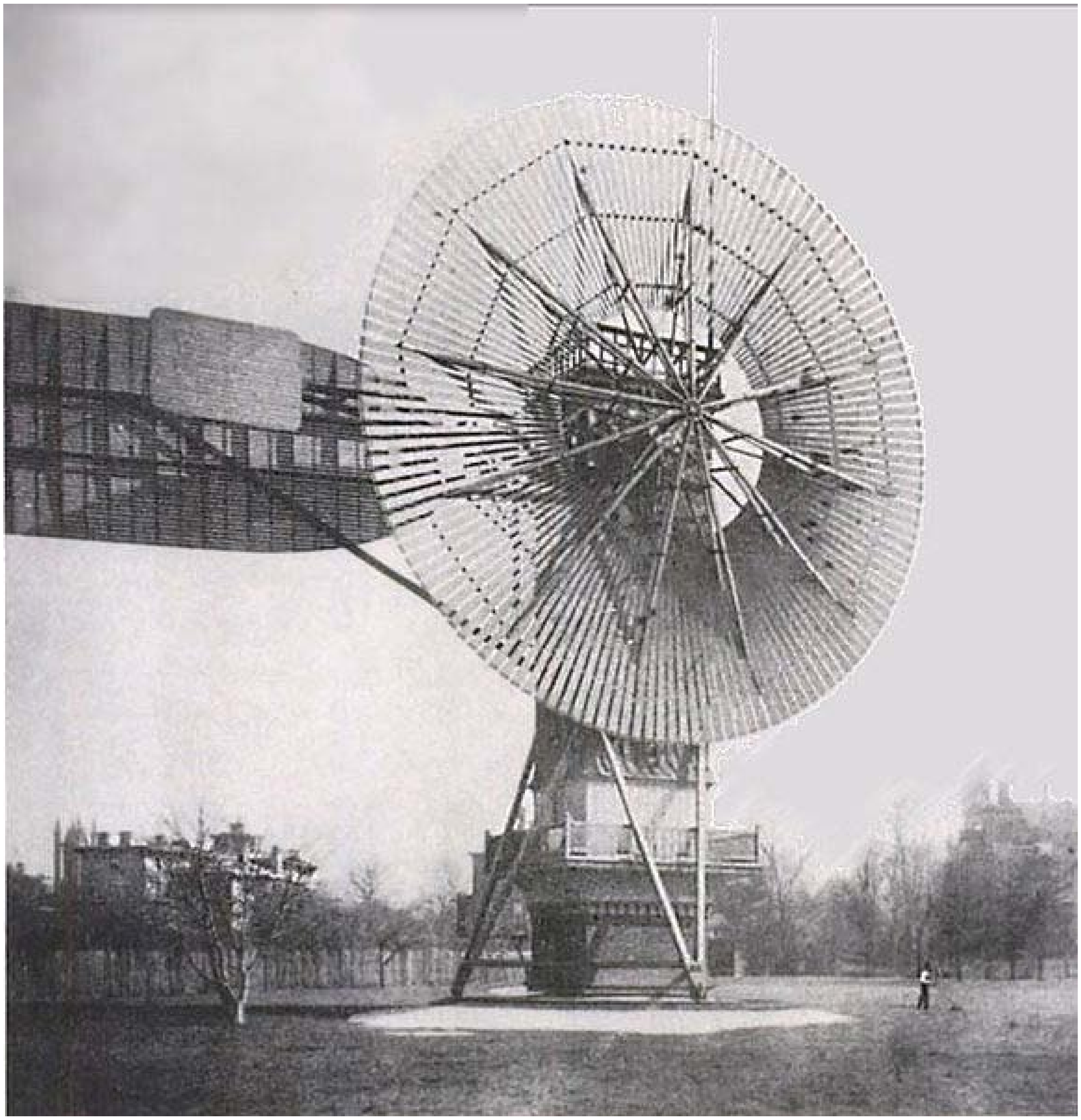
History



James Blyth's electricity generating wind turbine photographed in 1891

Windmills were used in Persia (present-day Iran) as early as 200 B.C. The windwheel of Heron of Alexandria marks one of the first known instances of wind powering a machine in history. However, the first known practical windmills were built in Sistan, a region between Afghanistan and Iran, from the 7th century. These "Panemone" were vertical axle windmills, which had long vertical driveshafts with rectangular blades. Made of six to twelve sails covered in reed matting or cloth material, these windmills were used to grind corn or draw up water, and were used in the gristmilling and sugarcane industries.

