Handbook of Automotive Handling, Parts and Technology

Twana Mosley
Arjun Applegate

First Edition, 2012

ISBN 978-81-323-0895-9

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

Table of Contents

Chapter- 1

Automobile Handling

Automobile handling and **vehicle handling** are descriptions of the way wheeled vehicles perform transverse to their direction of motion, particularly during cornering and swerving. It also includes their stability when moving at rest. Handling and braking are the major components of a vehicle's "active" safety. The maximum lateral acceleration is sometimes discussed separately as "road holding". Automobiles driven on public roads whose engineering requirements emphasize handling over comfort and passenger space are named sports cars.

Factors that affect a car's handling

Weight distribution

Center of mass height

The center of mass height, relative to the track, determines load transfer, (related to, but not exactly weight transfer), from side to side and causes body lean. When tires of a vehicle provide a centripetal force to pull it around a turn, the momentum of the vehicle actuates load transfer in a direction going from the vehicle's current position to a point on a path tangent to the vehicle's path. This load transfer presents itself in the form of body lean.

Height of the center of mass relative to the wheelbase determines load transfer between front and rear. The car's momentum acts at its center of mass to tilt the car forward or backward, respectively during braking and acceleration. Since it is only the downward force that changes and not the location of the center of mass, the effect on over/under steer is *opposite* to that of an actual change in the center of mass. When a car is braking, the downward load on the front tires increases and that on the rear decreases, with corresponding change in their ability to take sideways load.

A lower center of mass a principal performance advantage of sports cars, compared to sedans and (especially) SUVs. Some cars have body panels made of lightweight materials partly for this reason.

Body lean can also be controlled by the springs, anti-roll bars or the roll center heights.

Center of mass

The ideal weight distribution is "50/50" (i.e. the center of mass is mid-way between the front and rear axles). In steady-state cornering, front-heavy cars tend to understeer and rear-heavy cars to oversteer, all other things being equal. The mid-engine design seeks to achieve the ideal center of mass, though front-engine design has the advantage of permitting a more practical engine-passenger-baggage layout, and when engineered correctly can have as neutral (close to or at 50:50) weight distribution as a mid engine car. All other parameters being equal, at the hands of an expert professional driver a neutrally balanced mid-engine car can corner faster, but a FR layout car is easier to drive at the limit. Some good examples of these layouts are Ferrari F360 and Toyota MR2 for mid-engine design, and BMW M3 and Mazda MX-5 for the second.

The rearward weight bias preferred by sports and racing cars results from handling effects during the transition from straight-ahead to cornering. During corner entry the front tires, in addition to generating part of the lateral force required to accelerate the car's center of mass into the turn, also generate a torque about the car's vertical axis that starts the car rotating into the turn. However, the lateral force being generated by the rear tires is acting in the opposite torsional sense, trying to rotate the car out of the turn. For this reason, a car with "50/50" weight distribution will understeer on initial corner entry. To avoid this problem, sports and racing cars often have a more rearward weight distribution. In the case of pure racing cars, this is typically between "40/60" and "35/65". This gives the front tires an advantage in overcoming the car's moment of inertia (yaw angular inertia), thus reducing corner-entry understeer.

Using wheels and tires of different sizes (proportional to the weight carried by each end) is a lever automakers can use to fine tune the resulting over/understeer characteristics.

Roll angular inertia

This increases the time it takes to settle down and follow the steering. It depends on the (square of the) height and width, and (for a uniform mass distribution) can be approximately calculated by the equation: $I = M(height^2 + width^2) / 12$.

Greater width, then, though it counteracts center of gravity height, hurts handling by increasing angular inertia. Some high performance cars have light materials in their fenders and roofs partly for this reason.

Yaw and pitch angular inertia (polar moment)

Unless the vehicle is very short, compared to its height or width, these are about equal. Angular inertia determines the rotational inertia of an object for a given rate of rotation. The yaw angular inertia tends to keep the direction the car is pointing changing at a constant rate. This makes it slower to swerve or go into a tight curve, and it also makes it

slower to turn straight again. The pitch angular inertia detracts from the ability of the suspension to keep front and back tire loadings constant on uneven surfaces and therefore contributes to bump steer. Angular inertia is an integral over the *square* of the distance from the center of gravity, so it favors small cars even though the lever arms (wheelbase and track) also increase with scale. (Since cars have reasonable symmetrical shapes, the off-diagonal terms of the angular inertia tensor can usually be ignored.) Mass near the ends of a car can be avoided, without re-designing it to be shorter, by the use of light materials for bumpers and fenders or by deleting them entirely.

Suspension

Automobile suspensions have many variable characteristics, which are generally different in the front and rear and all of which affect handling. Some of these are: spring rate, damping, straight ahead camber angle, camber change with wheel travel, roll center height and the flexibility and vibration modes of the suspension elements. Suspension also affects unsprung weight.

Many cars have suspension that connects the wheels on the two sides, either by a sway bar and/or by a solid axle. The Citroën 2CV has interaction between the front and rear suspension.

The flexing of the frame interacts with the suspension.

Suspension travel

The severe handling vice of the TR3 and related cars was caused by running out of suspension travel. Other vehicles will run out of suspension travel with some combination of bumps and turns, with similarly catastrophic effect. Excessively modified cars also may encounter this problem.

Tires and wheels

In general softer rubber, higher hysteresis rubber and stiffer cord configurations increase road holding and improve handling. On most types of poor surfaces, large diameter wheels perform better than lower wider wheels. The depth of tread remaining greatly affects aquaplaning (riding over deep water without reaching the road surface). Increasing tire pressures reduces their slip angle, but lessening the contact area is detrimental in usual surface conditions and should be used with caution.

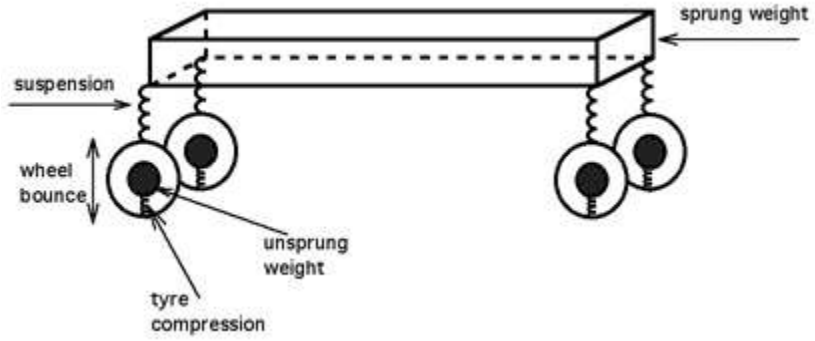
The amount a tire meets the road is an equation between the weight of the car and the type (and size) of its tire. A 1000kg car can depress a 185/65/15 tire more than a 215/45/15 tire longitudenally thus having better linear grip and better braking distance not to mention better aquaplaning performance, while the wider tires having better (dry) cornering resistance.

The contemporary chemical make-up of tires are dependant of the ambient and road temperatures. Ideally a tire should be soft enough to conform to the road surface (thus having good grip), but be hard enough to last for enough duration (distance) to be economically feasible. It is usually a good idea having different set of summer and winter tires for climates having these temperatures.

Track and wheelbase

The axle track provides the resistance to sideways weight transfer and body lean. The wheelbase provides resistance to front/back weight transfer and to pitch angular inertia, and provides the torque lever arm to rotate the car when swerving. The wheelbase, however, is less important than angular inertia (polar moment) to the vehicle's ability to swerve quickly.

Unsprung weight



Ignoring the flexing of other components, a car can be modeled as the sprung weight, carried by the springs, carried by the unsprung weight, carried by the tires, carried by the road. Unsprung weight is more properly regarded as a mass which has its own inherent inertia separate from the rest of the vehicle. When a wheel is pushed upwards by a bump in the road, the inertia of the wheel will cause it to be carried further upward above the height of the bump. If the force of the push is sufficiently large, the inertia of the wheel will cause the tire to completely lift off the road surface resulting in a loss of traction and control. Similarly when crossing into a sudden ground depression, the inertia of the wheel slows the rate at which it descends. If the wheel inertia is large enough, the wheel may be temporarily separated from the road surface before it has descended back into contact with the road surface.

This unsprung weight is cushioned from uneven road surfaces only by the compressive resilience of the tire (and wire wheels if fitted), and which aids the wheel in remaining in contact with the road surface when the wheel inertia prevents close-following of the ground surface. However, the compressive resilience of the tire results in rolling resistance which requires additional kinetic energy to overcome, and the rolling resistance is expended in the tire as heat due to the flexing of the rubber and steel bands in the sidewalls of the tires. To reduce rolling resistance for improved fuel economy and

to avoid overheating and failure of tires at high speed, tires are designed to have limited internal damping.

So the "wheel bounce" due to wheel inertia, or resonant motion of the unsprung weight moving up and down on the springiness of the tire, is only poorly damped, mainly by the dampers or shock absorbers of the suspension. For these reasons, high unsprung weight reduces road holding and increases unpredictable changes in direction on rough surfaces (as well as degrading ride comfort and increasing mechanical loads).

This unsprung weight includes the wheels and tires, usually the brakes, plus some percentage of the suspension, depending on how much of the suspension moves with the body and how much with the wheels; for instance a solid axle is completely unsprung. The main factors that improve unsprung weight are a sprung differential (as opposed to live axle) and inboard brakes. (The De Dion tube suspension operates much as a live axle does, but represents an improvement because the diff is mounted to the body, thereby reducing the unsprung weight.) Aluminum wheels also help. Magnesium alloy wheels are even lighter but corrode easily.

Since only the brakes on the driving wheels can easily be inboard, the Citroën 2CV had inertial dampers on its rear wheel hubs to damp only wheel bounce.

Aerodynamics

Aerodynamic forces are generally proportional to the square of the air speed, therefore car aerodynamics become rapidly more important as speed increases. Like darts, aeroplanes, etc., cars can be stabilised by fins and other rear aerodynamic devices. However, in addition to this cars also use downforce or "negative lift" to improve road holding. This is prominent on many types of racing cars, but is also used on most passenger cars to some degree, if only to counteract the tendency for the car to otherwise produce positive lift.

In addition to providing increased adhesion, car aerodynamics are frequently designed to compensate for the inherent increase in oversteer as cornering speed increases. When a car corners, it must rotate about its vertical axis as well as translate its center of mass in an arc. However, in a tight-radius (lower speed) corner the angular velocity of the car is high, while in a longer-radius (higher speed) corner the angular velocity is much lower. Therefore, the front tires have a more difficult time overcoming the car's moment of inertia during corner entry at low speed, and much less difficulty as the cornering speed increases. So the natural tendency of any car is to understeer on entry to low-speed corners and oversteer on entry to high-speed corners. To compensate for this unavoidable effect, car designers often bias the car's handling toward less corner-entry understeer (such as by lowering the front roll center), and add rearward bias to the aerodynamic downforce to compensate in higher-speed corners. The rearward aerodynamic bias may be achieved by an airfoil or "spoiler" mounted near the rear of the car, but a useful effect can also be achieved by careful shaping of the body as a whole, particularly the aft areas

In recent years, aerodynamics have become an area of increasing focus by racing teams as well as car manufacturers. Advanced tools such as wind tunnels and computational fluid dynamics (CFD) have allowed engineers to optimize the handling characteristics of vehicles. Advanced wind tunnels such as Wind Shear's Full Scale, Rolling Road, Automotive Wind Tunnel recently built in Concord, North Carolina have taken the simulation of on-road conditions to the ultimate level of accuracy and repeatability under very controlled conditions. CFD has similarly been used as a tool to simulate aerodynamic conditions but through the use of extremely advanced computers and software to duplicate the car's design digitally then "test" that design on the computer.