Windmills first appeared in Europe during the middle ages. The first historical records for their use in England date to the 11th or 12th centuries and there are reports of German crusaders taking their windmill-making skills to Syria around 1190. By the 14th century, Dutch windmills were in use to drain areas of the Rhine delta.



The first automatically operated wind turbine, built in Cleveland in 1887 by Charles F. Brush. It was 60 feet (18 m) tall, weighed 4 tons (3.6 metric tonnes) and powered a 12kW generator.

The first electricity generating wind turbine, was a battery charging machine installed in July 1887 by Scottish academic, James Blyth to light his holiday home in Marykirk, Scotland. Some months later American inventor Charles F Brush built the first automatically operated wind turbine for electricity production in Cleveland, Ohio. Although Blyth's turbine was considered uneconomical in the United Kingdom electricity generation by wind turbines was more cost effective in countries with widely scattered populations. In Denmark by 1900, there were about 2500 windmills for mechanical loads such as pumps and mills, producing an estimated combined peak power of about 30 MW. The largest machines were on 24-metre (79 ft) towers with four-bladed 23-metre (75 ft) diameter rotors. By 1908 there were 72 wind-driven electric generators operating in the

US from 5 kW to 25 kW. Around the time of World War I, American windmill makers were producing 100,000 farm windmills each year, mostly for water-pumping. By the 1930s, windmills for electricity were common on farms, mostly in the United States where distribution systems had not yet been installed. In this period, high-tensile steel was cheap, and windmills were placed atop prefabricated open steel lattice towers.

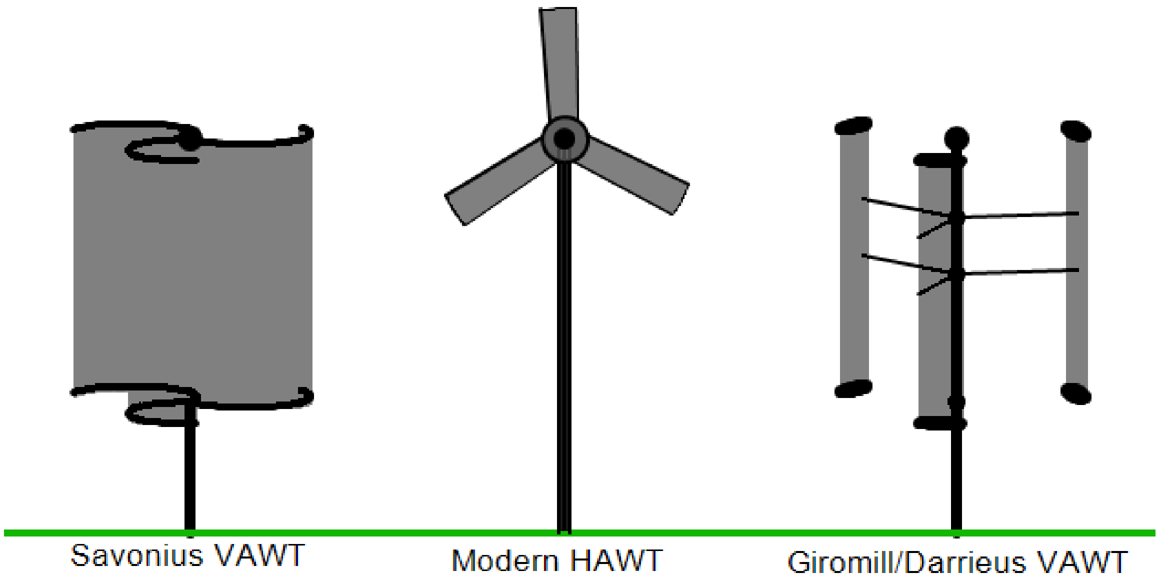
A forerunner of modern horizontal-axis wind generators was in service at Yalta, USSR in 1931. This was a 100 kW generator on a 30-metre (98 ft) tower, connected to the local 6.3 kV distribution system. It was reported to have an annual capacity factor of 32 per cent, not much different from current wind machines. In the fall of 1941, the first megawatt-class wind turbine was synchronized to a utility grid in Vermont. The Smith-Putnam wind turbine only ran for 1,100 hours before suffering a critical failure. The unit was not repaired because of shortage of materials during the war.

The first utility grid-connected wind turbine to operate in the U.K. was built by John Brown & Company in 1951 in the Orkney Islands.

Resources

A quantitative measure of the wind energy available at any location is called the Wind Power Density (WPD) It is a calculation of the mean annual power available per square meter of swept area of a turbine, and is tabulated for different heights above ground. Calculation of wind power density includes the effect of wind velocity and air density. Color-coded maps are prepared for a particular area described, for example, as "Mean Annual Power Density at 50 Meters." In the United States, the results of the above calculation are included in an index developed by the U.S. National Renewable Energy Lab and referred to as "NREL CLASS." The larger the WPD calculation, the higher it is rated by class. Classes range from Class 1 (200 watts/square meter or less at 50 meters altitude) to Class 7 (800 to 2000 watts/square meter). Commercial wind farms generally are sited in Class 3 or higher areas, although isolated points in an otherwise Class 1 area may be practical to exploit.

Types



The three primary types:VAWT Savonius, HAWT towered; VAWT Darrieus as they appear in operation.

Wind turbines can rotate about either a horizontal or a vertical axis, the former being both older and more common.

Horizontal axis

Components of a horizontal axis wind turbine (gearbox, rotor shaft and brake assembly) being lifted into position

Horizontal-axis wind turbines (HAWT) have the main rotor shaft and electrical generator at the top of a tower, and must be pointed into the wind. Small turbines are pointed by a simple wind vane, while large turbines generally use a wind sensor coupled with a servo motor. Most have a gearbox, which turns the slow rotation of the blades into a quicker rotation that is more suitable to drive an electrical generator.

Since a tower produces turbulence behind it, the turbine is usually positioned upwind of its supporting tower. Turbine blades are made stiff to prevent the blades from being pushed into the tower by high winds. Additionally, the blades are placed a considerable distance in front of the tower and are sometimes tilted forward into the wind a small amount.

Downwind machines have been built, despite the problem of turbulence (mast wake), because they don't need an additional mechanism for keeping them in line with the wind, and because in high winds the blades can be allowed to bend which reduces their swept area and thus their wind resistance. Since cyclical (that is repetitive) turbulence may lead to fatigue failures, most HAWTs are of upwind design.

11 x 7,5 MW E126 Estinnes Windfarm, Belgium, July 2010, one month before completion, with unique 2 part blades.

Modern wind turbines

Turbine blade convoy passing through Edenfield in the UK

Turbines used in wind farms for commercial production of electric power are usually three-bladed and pointed into the wind by computer-controlled motors. These have high tip speeds of over 320 kilometres per hour (200 mph), high efficiency, and low torque ripple, which contribute to good reliability. The blades are usually colored light gray to blend in with the clouds and range in length from 20 to 40 metres (66 to 130 ft) or more. The tubular steel towers range from 60 to 90 metres (200 to 300 ft) tall. The blades rotate at 10-22 revolutions per minute. At 22 rotations per minute the tip speed exceeds 300 feet per second (91 m/s). A gear box is commonly used for stepping up the speed of the generator, although designs may also use direct drive of an annular generator. Some models operate at constant speed, but more energy can be collected by variable-speed turbines which use a solid-state power converter to interface to the transmission system. All turbines are equipped with protective features to avoid damage at high wind speeds, by feathering the blades into the wind which ceases their rotation, supplemented by brakes.

Vertical axis design

Vertical-axis wind turbines (or VAWTs) have the main rotor shaft arranged vertically. Key advantages of this arrangement are that the turbine does not need to be pointed into the wind to be effective. This is an advantage on sites where the wind direction is highly variable, for example when integrated into buildings. The key disadvantages include the low rotational speed with the consequential higher torque and hence higher cost of the drive train, the inherently lower power coefficient, the 360 degree rotation of the aerofoil within the wind flow during each cycle and hence the highly dynamic loading on the blade, the pulsating torque generated by some rotor designs on the drive train, and the difficulty of modelling the wind flow accurately and hence the challenges of analysing and designing the rotor prior to fabricating a prototype.