Delivery of power to the wheels and brakes

The coefficient of friction of rubber on the road limits the magnitude of the vector sum of the transverse and longitudinal force. So the driven wheels or those supplying the most braking tend to slip sideways. This phenomenon is often explained by use of the circle of forces model.

One reason that sports cars are usually rear wheel drive is that power induced oversteer is useful, to a skilled driver, for tight curves. The weight transfer under acceleration has the opposite effect and either may dominate, depending on the conditions. Inducing understeer by applying power in a front wheel drive car is useful via proper use of "Left-foot braking." In any case, this is not an important safety issue, because power is not normally used in emergency situations. Using low gears down steep hills may cause some oversteer.

The effect of braking on handling is complicated by load transfer, which is proportional to the (negative) acceleration times the ratio of the center of gravity height to the wheelbase. The difficulty is that the acceleration at the limit of adhesion depends on the road surface, so with the same ratio of front to back braking force, a car will understeer under braking on slick surfaces and oversteer under hard braking on solid surfaces. Most modern cars combat this by varying the distribution of braking in some way. This is important with a high center of gravity, but it is also done on low center of gravity cars, from which a higher level of performance is expected.

Steering

Depending on the driver, steering force and transmission of road forces back to the steering wheel and the steering ratio of turns of the steering wheel to turns of the road wheels affect control and awareness. Play — free rotation of the steering wheel before the wheels rotate — is a common problem, especially in older model and worn cars. Another is friction. Rack and pinion steering is generally considered the best type of mechanism for control effectiveness. The linkage also contributes play and friction. Caster — offset of the steering axis from the contact patch — provides some of the self-centering tendency.

Precision of the steering is particularly important on ice or hard packed snow where the slip angle at the limit of adhesion is smaller than on dry roads.

The steering effort depends on the downward force on the steering tires and on the radius of the contact patch. So for constant tire pressure, it goes like the 1.5 power of the vehicle's weight. The driver's ability to exert torque on the wheel scales similarly with his size. The wheels must be rotated farther on a longer car to turn with a given radius. Power steering reduces the required force at the expense of feel. It is useful, mostly in parking, when the weight of a front-heavy vehicle exceeds about ten or fifteen times the driver's weight, for physically impaired drivers and when there is much friction in the steering mechanism.

Four-wheel steering has begun to be used on road cars (Some WW II reconnaissance vehicles had it). It relieves the effect of angular inertia by starting the whole car moving before it rotates toward the desired direction. It can also be used, in the other direction, to reduce the turning radius. Some cars will do one or the other, depending on the speed.

Steering geometry changes due to bumps in the road may cause the front wheels to steer in different directions together or independent of each other. The steering linkage should be designed to minimize this effect.

Electronic stability control

Electronic stability control (ESC) is a computerized technology that improves the safety of a vehicle's stability by attempting to detect and prevent skids. When ESC detects loss of steering control, the system applies individual brakes to help "steer" the vehicle where the driver wants to go. Braking is automatically applied to individual wheels, such as the outer front wheel to counter oversteer, or the inner rear wheel to counter understeer.

The stability control of some cars may not be compatible with some driving techniques, such as power induced over-steer. It is therefore, at least from a sporting point of view, preferable that it can be disabled.

Static alignment of the wheels

Of course things should be the same, left and right, for road cars. Camber affects steering because a tire generates a force towards the side that the top is leaning towards. This is called camber thrust. Additional front negative camber is used to improve the cornering ability of cars with insufficient camber gain.

Rigidity of the frame

The frame may flex with load, especially twisting on bumps. Rigidity is considered to help handling. At least it simplifies the suspension engineers work. Some cars, such as the Mercedes-Benz 300SL have had high doors to allow a stiffer frame.

Driver handling the car

Handling is a property of the car, but different characteristics will work well with different drivers.

Familiarity

A person learns to control a car much as he learns to control his body, so the more he has driven a car or type of car the better it will handle for them. One needs to take extra care for the first few months after buying a car, especially if it differs in design from those they are used to. Other things that a driver must adjust to include changes in tires, tire pressures and load. That is, handling is not just good or bad; it is also the same or different.

Position and support for the driver

Having to take up "g forces" in his/her arms interferes with a driver's precise steering. In a similar manner, a lack of support for the seating position of the driver may cause them to move around as the car undergoes rapid acceleration (through cornering, taking off or braking). This interferes with precise control inputs, making the car more difficult to control.

Being able to reach the controls easily is also an important consideration, especially if a car is being driven hard.

In some circumstances, good support may allow a driver to retain some control, even after a minor accident or after the first stage of an accident.

External conditions that affect handling

Weather

Weather affects handling by changing the amount of available traction on a surface. Different tires do best in different weather. Deep water is an exception to the rule that wider tires improve road holding.

Road condition

Cars with relatively soft suspension and with low unsprung weight are least affected by uneven surfaces, while on flat smooth surfaces the stiffer the better. Unexpected water, ice, oil, etc. are hazards.

Common handling problems

When any wheel leaves contact with the road there is a change in handling, so the suspension should keep all four (or three) wheels on the road in spite of hard cornering,

swerving and bumps in the road. It is very important for handling, as well as other reasons, not to run out of suspension travel and "bottom" or "top".

It is usually most desirable to have the car adjusted for a small amount of understeer, so that it responds predictably to a turn of the steering wheel and the rear wheels have a smaller slip angle than the front wheels. However this may not be achievable for all loading, road and weather conditions, speed ranges, or while turning under acceleration or braking. Ideally, a car should carry passengers and baggage near its center of gravity and have similar tire loading, camber angle and roll stiffness in front and back to minimise the variation in handling characteristics. A driver can learn to deal with excessive oversteer or understeer, but not if it varies greatly in a short period of time.

The most important common handling failings are;

- Understeer – the front wheels tend to crawl slightly or even slip and drift towards the outside of the turn. The driver can compensate by turning a little more tightly, but road-holding is reduced, the car's behaviour is less predictable and the tires are liable to wear more quickly.
- Oversteer – the rear wheels tend to crawl or slip towards the outside of the turn more than the front. The driver must correct by steering away from the corner, otherwise the car is liable to spin, if pushed to its limit. Oversteer is sometimes useful, to assist in steering, especially if it occurs only when the driver chooses it by applying power.
- Bump steer – the effect of irregularity of a road surface on the angle or motion of a car. It may be the result of the kinematic motion of the suspension rising or falling, causing toe-in or toe-out at the loaded wheel, ultimately affecting the yaw angle (heading) of the car. It may also be caused by defective or worn out suspension components. This will always happen under some conditions but depends on suspension, steering linkage, unsprung weight, angular inertia, differential type, frame rigidity, tires and tire pressures. If suspension travel is exhausted the wheel either bottoms or loses contact with the road. As with hard turning on flat roads, it is better if the wheel picks up by the spring reaching its neutral shape, rather than by suddenly contacting a limiting structure of the suspension.
- Body roll – the car leans towards the outside of the curve. This interferes with the driver's control, because he must wait for the car to finish leaning before he can fully judge the effect of his steering change. It also adds to the delay before the car moves in the desired direction. It also slightly changes the weight borne by the tires as described in weight transfer.
- Excessive load transfer – On any vehicle that is cornering, the outside wheels are more heavily loaded than the inside due to the CG being above the ground. Total weight transfer (sum of front and back), in steady cornering, is determined by the ratio of the height of a car's center of gravity to its axle track. When the weight transfer equals half the vehicle's loaded weight, it will start to roll over. This can be avoided by manually or automatically reducing the turn rate, but this causes further reduction in road-holding.

- Slow response – sideways acceleration does not start immediately when the steering is turned and may not stop immediately when it is returned to center. This is partly caused by body roll. Other causes include tires with high slip angle, and yaw and roll angular inertia. Roll angular inertia aggravates body roll by delaying it. Soft tires aggravate yaw angular inertia by waiting for the car to reach their slip angle before turning the car.

Compromises

Ride quality and handling have always been a compromise - technology has over time allowed automakers to combine more of both features in the same vehicle. High levels of comfort are difficult to reconcile with a low center of gravity, body roll resistance, low angular inertia, support for the driver, steering feel and other characteristics that make a car handle well.

For ordinary production cars, manufactures err towards deliberate understeer as this is safer for inexperienced or inattentive drivers than is oversteer. Other compromises involve comfort and utility, such as preference for a softer smoother ride or more seating capacity.

Inboard brakes improve both handling and comfort but take up space and are harder to cool. Large engines tend to make cars front or rear heavy. In tires, fuel economy, staying cool at high speeds, ride comfort and long wear all tend to conflict with road holding, while wet, dry, deep water and snow road holding are not exactly compatible. A-arm or wishbone front suspension tends to give better handling, because it provides the engineers more freedom to choose the geometry, and more road holding, because the camber is better suited to radial tires, than MacPherson strut, but it takes more space.

The older Live axle rear suspension technology, familiar from the Ford Model T, is still widely used in most sport utility vehicles and trucks, often for the purposes of durability (and cost). The live axle suspension is still used in some sports cars, like the Ford Mustang, and is better for drag racing, but generally has problems with grip on bumpy corners, fast corners and stability at high speeds on bumpy straights.

Aftermarket modifications and adjustments

Lowering the center of gravity will always help the handling (as well as reduce the chance of roll-over). This can be done to some extent by using plastic windows (or none) and light roof, hood (bonnet) and trunk (boot) lid materials, by reducing the ground clearance, etc. Increasing the track with "reversed" wheels will have a similar effect, but remember that the wider the car the less spare room it has on the road and the farther you may have to swerve to miss an obstacle. Stiffer springs and/or shocks, both front and rear, will generally improve handling on close to perfect surfaces, while worsening handling on less-than-perfect road conditions by "skipping" the car (and destroying grip), thus making handling the vehicle difficult. Aftermarket performance suspension kits are usually readily available.

Lighter (mostly aluminum or magnesium alloy) wheels improve handling as well as ride comfort, by lessening unsprung weight.

Moment of inertia can be reduced by using lighter bumpers and wings (fenders), or none at all.

Component	Reduce Under-steer	Reduce Over-steer
Weight distribution	center of gravity towards rear	center of gravity towards front
Front shock absorber	softer	stiffer
Rear shock absorber	stiffer	softer
Front sway bar	softer	stiffer
Rear sway bar	stiffer	softer
Front tire selection[1]	larger contact area[2]	smaller contact area
Rear tire selection	smaller contact area	larger contact area[2]
Front wheel rim width or diameter	larger[2]	smaller
Rear wheel rim width or diameter	smaller	larger[2]
Front tire pressure	lower pressure	increase pressure
Rear tire pressure	increase pressure	lower pressure
Front wheel camber	increase negative camber	reduce negative camber
Rear wheel camber	reduce negative camber	increase negative camber
Rear spoiler	smaller	larger
Front height (because these usually affect camber and roll resistance)	lower front end	raise front end
Rear height	raise rear end	lower rear end
Front toe in	decrease	increase
Rear toe in	decrease	increase
1) Tire contact area can be increased by using wider tires, or tires with fewer		

grooves in the tread pattern. Of course fewer grooves has the opposite effect in wet weather or other poor road conditions.

2) These also improve road holding, under most conditions.