With a vertical axis, the generator and gearbox can be placed near the ground, hence avoiding the need of a tower and improving accessibilty for maintenance. Drawbacks of this configuration include (i) wind speeds are lower close to the ground, so less wind energy is available for a given size turbine, and (ii) wind shear is more severe close to the ground, so the rotor experiences higher loads. Air flow near the ground and other objects can create turbulent flow, which can introduce problems associated with vibration, such as noise and bearing wear which may increase the maintenance or shorten the service life. However, when a turbine is mounted on a rooftop, the building generally redirects wind over the roof and this can double the wind speed at the turbine. If the height of the rooftop mounted turbine tower is approximately 50% of the building height, this is near the optimum for maximum wind energy and minimum wind turbulence. It should be borne in mind that wind speeds within the built environment are generally much lower than at exposed rural sites.

Subtypes

Darrieus wind turbine of 30 m in the Magdalen Islands

Darrieus wind turbine

"Eggbeater" turbines, or Darrieus turbines, were named after the French inventor, Georges Darrieus. They have good efficiency, but produce large torque ripple and cyclical stress on the tower, which contributes to poor reliability. They also generally require some external power source, or an additional Savonius rotor to start turning, because the starting torque is very low. The torque ripple is reduced by using three or more blades which results in greater solidity of the rotor. Solidity is measured by blade area divided by the rotor area. Newer Darrieus type turbines are not held up by guy-wires but have an external superstructure connected to the top bearing.

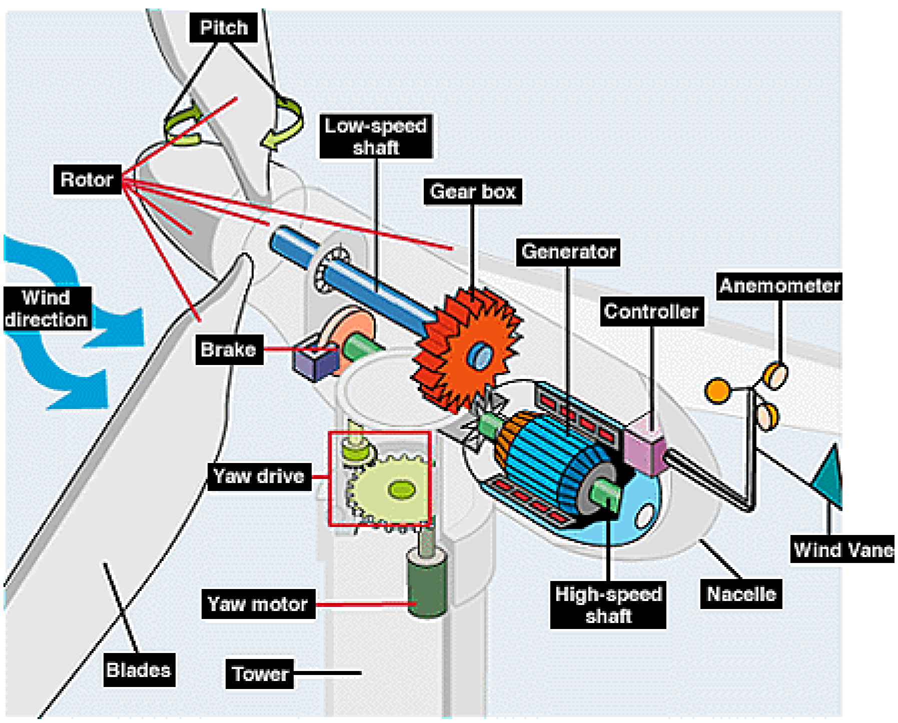
Giromill

A subtype of Darrieus turbine with straight, as opposed to curved, blades. The cycloturbine variety has variable pitch to reduce the torque pulsation and is self-starting. The advantages of variable pitch are: high starting torque; a wide, relatively flat torque curve; a lower blade speed ratio; a higher coefficient of performance; more efficient operation in turbulent winds; and a lower blade speed ratio which lowers blade bending stresses. Straight, V, or curved blades may be used.

Savonius wind turbine

These are drag-type devices with two (or more) scoops that are used in anemometers, *Flettner* vents (commonly seen on bus and van roofs), and in some high-reliability low-efficiency power turbines. They are always self-starting if there are at least three scoops. They sometimes have long helical scoops to give a smooth torque.

Turbine design and construction



Components of a horizontal-axis wind turbine

Wind turbines are designed to exploit the wind energy that exists at a location. Aerodynamic modeling is used to determine the optimum tower height, control systems, number of blades and blade shape.

Wind turbines convert wind energy to electricity for distribution. Conventional horizontal axis turbines can be divided into three components.

- The rotor component, which is approximately 20% of the wind turbine cost, includes the blades for converting wind energy to low speed rotational energy.
- The generator component, which is approximately 34% of the wind turbine cost, includes the electrical generator, the control electronics, and most likely a gearbox (e.g. planetary gearbox, adjustable-speed drive or continuously variable transmission) component for converting the low speed incoming rotation to high speed rotation suitable for generating electricity.
- The structural support component, which is approximately 15% of the wind turbine cost, includes the tower and rotor yaw mechanism.

A 1.5 MW wind turbine of a type frequently seen in the United States has a tower 80 meters high. The rotor assembly (blades and hub) weighs 48,000 pounds (22,000 kg). The nacelle, which contains the generator component, weighs 115,000 pounds (52,000 kg). The concrete base for the tower is constructed using 58,000 pounds (26,000 kg) of reinforcing steel and contains 250 cubic yards of concrete. The base is 50 feet (15 m) in diameter and 8 feet (2.4 m) thick near the center.

Unconventional wind turbines

One E-66 wind turbine at Windpark Holtriem, Germany, carries an observation deck, open for visitors. Another turbine of the same type, with an observation deck, is located in Swaffham, England. Airborne wind turbines have been investigated many times but have yet to produce significant energy. Conceptually, wind turbines may also be used in conjunction with a large vertical solar updraft tower to extract the energy due to air heated by the sun.

Wind turbines which utilise the Magnus effect have been developed.

Small wind turbines

A small wind turbine being used in Australia

Small wind turbines may be as small as a fifty-watt generator for boat or caravan use. Small units often have direct drive generators, direct current output, aeroelastic blades, lifetime bearings and use a vane to point into the wind.

Larger, more costly turbines generally have geared power trains, alternating current output, flaps and are actively pointed into the wind. Direct drive generators and aeroelastic blades for large wind turbines are being researched.

Record-holding turbines

Largest capacity

The Enercon E-126 has a rated capacity of 7.58 MW , has an overall height of 198 m (650 ft), a diameter of 126 m (413 ft), and is the world's largest-capacity wind turbine since its introduction in 2007.

At least four companies are working on the development of a 10MW turbine:

- American Superconductor
- Wind Power Ltd are developing a 10 MW VAWT, the Aerogenerator X
- Clipper Windpower are developing the Britannia 10 MW HAWT
- Sway AS announced the proposed development of a prototype 10 MW wind turbine with a height of 162.5 m (533 ft) and a rotor diameter of 145 m (475 ft).

Largest swept area

The turbine with the largest swept area is a prototype installed by Gamesa at Jaulín, Zaragoza, Spain in 2009. The G10X – 4.5 MW has a rotor diameter of 128m.

Tallest

The tallest wind turbine is Fuhrländer Wind Turbine Laasow. Its axis is 160 meters above ground and its rotor tips can reach a height of 205 meters. It is the only wind turbine in the world taller than 200 meters.

Largest vertical-axis

Le Nordais wind farm in Cap-Chat, Quebec has a vertical axis wind turbine (VAWT) named Éole, which is the world's largest at 110 m. It has a nameplate capacity of 3.8MW.

Most southerly

The turbines currently operating closest to the South Pole are three *Enercon E-33* in Antarctica, powering New Zealand's Scott Base and the United States' McMurdo Station since December 2009 although a modified HR3 turbine from Northern Power Systems operated at the Amundsen-Scott South Pole Station in 1997 and 1998. In March 2010 CITEDEF designed, built and installed a wind turbine in Argentine Marambio Base.

Most productive

Four turbines at Rønland wind farm in Denmark share the record for the most productive wind turbines, with each having generated 63.2 GWh by June 2010

Highest-situated

The world's highest-situated wind turbine is made by DeWind and located in the Andes, Argentina around 4,100 metres (13,500 ft) above sea level. The site uses a type D8.2 - 2000 kW / 50 Hz turbine. This turbine has a new drive train concept with a special torque converter (WinDrive) made by Voith and a synchronous generator. The WKA was put into operation in December 2007 and has supplied the Veladero mine of Barrick Gold with electricity since then.

Gallery of record-holders

Enercon E-126, highest rated capacity

Fuhrländer Wind Turbine Laasow, world's tallest

Éole, the largest vertical axis wind turbine, in Cap-Chat, Quebec

Highest-situated wind turbine, at the Veladero mine in San Juan Province, Argentina

Rønland, most productive turbines, in Denmark