Cars with unusual handling problems

Certain vehicles can be involved in a disproportionate share of single-vehicle accidents; their handling characteristics may play a role:

- Early Porsche 911s — suffered from treacherous lift off oversteer (where the rear of the car loses grip as the driver lifts off the accelerator); also the inside front wheel leaves the road during hard cornering on dry pavement, causing increasing understeer. The roll bar stiffness at the front is set to compensate for the rear-heaviness and gives neutral handling in ordinary driving. This compensation starts to give out when the wheel lifts. A skilled driver can use the 911's other features to his/her advantage, making the 911 an extremely capable sports car in expert hands. Later 911s have had increasingly sophisticated rear suspensions and larger rear tires, eliminating these problems.
- Triumph TR2, and TR3 — began to oversteer more suddenly when their inside rear wheel lifted.
- Volkswagen Beetle — (original Beetle) senstitivity to crosswinds, due to the lightness of the front of the rear engine car; and poor roll stability due to the swing axle suspension. People who drove them hard fitted reversed wheels and bigger rear tires and rims to ameliorate.
- Chevrolet Corvair - cited for dangerous handling in Unsafe at Any Speed caused by poor roll stability due to the swing axle rear suspension similar to that used in the Volkswagen Beetle. These problems were corrected with the redesign of the Corvair for 1965, however, it died from its negative publicity.
- The large, rear-engine Tatra (known as the 'Czech secret weapon') killed so many Nazi officers during World War II that the German Army eventually forbade its officers from driving the Tatra.
- Some 1950s American "full size" cars responded very slowly to steering changes because of their very large angular inertia, softly tuned suspension which made ride quality a priority over cornering, and comfort oriented cross bias tires. *Auto Motor und Sport* reported on one of these that they lacked the courage to test it for top speed, probably due to their familiarity with smaller European cars and their unfamiliarity with large American cars.
- Dodge Omni and Plymouth Horizon — these early American responses to the Volkswagen Rabbit were found "unacceptable" in their initial testing by Consumer Reports, due to an observed tendency to display an uncontrollable oscillating yaw from side to side under certain steering inputs. While Chrysler's denials of this behaviour were countered by a persistent trickle of independent reports of this behaviour, production of the cars was altered to equip them with both a lighter weight steering wheel and a steering damper, and no further reports of this problem were heard.

- The Suzuki Samurai — was similarly reported by Consumer Reports to exhibit a propensity to tipping over onto two wheels, to the point where Consumer Reports claimed they were afraid to continue testing the vehicle without the attachment of outrigger wheels to catch it from completely rolling over. In its first set of tests, the Samurai performed well. R. David Little, Consumers Union's technical director, drove the light SUV through several short, hard turns, designed to simulate an emergency, such as trying to avoid a child running in front of the car. An article published several years later in a Consumer Reports anniversary issue prompted Suzuki to sue. The suit was based on the perception that Consumer Reports rigged the results: "This case is about lying and cheating by Consumers Union for its own financial motives," George F. Ball, Suzuki's managing counsel, said Monday. "They were in debt [in 1988], and they needed a blockbuster story to raise and solicit funds." Entrepreneur Magazine reported that "Suzuki's case centered on a change CU made while testing the vehicle. After the Samurai and other SUVs completed the standard course without threatening to roll over, CU altered the course to make the turns more abrupt. The other vehicles didn't show a problem, but the Samurai tipped up and would have rolled over but for outriggers set up to prevent that outcome" After eight years in court the parties consented to a settlement which did not include monetary damages nor a retraction.. Commenting on the settlement, Consumer Union said,"Consumers Union also says in the agreement that it "never intended to imply that the Samurai easily rolls over in routine driving conditions." CU Vice President of Technical Policy further stated: "There is no apology. "We stand fully behind our testing and rating of the Samurai." In a joint press statement Suzuki recognized "CU's stated commitment for objective and unbiased testing and reporting."

- Mercedes-Benz A-Class — a tall car with a high center of gravity; early models showed excessive body roll during sharp swerving manoeuvres and rolled over, most particularly during the Swedish moose test. This was later corrected using Electronic Stability Control and retrofitted at great expense to earlier cars.
- Ford Explorer — a dangerous tendency to blow a rear tire and flip over. Ford had constructed a vehicle with a high center of gravity - the tendency to roll over on sharp changes in direction is built in to the vehicle. Ford attempted to counteract the forces of nature by specifying lower than optimum pressures, in the tires in order to induce them to lose traction and slide under sideways forces rather than to grip and force the vehicle to roll over. For reasons that were never entirely clear, tires from one factory tended to blow out when under inflated, these vehicles then rolled over, which led to a spate of well publicized single-vehicle accidents.

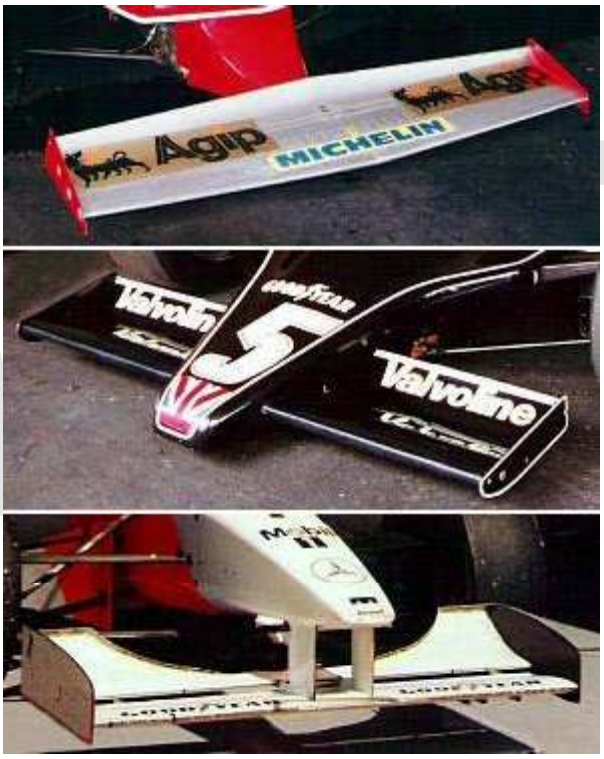
  Ford and Firestone, the makers of the tires, pointed fingers at each other, with the final blame being assigned to quality control practices at a Firestone plant which was undergoing a strike. Tires from a different Firestone plant were not associated with this problem. An internal document dated 1989 states
  Engineering has recommended use of tire pressures below maximum allowable inflation levels for all UN46 tires. As described previously, the reduced tire

pressures increase understeer and reduce maximum cornering capacity (both 'stabilising' influences). This practice has been used routinely in heavy duty pick-up truck and car station wagon applications to assure adequate understeer under all loading conditions. Nissan (Pathfinder), Toyota, Chevrolet, and Dodge also reduce tire pressures for selected applications. While we cannot be sure of their reasons, similarities in vehicle loading suggest that maintaining a minimal level of understeer under rear-loaded conditions may be the compelling factor.
This contributed to build-up of heat and tire deterioration under sustained high speed use, and eventual failure of the most highly stressed tire. Of course, the possibility that slightly substandard tire construction and slightly higher than average tire stress, neither of which would be problematic in themselves, would in combination result in tire failure is quite likely. The controversy continues without unequivocal conclusions, but it also brought public attention to a generally high incidence of rollover accidents involving SUVs, which the manufacturers continue to address in various ways. A subsequent NHTSA investigation of real world accident data showed that the SUVs in question were no more likely to roll over than any other SUV, after a tread separation.

- The Jensen GT (hatchback coupe) — was introduced in attempt to broaden the sales base of the Jensen Healey, which had up to that time been a roadster or convertible. Its road test report in *Motor Magazine* and a very similar one, soon after, in *Road & Track* concluded that it was no longer fun enough to drive to be worth that much money. They blamed it on minor suspension changes. Much more likely, the change in weight distribution was at fault. The Jensen Healey was a rather low and wide fairly expensive sports car, but the specifications of its suspension were not particularly impressive, having a solid rear axle. Unlike the AC Ace, with its double transverse leaf rear suspension and aluminium body, the Jensen Healey could not stand the weight of that high up metal and glass and still earn a premium price for its handling. The changes also included a cast iron exhaust manifold replacing the aluminium one, probably to partly balance the high and far back weight of the top. The factory building was used to build multi-tub truck frames.
- The rear engined Renault Dauphine earned in Spain the sobriquet of the *"widow's car"*, due to its bad handling.
- Three-wheeled cars/vehicles have unique handling issues, especially considering whether the single wheel is at the front or back. (Motorcycles with sidecars; another matter.) Buckminster Fuller's Dymaxion car caused a sensation, but ignorance of the problems of rear-wheel-steering led to a fatal crash that destroyed its reputation.

Chapter- 2

Downforce



Three different styles of front wings from three different Formula One eras, all designed to produce downforce at the front end of the respective race cars. *Top to bottom:* Ferrari 312 (1979), Lotus 79 (1978), McLaren MP4-10 (1995)

Downforce is a downwards thrust created by the aerodynamic characteristics of a car. The purpose of downforce is to allow a car to travel faster through a corner by increasing the vertical force on the tires, thus creating more grip.

Fundamental principles

The same principle that allows an airplane to rise off the ground by creating lift from its wings is used in reverse to apply force that presses the race car against the surface of the track. This effect is referred to as "aerodynamic grip" and is distinguished from "mechanical grip," which is a function of the car mass repartition, tires and suspension. The creation of downforce by passive devices almost always can only be achieved at the cost of increased aerodynamic drag (or friction), and the optimum setup is almost always a compromise between the two. The aerodynamic setup for a car can vary considerably between race tracks, depending on the length of the straights and the types of corners; some drivers also make different choices on setup. Because it is a function of the flow of air over and under the car, and because aerodynamic forces increase with the square of velocity, downforce increases with the square of the car's speed and requires a certain minimum speed in order to produce a significant effect. But some cars have had rather unstable aerodynamics, such that a minor change in angle of attack or height of the vehicle and this can cause large changes in the downforce. In the very worse cases the can cause the car to experience lift, not downforce, for example, caused by a bump on the track or slipstreaming over a crest, and sometimes can have disastrous consequences. A notorious example of this was Peter Dumbreck's Mercedes-Benz CLR in the 1999 Le Mans 24 hours, which flipped spectacularly after closely following a competitor car over a hump.

Two primary components of a racing car can be used to create downforce when the car is travelling at racing speed:

- the shape of the body, and
- the use of airfoils.

Most racing formulae have a ban on aerodynamic devices that can be adjusted during a race, except at pit stops.

The bottom panel of the Panoz DP01 ChampCar exhibiting complex aerodynamic design.

The underside curves of the Panoz DP01 ChampCar.

The formula for downforce of a wing is given by:

$$D = \frac{1}{2} \times (WS \times \ H \times \ AoA) \times F \times \rho \times V^2$$

Where:

- D is downforce in newtons
- WS is wingspan in metres
- H is height in metres
- AoA is angle of attack
- F is drag coefficient
- ρ is air density in kg/m³
- V is velocity in m/s

Body

The rounded and tapered shape of the top of the car is designed to slice through the air and minimize wind resistance. Detailed pieces of bodywork on top of the car can be added to allow a smooth flow of air to reach the downforce-creating elements (i.e., wings or spoilers, and underbody tunnels).

The overall shape of a street car resembles an airplane wing with air flowing over it faster than the air flows under it causing a difference in air pressure. Almost all street cars have aerodynamic lift as a result of this shape. There are many techniques that are used to counter-balance a street car. Looking at the profile of most street cars, the front bumper has the lowest ground clearance followed by the section between the front and rear tires, and followed yet by a rear bumper usually with the highest clearance. Using this method, the air flowing under the front bumper will make its way back to the rear bumper where it has a larger volume and thus a lower pressure. Race cars will exemplify this effect by adding a rear diffuser to better control the pressures directly under the rear bumper. Other aerodynamic components can be found on the underside to improve downforce and/or reduce drag include a splitter and a diffuser and vortex generators.

Airfoils

The amount of downforce created by the wings or spoilers on a car is dependent primarily on two things:

- The shape, including surface area, aspect ratio and cross-section of the device, and
- The device's orientation (or angle of attack).

A larger surface area creates greater downforce and greater drag (also known as air resistance). The aspect ratio is the width of the airfoil divided by its depth. The aspect ratio formula is written like AR=b squared/s, where AR=aspect ratio, b=spand squared, and s=wing area. Also, a greater angle of attack (or tilt) of the wing or spoiler, creates more downforce, which puts more pressure on the rear wheels and more drag.

The rear wing of a 1998 Formula One car, with three aerodynamic elements (1, 2, 3). The rows of holes for adjustment of the angle of attack (4) and installation of another element (5) are visible on the wing's endplate.

Front

The function of the airfoils at the front of the car is twofold. They create downforce that enhances the grip of the front tires, while also optimizing (or minimizing disturbance to) the flow of air to the rest of the car. The front wings on an open-wheeled car undergo constant modification as data is gathered from race to race, and are customized for every characteristic of a particular circuit. In most series, the wings are even designed for adjustment during the race itself when the car is serviced.

Rear

The flow of air at the rear of the car is affected by the front wings, front wheels, mirrors, driver's helmet, side pods and exhaust. This causes the rear wing to be less aerodynamically efficient than the front wing, Yet, because it must generate more than twice as much downforce as the front wings in order to maintain the handling to balance the car, the rear wing typically has a much larger aspect ratio, and often uses two or more elements to compound the amount of downforce created. Like the front wings, each of these elements can often be adjusted when the car is serviced, before or even during a race, and are the object of constant attention and modification.

Wings in unusual places

Partly as a consequence of rules aimed at reducing downforce from the front and rear wings of F1 cars, several teams have sought to find other places to position wings. Small wings mounted on the rear of the cars' sidepods began to appear in mid-1994, and were virtually standard on all F1 cars in one form or another, until all such devices were outlawed in 2009. Other wings have sprung up in various other places about the car, but these modifications are usually only used at circuits where downforce is most sought, particularly the twisty Hungary and Monaco racetracks.

The 1995 McLaren Mercedes MP4/10 was one of the first cars to feature a "midwing", using a loophole in the regulations to mount a wing on top of the engine cover. This arrangement has since been used by every team on the grid at one time or another, and in the 2007 Monaco Grand Prix all but two teams used them. These midwings are not to be confused either with the roll-hoop mounted cameras which each car carries as standard in all races, or with the bull-horn shaped flow controllers first used by McLaren and since by BMW Sauber, whose primary function is to smooth and redirect the airflow in order to make the rear wing more effective rather than to generate downforce themselves.

A variation on this theme was "X-wings", high wings mounted on the front of the sidepods which used a similar loophole to midwings. These were first used by Tyrrell in

1997, and were last used in the 1998 San Marino Grand Prix, by which time Ferrari, Sauber, Jordan and others had used such an arrangement. However it was decided they would have to be banned in view of the obstruction they caused during refueling and the risk they posed to the driver should a car roll over. (It is rumored that Bernie Ecclestone saw them as being too ugly on television and therefore had them banned.)

Various other extra wings have been tried from time to time, but nowadays it is more common for teams to seek to improve the performance of the front and rear wings by the use of various flow controllers such as the afore-mentioned "bull-horns" used by McLaren.

Chapter- 3

Inboard Brake and Oversteer

Inboard brake

An **inboard braking system** is an automobile technology wherein the brakes are mounted between the bearings of the wheels that constitute an axle.

Excepting the case of vehicles with beam axles and vehicles having no suspension, in practice it is normal for inboard brakes to be mounted rigidly with respect to the body of the vehicle, often to the differential casing. This is done to move the weight of the braking mechanism from being carried by the wheels directly (unsprung weight), to being carried indirectly by the wheels via the suspension (sprung mass). This then necessitates a means of transferring braking torque from the brake mechanism to the wheel, which is capable of operating despite the relative movement between body and wheel. Driven wheels already have shafting (or in older vehicles chains) which serve this purpose so there is no penalty for them, but undriven wheels require a similar mechanism which is then called a brake shaft.

The benefit of such a system is primarily the reduction of unsprung weight which improves handling and ride. The suspension does not have to resist twisting when the brakes are applied. The wheels don't enclose the brake mechanism allowing greater flexibility in wheel offset, and placement of suspension members. It is also much easier to protect the brake mechanism from the outside environment, and protect it from water, dust, and oil. Of secondary importance is flexible brake pipes are avoided and rigid pipes allow increases in brake fluid pressure, allowing a smaller disks for a given braking torque.

The mechanical disadvantages are largely those of added complexity. Undriven wheels require a brake shaft. It is more difficult to arrange for cooling air to flow over the wheel, and air ducting can be required, to prevent brake fade.

There can be practical difficulties,in servicing the brake mechanism. Instead of simply removing a wheel to renew pads and discs, the vehicle may need to be jacked up, so a mechanic can work underneath the vehicle. Additionally renewing brake discs can

require dismantling the half axle. This greatly discourages their use in motorsport, and the additional time makes for greater labour cost when servicing these parts.

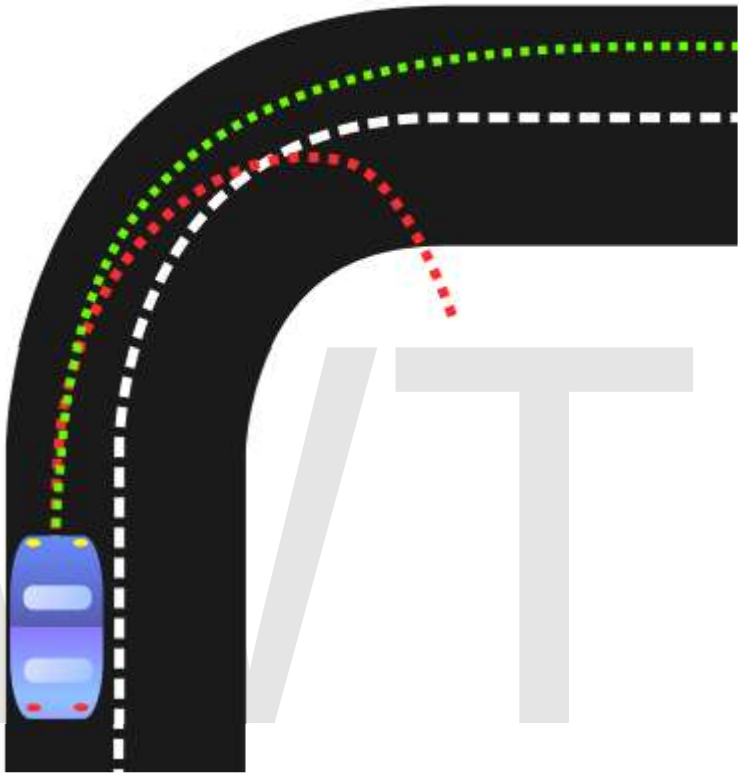
This system was more common in the 1960s, found on such cars as the Jaguar E-Type and Citroën 2CV. The Hummer H1 is one of the few modern vehicles fitted with inboard brakes, to accommodate each wheel's portal gear system.

Hybrid gasoline/electric vehicles may be considered to have partial inboard braking, because the devices used for the regenerative part of the braking are usually mounted inboard.

Cars with inboard brakes

- Alfa Romeo Alfasud Alfetta GTV GTV6 Giulietta 75 90 SZ Milano
- Audi 100 LS
- British Racing Motors: Some BRM racing cars had a single inboard disk brake, acting on both back wheels.
- Citroën 2CV, DS, GS, SM, Ami, Dyane, Axel and other Citroën models
- DKW Junior plus other models
- Hummer H1
- Jaguar E-Type, Jaguar XJ (until XJ40), Jaguar XJ-S
- Lancia Aurelia
- Lotus 72 Formula One racing car
- Lotus Elite, Elan and Esprit
- Mercedes-Benz W196 and 300SLR
- Monteverdi Hai 450 SS
- Oldsmobile Toronado
- Oltcit
- Rover P6
- Subaru G

Oversteer



Oversteer

Oversteer is a term for a car handling condition in which the slip angle of the rear tires is greater than the slip angle of the front tires. In other words, the amount that the car *steers* is *over* that commanded by the driver. The effect is opposite to that of understeer.

An oversteering car is referred to as "loose" or "free". Oversteer is a dynamically unstable condition; in other words, if control is lost, the vehicle will spin.

Causes

The tendency of a car to oversteer is affected by several factors such as mechanical traction, aerodynamics and suspension, and driver control, and may be applicable at any level of lateral acceleration. Generally, oversteer is the condition when the slip angle of the rear tires exceeds that of the front tires, even when they are both small. *Limit oversteer* occurs when the rear tires reach the limits of their lateral traction during a

cornering situation but the front tires have not, thus causing the rear of the vehicle to head towards the outside of the corner. The driving technique called opposite lock is meant to cope in this circumstance. *Trailing Throttle Oversteer* (TTO), a.k.a. "snap-oversteer" is induced by the weight balance of the car shifting from the rear to the front, this may happen if the car is cornering under throttle, causing the car to settle on the rear, if the throttle application would be removed—e.g. as to reduce the radius of the turn—the balance would suddenly shift to the front, giving less traction on the rear, if the car was already at the traction limit before the driver lifted the throttle it is very likely to cause a TTO. Rear wheel drive cars are more prone to oversteer, in particular when applying power in a tight corner. This occurs because the rear tires must handle both the lateral cornering force and engine torque.

Yaw rate

The terms oversteer and understeer are related to yaw rate and not to *sideways movement*. A car undergoes a circular spinning motion (yaw) as it turns, as well as *sideways movement* (towards the inside of the corner). Understeer and oversteer refer to the yaw motion. The difference between yaw and sideways movement is best demonstrated by practising turning an aircraft, because separate controls control each of the two movement types in aircraft. Consider a car with its steering wheel turned part way to one side and locked in that position. Now imagine that car rolling forward very slowly on a flat surface. It will move along an arc of a circle whose radius is determined solely by the position of the wheels, since centrifugal force is minimal. Its sideways motion and yaw rate are hence interlinked and set by the steering wheel position. However, the wheels can only provide a limited amount of sideways force before they slide. This sliding will happen at a larger radius as either the speed increases, the friction coefficient decreases, or the normal force decreases. Once this sliding occurs, the sideways movement and yaw rate may become unlinked. If the yaw rate of the car tends towards a larger radius than the radius set by the wheels, it is said to understeer. If the yaw rate radius is smaller (spinning too fast), it is called oversteer. During oversteer or understeer the sideways movement of the car may also follow a different radius to that set by the steering wheel, but this does not affect the definition of oversteer or understeer.

Critical speed

Oversteering cars have an associated instability mode, which occurs at and above the critical speed. As this speed is approached, with the car on an approximately straight course, the steering becomes progressively more sensitive. At the critical speed the yaw velocity gain becomes infinite, that is, the car will turn violently in response to the slightest steering input or external disturbance. Above the critical speed analysis shows that the yaw response will be reversed for a given steering wheel input, such as a car turning left in response to turning the wheel to the right. This is an oversimplification, however, as the model used is linearised in many important ways. Understeering cars do not suffer from this, which is one of the reasons why high speed cars tend to be set up to understeer.

In road cars

A Mercedes-Benz CLS AMG 55 oversteering on a wet surface

Contrary to popular opinion, modern rear-wheel-drive cars are much more user-friendly in regard to oversteer. Their suspension is not balanced heavily toward understeer, in fact with today's experience in making cars, most manufacturers try to achieve neutrality from the respective configurations so that they are largely capable of oversteering especially when the driver attempts to invoke it on purpose.

The natural reaction of most drivers to the perception of loss of control during oversteer is to immediately lift their foot off the gas pedal. Cutting the power mid-corner can induce more oversteer, known as lift-off oversteer. The correct reaction to oversteer is to gently steer into the slide and take the power away as needed without pitching the car forward. Indeed, "trail braking", or continuing to apply brake pressure after turning into a curve, can induce oversteer by transferring weight off the rear tires, regardless of whether the car is front, rear or all-wheel drive.

Braking may or may not improve the situation. Most modern cars have a brake bias which tends to straighten out the car. However, there are two factors working against this. Most drivers must lift their foot from the gas pedal in order to press the brake, inducing the spin as described above. The second is that braking transfers more of the vehicle's weight forward which tends to worsen oversteer. Even so, the brake bias may be enough to help or at least not make it worse.

In race cars

A car that tends neither to oversteer nor understeer when pushed to the limit is said to have neutral handling. It seems intuitive that race drivers would prefer a slight oversteer condition to rotate the car around a corner, but this isn't usually the case for two reasons. Accelerating early as the car passes the apex of a corner allows it to gain extra speed down the following straight. The driver who accelerates sooner and/or harder has a large advantage. The rear tires need some excess traction to accelerate the car in this critical phase of the corner, while the front tires can devote all their traction to turning. So the car must be set up with a slight understeer or "tight" tendency. Also, an oversteering car tends to be twitchy and ill tempered, making a race car driver more likely to lose control during a long race or when reacting to sudden situations in traffic.

Carroll Smith, in his book "Drive to Win", provides a detailed explanation of why a fast race car must have a bit of understeer. Note that this applies only to road racing. Dirt racing is a different matter.

Even so, some successful race car drivers do prefer a bit of oversteer in their cars, preferring a car which is less sedate and more willing to turn into corners (or inside their opponents). The judgement of a car's handling balance is not an objective one. Driving style is a major factor in the apparent balance of a car. This is why two drivers with identical cars on the same race team often run with rather different balance settings from each other. And both may call the balance of their cars 'neutral'.

Aerodynamic stability

The importance of the position of a fast car's aerodynamic centre of pressure to its directional stability was not understood at first. In the late 1950s, cars such as the 120 mph Jaguar 3.4-litre saloon / sedan were reported to feel directionally unstable at high speeds, and were badly affected by gusts.

Simple streamlining tends to lift the back of a car, reducing the downforce on its back wheels relative to the front wheels, resulting in oversteer. Streamlining also moves the centre of pressure well forward, causing directional instability in cross winds.

At first, aerodynamic oversteer was counteracted by setting the cars up with strong mechanical understeer, resulting in excessive understeer at lower speeds. Various means of achieving aerodynamic stability have since been developed, such as tail fins to move the centre of pressure back, the Kamm tail and the spoiler to reduce lift, rear wings to generate downward acting lift force, and air dams and skirts to reduce air pressure under the car, causing down force due to ground effect. Most of those features improve stability but increase drag, reducing top speed and increasing fuel consumption. However an early example of a fin used for directional stability without reducing top speed is provided by the Jaguar D-Type.

Usually these features are little more than styling gimmicks, the cars not being fast enough to benefit from them.

In modern race cars, especially open-wheel race cars, oversteering in high speed turns is caused mainly by aerodynamic configuration. A heavier aerodynamic load on the front of the car relative to the rear causes it to oversteer. Oversteer in low speed turns is often reduced or eliminated electronically through traction control (if the sanctioning body allows their use). The front/rear balance required to make the cars fast through corners is obtained by setting up the aerodynamics and balancing the suspension. The car's tendency toward oversteer is generally increased by softening the front suspension or stiffening the rear suspension in roll. The suspension's roll stiffness may be adjusted independently of pitch stiffness by means of adjustable or interchangeable anti-roll bars at one or both ends of the car. Camber angle, ride height, and tire pressures can also be used to tune the balance of the car.

Chapter- 4

Steering

Steering is the term applied to the collection of components, linkages, etc. which will allow a vessel (ship, boat) or vehicle (car, motorcycle, bicycle) to follow the desired course. An exception is the case of rail transport by which rail tracks combined together with railroad switches (and also known as 'points' in British English) provide the steering function.

Part of car steering mechanism: tie rod, steering arm, king pin axis (using ball joints).

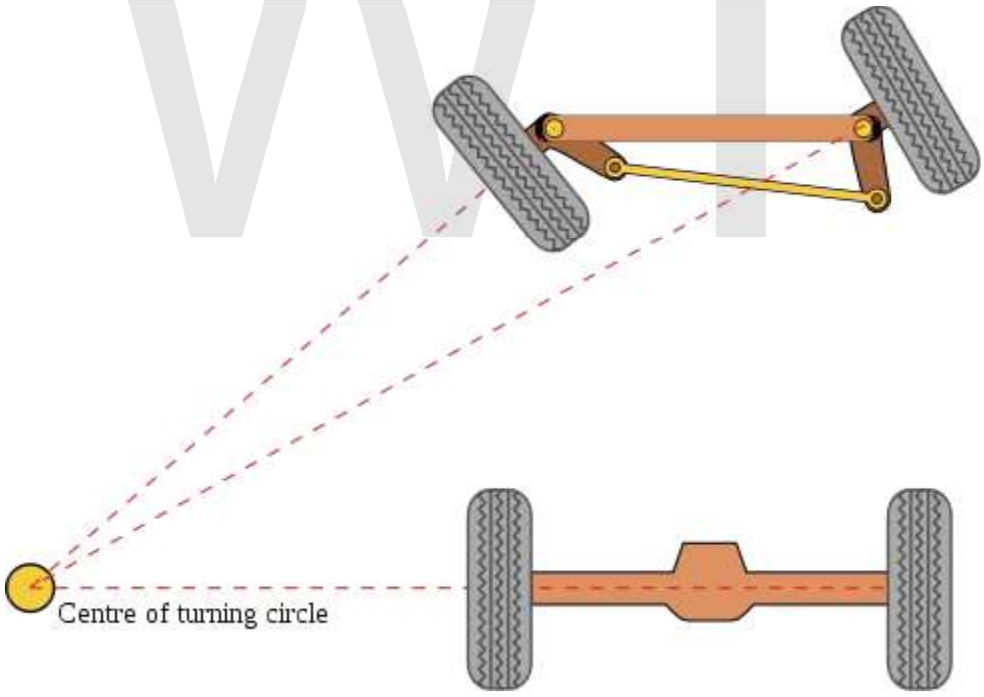
Introduction

The most conventional steering arrangement is to turn the front wheels using a hand–operated steering wheel which is positioned in front of the driver, via the steering column, which may contain universal joints (which may also be part of the collapsible steering column design), to allow it to deviate somewhat from a straight line. Other arrangements are sometimes found on different types of vehicles, for example, a tiller or rear–wheel steering. Tracked vehicles such as tanks usually employ differential steering — that is, the tracks are made to move at different speeds or even in opposite directions to bring about a change of course or direction.

Wheeled vehicle steering

Basic geometry

The basic aim of steering is to ensure that the wheels are pointing in the desired directions. This is typically achieved by a series of linkages, rods, pivots and gears. One of the fundamental concepts is that of *caster angle*- each wheel is steered with a pivot point ahead of the wheel; this makes the steering tend to be self-centering towards the direction travel.



Ackermann steering geometry

The steering linkages connecting the steering box and the wheels usually conforms to a variation of Ackermann steering geometry, to account for the fact that in a turn, the inner

wheel is actually travelling a path of smaller radius than the outer wheel, so that the degree of toe suitable for driving in a straight path is not suitable for turns.

Rack and pinion, recirculating ball, worm and sector

Rack and pinion

Rack and pinion unit mounted in the cockpit of an Ariel Atom sports car chassis. For most high volume production, this is usually mounted on the other side of this panel

Many modern cars use rack and pinion steering mechanisms, where the steering wheel turns the pinion gear; the pinion moves the rack, which is a linear gear that meshes with

the pinion, converting circular motion into linear motion along the transverse axis of the car (side to side motion). This motion applies steering torque to the swivel pin ball joints that replaced previously used kingpins of the stub axle of the steered wheels via tie rods and a short lever arm called the steering arm.

The rack and pinion design has the advantages of a large degree of feedback and direct steering "feel". A disadvantage is that it is not adjustable, so that when it does wear and develop lash, the only cure is replacement.

Older designs often use the recirculating ball mechanism, which is still found on trucks and utility vehicles. This is a variation on the older worm and sector design; the steering column turns a large screw (the "worm gear") which meshes with a sector of a gear, causing it to rotate about its axis as the worm gear is turned; an arm attached to the axis of the sector moves the Pitman arm, which is connected to the steering linkage and thus steers the wheels. The recirculating ball version of this apparatus reduces the considerable friction by placing large ball bearings between the teeth of the worm and those of the screw; at either end of the apparatus the balls exit from between the two pieces into a channel internal to the box which connects them with the other end of the apparatus, thus they are "recirculated".

The recirculating ball mechanism has the advantage of a much greater mechanical advantage, so that it was found on larger, heavier vehicles while the rack and pinion was originally limited to smaller and lighter ones; due to the almost universal adoption of power steering, however, this is no longer an important advantage, leading to the increasing use of rack and pinion on newer cars. The recirculating ball design also has a perceptible lash, or "dead spot" on center, where a minute turn of the steering wheel in either direction does not move the steering apparatus; this is easily adjustable via a screw on the end of the steering box to account for wear, but it cannot be entirely eliminated because it will create excessive internal forces at other positions and the mechanism will wear very rapidly. This design is still in use in trucks and other large vehicles, where rapidity of steering and direct feel are less important than robustness, maintainability, and mechanical advantage. The much smaller degree of feedback with this design can also sometimes be an advantage; drivers of vehicles with rack and pinion steering can have their thumbs broken when a front wheel hits a bump, causing the steering wheel to kick to one side suddenly (leading to driving instructors telling students to keep their thumbs on the front of the steering wheel, rather than wrapping around the inside of the rim). This effect is even stronger with a heavy vehicle like a truck; recirculating ball steering prevents this degree of feedback, just as it prevents desirable feedback under normal circumstances.

The worm and sector was an older design, used for example in Willys and Chrysler vehicles, and the Ford Falcon (1960s).

Other systems for steering exist, but are uncommon on road vehicles. Children's toys and go karts often use a very direct linkage in the form of a bellcrank (also commonly known as a Pitman arm) attached directly between the steering column and the steering arms,

and the use of cable-operated steering linkages (e.g. the Capstan and Bowstring mechanism) is also found on some home-built vehicles such as soapbox cars and recumbent tricycles.

Power steering

Power steering, assists the driver of an automobile in steering by directing a portion of the vehicle's power to traverse the axis of one or more of the roadwheels. As vehicles have become heavier and switched to front wheel drive, particularly using negative offset geometry, along with increases in tire width and diameter, the effort needed to turn the steering wheel manually has increased — often to the point where major physical exertion is required. To alleviate this, auto makers have developed power steering systems: or more correctly power-assisted steering — on road going vehicles there has to be a mechanical linkage as a fail safe. There are two types of power steering systems— hydraulic and electric/electronic. A hydraulic-electric hybrid system is also possible.

A hydraulic power steering (HPS) uses hydraulic pressure supplied by an engine-driven pump to assist the motion of turning the steering wheel. Electric power steering (EPS) is more efficient than the hydraulic power steering, since the electric power steering motor only needs to provide assistance when the steering wheel is turned, whereas the hydraulic pump must run constantly. In EPS, the assist level is easily tunable to the vehicle type, road speed, and even driver preference. An added benefit is the elimination of environmental hazard posed by leakage and disposal of hydraulic power steering fluid. Also in the event of the engine cutting out, assist will not be lost - whereas hydraulic will stop working, as well as making the steering doubly heavy as the driver has to turn the power-assist mechanism on top of the steering system itself.

Speed Sensitive Steering

An outgrowth of power steering is speed sensitive steering, where the steering is heavily assisted at low speed and lightly assisted at high speed. The auto makers perceive that motorists might need to make large steering inputs while manoeuvering for parking, but not while traveling at high speed. The first vehicle with this feature was the Citroën SM with its Diravi layout, although rather than altering the amount of assistance as in modern power steering systems, it altered the pressure on a centring cam which made the steering wheel try to "spring" back to the straight-ahead position. Modern speed-sensitive power steering systems reduce the mechanical or electrical assistance as the vehicle speed increases, giving a more direct feel. This feature is gradually becoming more common.

Four-wheel steering

Four-wheel steering (or all-wheel steering) is a system employed by some vehicles to improve steering response, increase vehicle stability while maneuvering at high speed, or to decrease turning radius at low speed.

Sierra Denali with Quadrasteer, rear steering angle

In most *active* four-wheel steering systems, the rear wheels are steered by a computer and actuators. The rear wheels generally cannot turn as far as the front wheels. Some systems, including Delphi's Quadrasteer and the system in Honda's Prelude line, allow the rear wheels to be steered in the opposite direction as the front wheels during low speeds. This allows the vehicle to turn in a significantly smaller radius — sometimes critical for large trucks or tractors and vehicles with trailers.

Many modern vehicles offer a form of *passive* rear steering to counteract normal vehicle tendencies. For example, Subaru used a passive steering system to correct for the rear wheel's tendency to toe-out. On many vehicles, when cornering, the rear wheels tend to steer slightly to the outside of a turn, which can reduce stability. The passive steering system uses the lateral forces generated in a turn (through suspension geometry) and the bushings to correct this tendency and steer the wheels slightly to the inside of the corner. This improves the stability of the car, through the turn. This effect is called compliance understeer and it, or its opposite, is present on all suspensions. Typical methods of achieving compliance understeer are to use a Watt's Link on a live rear axle, or the use of toe control bushings on a twist beam suspension. On an independent rear suspension it is normally achieved by changing the rates of the rubber bushings in the suspension. Some suspensions will always have compliance oversteer due to geometry, such as Hotchkiss live axles or a semi-trailing arm IRS.

Passive rear wheel steering is not a new concept, as it has been in use for many years, although not always recognised as such. For example, Jaguar independent rear suspension incorporated a small amount of passive rear wheel steering since 1961.

Recent application

In an active four-wheel steering system, all four wheels turn at the same time when the driver steers. There can be controls to switch off the rear steer and options to steer only the rear wheel independent of the front wheels. At slow speeds (*e.g.* parking) the rear wheels turn opposite of the front wheels, reducing the turning radius by up to twenty-five percent, while at higher speeds both front and rear wheels turn alike (electronically controlled), so that the vehicle may change position with less yaw, enhancing straight-line stability. The "Snaking effect" experienced during motorway drives while towing a travel trailer is thus largely nullified. Four-wheel steering found its most widespread use in monster trucks, where maneuverability in small arenas is critical, and it is also popular in large farm vehicles and trucks. Some of the modern European Intercity buses also utilize four-wheel steering to assist maneuverability in bus terminals, and also to improve road stability.

General Motors offers Delphi's Quadrasteer in their consumer Silverado/Sierra and Suburban/Yukon. However, only 16,500 vehicles have been sold with this system since its introduction in 2002 through 2004. Due to this low demand, GM discontinued the technology at the end of the 2005 model year.

Previously, Honda had four-wheel steering as an option in their 1987-2000 Prelude, and Mazda also offered four-wheel steering on the 626 and MX6 in 1988.

A new "Active Drive" system is introduced on the 2008 version of the Renault Laguna line. It was designed as one of several measures to increase security and stability. The Active Drive should lower the effects of under steer and decrease the chances of spinning by diverting part of the G-forces generated in a turn from the front to the rear tires. At low speeds the turning circle can be tightened so parking and maneuvering is easier.

Articulated steering

A front loader with articulated steering.

Articulated steering is a system by which a four-wheel drive vehicle is split into front and rear halves which are connected by a vertical hinge. The front and rear halves are connected with one or more hydraulic cylinders that change the angle between the halves, including the front and rear axles and wheels, thus steering the vehicle. This system does not use steering arms, king pins, tie rods, etc. as does four-wheel steering. If the vertical hinge is placed equidistant between the two axles, it also eliminates the need for a central differential, as both front and rear axles will follow the same path, and thus rotate at the same speed.

Rear wheel steering

A few types of vehicle use rear wheel steering, notably fork lift trucks, early pay loaders, Buckminster Fuller's Dymaxion car, and the ThrustSSC.

Rear wheel steering can tend to be unstable because in turns the steering geometry tends to decrease the turn radius (oversteer), rather than increase it (understeer). A rear wheel steered automobile exhibits non-minimum phase behavior. It turns in the direction opposite of how it is initially steered. A rapid steering input will cause two accelerations, first in the direction that the wheel is steered, and then in the opposite direction: a

"reverse response." This makes it harder to steer a rear wheel steered vehicle at high speed than a front wheel steered vehicle.

Steer-by-wire

The aim of *steer-by-wire* technology is to completely do away with as many mechanical components (steering shaft, column, gear reduction mechanism, etc.) as possible. Completely replacing conventional steering system with steer-by-wire holds several advantages, such as:

- The absence of steering column simplifies the car interior design.
- The absence of steering shaft, column and gear reduction mechanism allows much better space utilization in the engine compartment.
- The steering mechanism can be designed and installed as a modular unit.
- Without mechanical connection between the steering wheel and the road wheel, it is less likely that the impact of a frontal crash will force the steering wheel to intrude into the driver's survival space.
- Steering system characteristics can easily and infinitely be adjusted to optimize the steering response and feel.

As of 2007 there are no production cars available that rely solely on steer-by-wire technology due to safety, reliability and economic concerns, but this technology has been demonstrated in numerous concept cars and the similar *fly-by-wire* technology is in use in both military and civilian aviation applications. Removing the mechanical steering linkage in road going vehicles would require new legislation in most countries.

Safety

For safety reasons all modern cars feature a collapsible steering column (energy absorbing steering column) which will collapse in the event of a heavy frontal impact to avoid excessive injuries to the driver. Airbags are also generally fitted as standard. Non-collapsible steering columns fitted to older vehicles very often impaled drivers in frontal crashes, particularly when the steering box or rack was mounted in front of the front axle line, at the front of the crumple zone. This was particularly a problem on vehicles that had a rigid separate chassis frame, with no crumple zone. Most modern vehicle steering boxes/racks are mounted behind the front axle on the front bulkhead, at the rear of the front crumple zone.

Audi used a retractable steering wheel and seat belt tensioning system called procon-ten, but it has since been discontinued in favor of airbags and pyrotechnic seat belt pre-tensioners.

Collapsible steering columns were invented by Bela Barenyi and were introduced in the 1959 Mercedes-Benz W111 Fintail, along with Crumple zones.

This safety feature first appeared on cars built by General Motors after an extensive and very public lobbying campaign enacted by Ralph Nader.

Ford started to install collapsible steering columns in 1968.

Cycles

Steering is crucial to the stability of bicycles and motorcycles.

Ship and boat steering

Ships and boats are usually steered with a rudder. Depending on the size of the vessel, rudders can be manually actuated, or operated using a servomechanism, or a trim tab/servo tab system.

Chapter- 5

Suspension (Vehicle)

The front suspension components of a Ford Model T

The rear suspension on a truck: a leaf spring.

Part of car front suspension and steering mechanism: tie rod, steering arm, king pin axis (using ball joints).

Suspension is the term given to the system of springs, shock absorbers and linkages that connects a vehicle to its wheels. Suspension systems serve a dual purpose — contributing to the car's roadholding/handling and braking for good active safety and driving pleasure, and keeping vehicle occupants comfortable and reasonably well isolated from road noise, bumps, and vibrations,etc. These goals are generally at odds, so the tuning of suspensions involves finding the right compromise. It is important for the suspension to keep the road wheel in contact with the road surface as much as possible, because all the forces acting on the vehicle do so through the contact patches of the tires. The suspension also protects the vehicle itself and any cargo or luggage from damage and wear. The design of front and rear suspension of a car may be different.

History

Leaf springs have been around since the early Egyptians.

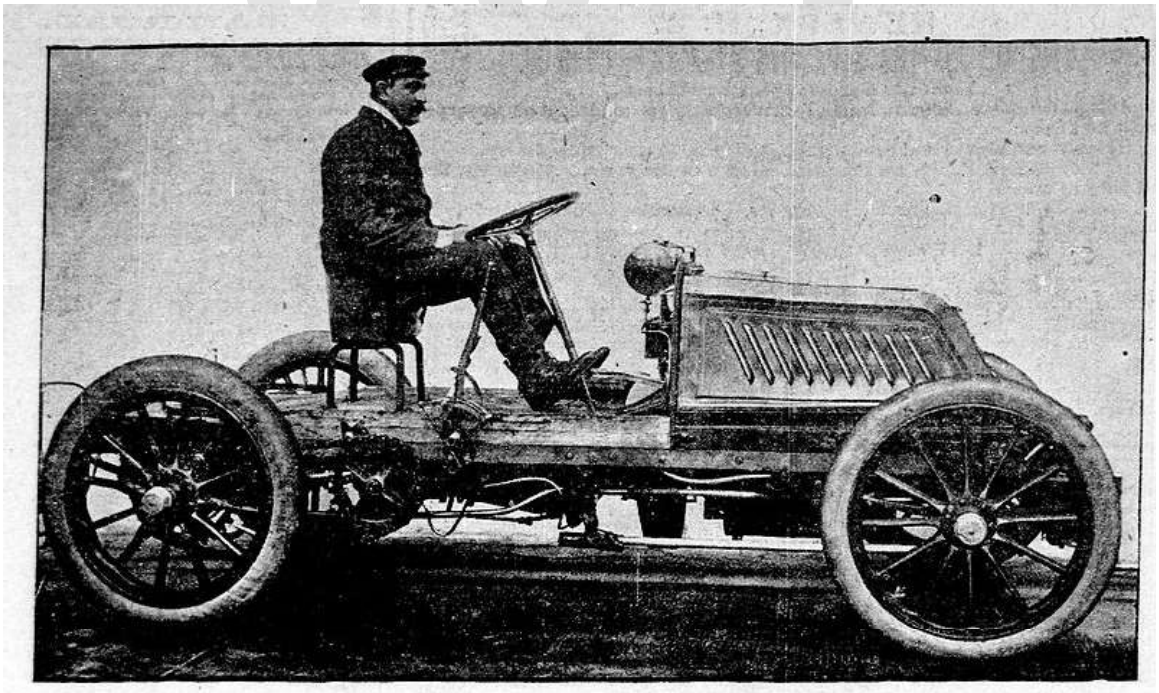
Ancient military engineers used leaf springs in the form of bows to power their siege engines, with little success at first. The use of leaf springs in catapults was later refined and made to work years later. Springs were not only made of metal, a sturdy tree branch could be used as a spring, such as with a bow.

Horse drawn vehicles

By the early 19th century, most British horse carriages were equipped with springs; wooden springs in the case of light one-horse vehicles to avoid taxation, and steel springs in larger vehicles. These were made of low-carbon steel and usually took the form of multiple layer leaf springs.

The British steel springs were not well suited for use on America's rough roads of the time, and could even cause coaches to collapse if cornered too fast. In the 1820s, the Abbot Downing Company of Concord, New Hampshire developed a system whereby the bodies of stagecoaches were supported on leather straps called "thoroughbraces", which gave a swinging motion instead of the jolting up and down of a spring suspension (the stagecoach itself was sometimes called a "thoroughbrace").

Automobiles



FOURNIER ON THE "MORS" MACHINE WITH WHICH HE WON THE PARIS-BORDEAUX AND PARIS-BERLIN RACES AND BEAT THE VANDERBILT RECORD FOR ONE KILOMETRE.

Henri Fournier on his uniquely dampened and racewinning 'Mors Machine', photo taken 1902

Automobiles were initially developed as self-propelled versions of horse drawn vehicles. However, horse drawn vehicles had been designed for relatively slow speeds and their suspension was not well suited to the higher speeds permitted by the internal combustion engine.

In 1901 Mors of Germany first fitted an automobile with shock absorbers. With the advantage of having a dampened suspension system in his 'Mors Machine', Henri Fournier was able to win the prestigegous Paris — Berlin race on June 20, 1901. Fourniers superior time was 11 hrs 46 min 10 sec, while the best competitor was Léonce Girardot in a Panhard at the time 12 hrs 15 min 40 sec.

In 1920, Leyland used torsion bars in a suspension system. In 1922, independent front suspension was pioneered on the Lancia Lambda and became more common in mass market cars from 1932.

Important properties

Citroën BX Hydropneumatic suspension - maximum to minimum demonstration

Spring rate

The spring rate (or suspension rate) is a component in setting the vehicle's ride height or its location in the suspension stroke. Vehicles which carry heavy loads will often have

heavier springs to compensate for the additional weight that would otherwise collapse a vehicle to the bottom of its travel (stroke). Heavier springs are also used in performance applications where the loading conditions experienced are more extreme.

Springs that are too hard or too soft cause the suspension to become ineffective because they fail to properly isolate the vehicle from the road. Vehicles that commonly experience suspension loads heavier than normal have heavy or hard springs with a spring rate close to the upper limit for that vehicle's weight. This allows the vehicle to perform properly under a heavy load when control is limited by the inertia of the load. Riding in an empty truck used for carrying loads can be uncomfortable for passengers because of its high spring rate relative to the weight of the vehicle. A race car would also be described as having heavy springs and would also be uncomfortably bumpy. However, even though we say they both have heavy springs, the actual spring rates for a 2,000 lb (910 kg) race car and a 10,000 lb (4,500 kg) truck are very different. A luxury car, taxi, or passenger bus would be described as having soft springs. Vehicles with worn out or damaged springs ride lower to the ground which reduces the overall amount of compression available to the suspension and increases the amount of body lean. Performance vehicles can sometimes have spring rate requirements other than vehicle weight and load.

Mathematics of the spring rate

Spring rate is a ratio used to measure how resistant a spring is to being compressed or expanded during the spring's deflection. The magnitude of the spring force increases as deflection increases according to Hooke's Law. Briefly, this can be stated as

$$F = -kx$$

where

F is the force the spring exerts
k is the spring rate of the spring.
x is the displacement from equilibrium length i.e. the length at which the spring is neither compressed or stretched.

Spring rate is confined to a narrow interval by the weight of the vehicle,load the vehicle will carry, and to a lesser extent by suspension geometry and performance desires.

Spring rates typically have units of N/mm (or lbf/in). An example of a linear spring rate is 500 lbf/in. For every inch the spring is compressed, it exerts 500 lbf. A non-linear spring rate is one for which the relation between the spring's compression and the force exerted cannot be fitted adequately to a linear model. For example, the first inch exerts 500 lbf force, the second inch exerts an additional 550 lbf (for a total of 1050 lbf), the third inch exerts another 600 lbf (for a total of 1650 lbf). In contrast a 500 lbf/in linear spring compressed to 3 inches will only exert 1500 lbf.

The spring rate of a coil spring may be calculated by a simple algebraic equation or it may be measured in a spring testing machine. The spring constant *k* can be calculated as follows:

$$k = \frac{d^4 G}{8 N D^3}$$

where *d* is the wire diameter, G is the spring's shear modulus (e.g., about 12,000,000 lbf/in² or 80 GPa for steel), and *N* is the number of wraps and *D* is the diameter of the coil.

Wheel rate

Wheel rate is the effective spring rate when measured at the wheel. This is as opposed to simply measuring the spring rate alone.

Wheel rate is usually equal to or considerably less than the spring rate. Commonly, springs are mounted on control arms, swing arms or some other pivoting suspension member. Consider the example above where the spring rate was calculated to be 500 lbs/inch, if you were to move the wheel 1 in (2.5 cm) (without moving the car), the spring more than likely compresses a smaller amount. Lets assume the spring moved 0.75 in (19 mm), the lever arm ratio would be 0.75:1. The wheel rate is calculated by taking the square of the ratio (0.5625) times the spring rate. Squaring the ratio is because the ratio has two effects on the wheel rate. The ratio applies to both the force and distance traveled.

Wheel rate on independent suspension is fairly straight-forward. However, special consideration must be taken with some non-independent suspension designs. Take the case of the straight axle. When viewed from the front or rear, the wheel rate can be measured by the means above. Yet because the wheels are not independent, when viewed from the side under acceleration or braking the pivot point is at infinity (because both wheels have moved) and the spring is directly inline with the wheel contact patch. The result is often that the effective wheel rate under cornering is different from what it is under acceleration and braking. This variation in wheel rate may be minimized by locating the spring as close to the wheel as possible.

Roll couple percentage

Roll couple percentage is the effective wheel rate, in roll, of each axle of the vehicle as a ratio of the vehicle's total roll rate. Roll couple percentage is critical in accurately balancing the handling of a vehicle. It is commonly adjusted through the use of anti-roll bars, but can also be changed through the use of different springs.

A vehicle with a roll couple percentage of 70% will transfer 70% of its sprung weight at the front of the vehicle during cornering. This is also commonly known as "Total Lateral Load Transfer Distribution" or "TLLTD".

Weight transfer

Weight transfer during cornering, acceleration or braking is usually calculated per individual wheel and compared with the static weights for the same wheels.

The total amount of weight transfer is only affected by four factors: the distance between wheel centers (wheelbase in the case of braking, or track width in the case of cornering) the height of the center of gravity, the mass of the vehicle, and the amount of acceleration experienced.

The speed at which weight transfer occurs as well as through which components it transfers is complex and is determined by many factors including but not limited to roll center height, spring and damper rates, anti-roll bar stiffness and the kinematic design of the suspension links.

Unsprung weight transfer

Unsprung weight transfer is calculated based on the weight of the vehicle's components that are not supported by the springs. This includes tires, wheels, brakes, spindles, half the control arm's weight and other components. These components are then (for calculation purposes) assumed to be connected to a vehicle with zero sprung weight. They are then put through the same dynamic loads. The weight transfer for cornering in the front would be equal to the total unsprung front weight times the G-Force times the front unsprung center of gravity height divided by the front track width. The same is true for the rear.

Sprung weight transfer

Sprung weight transfer is the weight transferred by only the weight of the vehicle resting on the springs, not the total vehicle weight. Calculating this requires knowing the vehicle's sprung weight (total weight less the unsprung weight), the front and rear roll center heights and the sprung center of gravity height (used to calculate the roll moment arm length). Calculating the front and rear sprung weight transfer will also require knowing the roll couple percentage.

The roll axis is the line through the front and rear roll centers that the vehicle rolls around during cornering. The distance from this axis to the sprung center of gravity height is the roll moment arm length. The total sprung weight transfer is equal to the G-force times the sprung weight times the roll moment arm length divided by the effective track width. The front sprung weight transfer is calculated by multiplying the roll couple percentage times the total sprung weight transfer. The rear is the total minus the front transfer.

Jacking forces

Jacking forces are the sum of the vertical force components experienced by the suspension links. The resultant force acts to lift the sprung mass if the roll center is above

ground, or compress it if underground. Generally, the higher the roll center, the more jacking force is experienced.

Travel

Travel is the measure of distance from the bottom of the suspension stroke (such as when the vehicle is on a jack and the wheel hangs freely) to the top of the suspension stroke (such as when the vehicle's wheel can no longer travel in an upward direction toward the vehicle). Bottoming or lifting a wheel can cause serious control problems or directly cause damage. "Bottoming" can be caused by the suspension, tires, fenders, etc. running out of space to move or the body or other components of the car hitting the road. The control problems caused by lifting a wheel are less severe if the wheel lifts when the spring reaches its unloaded shape than they are if travel is limited by contact of suspension members. Many off-road vehicles, such as desert racers, use straps called "limiting straps" to limit the suspensions downward travel to a point within safe limits for the linkages and shock absorbers. This is necessary, since these trucks are intended to travel over very rough terrain at high speeds, and even become airborne at times. Without something to limit the travel, the suspension bushings would take all the force when the suspension reaches "full droop", and it can even cause the coil springs to come out of their "buckets" if they are held in by compression forces only. A limiting strap is a simple strap, often nylon of a predetermined length, that stops the downward movement at a preset point before the theoretical maximum travel is reached. The opposite of this is the "bump-stop", which protects the suspension and vehicle (as well as the occupants) from violent "bottoming" of the suspension, caused when an obstruction (or hard landing) causes the suspension to run out of upward travel without fully absorbing the energy of the stroke. Without bump-stops, a vehicle that "bottoms out" will experience a very hard shock when the suspension contacts the bottom of the frame or body, which is transferred to the occupants and every connector and weld on the vehicle. Factory vehicles often come with plain rubber "nubs" to absorb the worst of the forces, and insulate the shock. A desert race vehicle, which must routinely absorb far higher impact forces, may be provided with pneumatic or hydro-pneumatic bump-stops. These are essentially miniature shock absorbers (dampeners) that are fixed to the vehicle in a location such that the suspension will contact the end of the piston when it nears the upward travel limit. These absorb the impact far more effectively than a solid rubber bump-stop will, essential because a rubber bump-stop is considered a "last-ditch" emergency insulator for the occasional accidental bottoming of the suspension; it is entirely insufficient to absorb repeated and heavy bottomings such as a high-speed off road vehicle encounters.

Damping

Damping is the control of motion or oscillation, as seen with the use of hydraulic gates and valves in a vehicles shock absorber. This may also vary, intentionally or unintentionally. Like spring rate, the optimal damping for comfort may be less than for control.

Damping controls the travel speed and resistance of the vehicle's suspension. An undamped car will oscillate up and down. With proper damping levels, the car will settle back to a normal state in a minimal amount of time. Most damping in modern vehicles can be controlled by increasing or decreasing the resistance to fluid flow in the shock absorber.

Camber control

Camber changes due to wheel travel, body roll and suspension system deflection or compliance. In general, a tire wears and brakes best at -1 to -2° of camber from vertical. Depending on the tire and the road surface, it may hold the road best at a slightly different angle. Small changes in camber, front and rear, can be used to tune handling. Some race cars are tuned with -2~-7° camber depending on the type of handling desired and the tire construction. Oftentimes, too much camber will result in the decrease of braking performance due to a reduced contact patch size through excessive camber variation in the suspension geometry. The amount of camber change in bump is determined by the instantaneous front view swing arm (FVSA) length of the suspension geometry, or in other words, the tendency of the tire to camber inward when compressed in bump.

Roll center height

This is important to body roll and to front to rear roll stiffness distribution. However, the roll stiffness distribution in most cars is set more by the antiroll bars than the RCH. The height of the roll center is related to the amount of jacking forces experienced.

Instant center

Due to the fact that the wheel and tire's motion is constrained by the suspension links on the vehicle, the motion of the wheel package in the front view will scribe an imaginary arc in space with an “instantaneous center" of rotation at any given point along its path. The instant center for any wheel package can be found by following imaginary lines drawn through the suspension links to their intersection point.

A component of the tire's force vector points from the contact patch of the tire through instant center. The larger this component is, the less suspension motion will occur. Theoretically, if the resultant of the vertical load on the tire and the lateral force generated by it points directly into the instant center, the suspension links will not move. In this case, all weight transfer at that end of the vehicle will be geometric in nature. This is key information used in finding the force-based roll center as well.

In this respect the instant centers are more important to the handling of the vehicle than the kinematic roll center alone, in that the ratio of geometric to elastic weight transfer is determined by the forces at the tires and their directions in relation to the position of their respective instant centers.

No-dive and No-squat

Anti-dive and anti-squat are percentages and refer to the front diving under braking and the rear squatting under acceleration. They can be thought of as the counterparts for braking and acceleration as jacking forces are to cornering. The main reason for the difference is due to the different design goals between front and rear suspension, whereas suspension is usually symmetrical between the left and right of the vehicle.

The method of determining the anti-dive or anti-squat depends on whether the suspension linkages react to the torque of braking and accelerating. For example, with inboard brakes and half-shaft driven rear wheels, the suspension linkages do not, but with outboard brakes and a swing-axle driveline, they do.

To determine the percentage of front suspension braking anti-dive for outboard brakes, it is first necessary to determine the tangent of the angle between a line drawn, in side view, through the front tire patch and the front suspension instant center, and the horizontal. In addition, the percentage of braking effort at the front wheels must be known. Then, multiply the tangent by the front wheel braking effort percentage and divide by the ratio of the center of gravity height to the wheelbase. A value of 50% would mean that half of the weight transfer to the front wheels, during braking, is being transmitted through the front suspension linkage and half is being transmitted through the front suspension springs.

For inboard brakes, the same procedure is followed but using the wheel center instead of contact patch center.

Forward acceleration anti-squat is calculated in a similar manner and with the same relationship between percentage and weight transfer. Anti-squat values of 100% and more are commonly used in dragracing, but values of 50% or less are more common in cars which have to undergo severe braking. Higher values of anti-squat commonly cause wheel hop during braking. It is important to note that, while the value of 100%...in either case...means that all of the weight transfer is being carried through the suspension linkage, this does not mean that the suspension is incapable of carrying additional loads (aerodynamic, cornering, etc.) during an episode of braking or forward acceleration. In other words, no "binding" of the suspension is to be implied.

Flexibility and vibration modes of the suspension elements

In modern cars, the flexibility is mainly in the rubber bushings. For high-stress suspensions, such as off-road vehicles, polyurethane bushings are available, which offer far more longevity under greater stresses.

Isolation from high frequency shock

For most purposes, the weight of the suspension components is unimportant, but at high frequencies, caused by road surface roughness, the parts isolated by rubber bushings act

as a multistage filter to suppress noise and vibration better than can be done with only the tires and springs. (The springs work mainly in the vertical direction.)

Contribution to unsprung weight and total weight

These are usually small, except that the suspension is related to whether the brakes and differential(s) are sprung.

Space occupied

Designs differ as to how much space they take up and where it is located. It is generally accepted that MacPherson struts are the most compact arrangement for front-engined vehicles, where space between the wheels is required to place the engine.

Force distribution

The suspension attachment must match the frame design in geometry, strength and rigidity.

Air resistance (drag)

Certain modern vehicles have height adjustable suspension in order to improve aerodynamics and fuel efficiency. And modern formula cars, that have exposed wheels and suspension, typically use streamlined tubing rather than simple round tubing for their suspension arms to reduce drag. Also typical is the use of rocker arm, push rod, or pull rod type suspensions, that among other things, places the spring/damper unit inboard and out of the air stream to further reduce air resistance.

Cost

Production methods improve, but cost is always a factor. The continued use of the solid rear axle, with unsprung differential, especially on heavy vehicles, seems to be the most obvious example.

Springs and dampers

Most conventional suspensions use passive springs to absorb impacts and dampers (or shock absorbers) to control spring motions.

Some notable exceptions are the hydropneumatic systems, which can be treated as an integrated unit of gas spring and damping components, used by the French manufacturer Citroën and the hydrolastic, hydragas and rubber cone systems used by the British Motor Corporation, most notably on the Mini. A number of different types of each have been used:

Passive suspensions

Traditional springs and dampers are referred to as passive suspensions — most vehicles are suspended in this manner.

Springs

Pneumatic spring on a semitrailer

- Leaf spring – AKA Hotchkiss, Cart, or semi-elliptical spring
- Torsion beam suspension
- Coil spring
- Rubber bushing
- Air spring

Dampers or shock absorbers

The shock absorbers damp out the (otherwise resonant) motions of a vehicle up and down on its springs. They also must damp out much of the wheel bounce when the unsprung weight of a wheel, hub, axle and sometimes brakes and differential bounces up and down on the springiness of a tire. The regular bumps found on dirt roads (nicknamed "corduroy", but properly corrugations or washboarding) are caused by this wheel bounce.

Semi-active and active suspensions

If the suspension is externally controlled then it is a semi-active or active suspension — the suspension is reacting to what are in effect "brain" signals. As electronics have become more sophisticated, the opportunities in this area have expanded.

For example, a hydropneumatic Citroën will "know" how far off the ground the car is supposed to be and constantly reset to achieve that level, regardless of load. It will *not* instantly compensate for body roll due to cornering however. Citroën's system adds about 1% to the cost of the car versus passive steel springs.

Semi-active suspensions include devices such as air springs and switchable shock absorbers, various self-levelling solutions, as well as systems like Hydropneumatic, Hydrolastic, and Hydragas suspensions. Mitsubishi developed the world's first production semi-active electronically controlled suspension system in passenger cars; the system was first incorporated in the 1987 Galant model. Delphi currently sells shock absorbers filled with a magneto-rheological fluid, whose viscosity can be changed electromagnetically, thereby giving variable control without switching valves, which is faster and thus more effective.

Fully active suspension systems use electronic monitoring of vehicle conditions, coupled with the means to impact vehicle suspension and behavior in real time to directly control the motion of the car. Lotus Cars developed several prototypes, from 1982 onwards, and introduced them to F1, where they have been fairly effective, but have now been banned. Nissan introduced a low bandwidth active suspension in circa 1990 as an option that added an extra 20% to the price of luxury models. Citroën has also developed several active suspension models. A recently publicised fully active system from Bose Corporation uses linear electric motors, i.e. solenoids, in place of hydraulic or pneumatic actuators that have generally been used up until recently. The most advanced suspension system is Active Body Control, introduced in 1999 on the top-of-the-line Mercedes-Benz CL-Class.

Several electromagnetic suspensions have also been developed for vehicles. Examples include the electromagnetic suspension of Bose, and the electromagnetic suspension developed by prof. Laurentiu Encica. In addition, the new Michelin wheel with embedded suspension working on a electromotor is also similar.

With the help of control system, various semi-active/active suspensions realize an improved design compromise among different vibrations modes of the vehicle, namely bounce, roll, pitch and warp modes. However, the applications of these advanced suspensions are constrained by the cost, packaging, weight, reliability, and/or the other challenges.

Interconnected suspensions

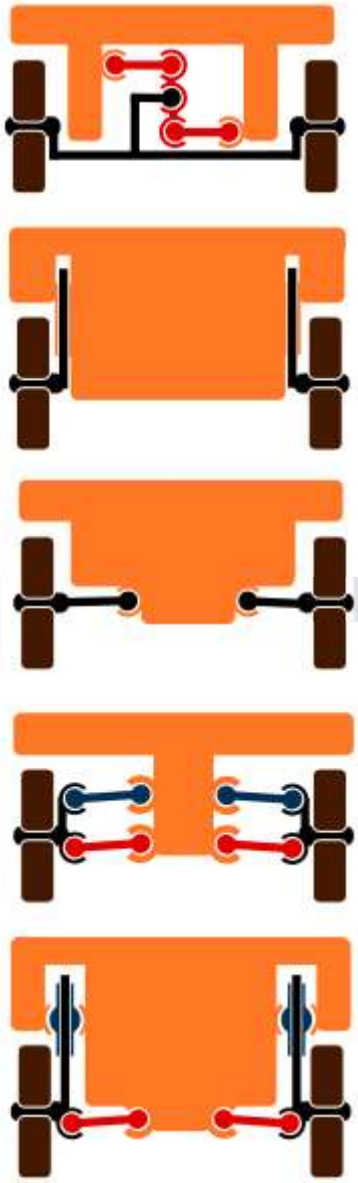
Interconnected suspension, unlike semi-active/active suspensions, could easily decouple different vehicle vibration modes in a passive manner. The interconnections can be realized by various means, such as mechanical, hydraulic and pneumatic. Anti-roll bars are one of the typical examples of mechanical interconnections, while it has been stated that fluidic interconnections offer greater potential and flexibility in improving both the stiffness and damping properties.

Considering the considerable commercial potentials of hydro-pneumatic technology (Corolla, 1996), interconnected hydropneumatic suspensions have also been explored in some recent studies, and their potential benefits in enhancing vehicle ride and handling have been demonstrated. The control system can also be used for further improving performance of interconnected suspensions. Apart from academic research, an Australian company, Kinetic, is having some success (WRC: 3 Championships, Dakar Rally: 2 Championships, Lexus GX470 2004 4x4 of the year with KDSS, 2005 PACE award) with various passive or semi-active systems, which generally decouple at least two vehicle modes (roll, warp (articulation), pitch and/or heave (bounce)) to simultaneously control each mode's stiffness and damping, by using interconnected shock absorbers, and other methods. In 1999, Kinetic was bought out by Tenneco.

Historically, the first mass production car with front to rear mechanical interconnected suspension was the 1948 Citroën 2CV. The suspension of the 2CV was extremely soft — it had low roll stiffness, but its pitch stiffness was increased by using an interconnected suspension. The leading arm / trailing arm swinging arm, fore-aft linked suspension system together with inboard front brakes had a much smaller unsprung weight than existing coil spring or leaf designs. The interconnection transmitted some of the force deflecting a front wheel up over a bump, to push the rear wheel down on the same side. When the rear wheel met that bump a moment later, it did the same in reverse, keeping the car level front to rear. The 2CV had a design brief to be able to be driven at speed over a ploughed field. It originally featured friction dampers and tuned mass dampers. Later models had tuned mass dampers at the front with telescopic dampers / shock absorbers front and rear.

Some of the last post-war Packard models also featured interconnected suspension. The original Mini and some more recent British Leyland models also featured interlinking, when fitted with Moulton's Hydrolastic or Hydragas suspensions.

Suspension Geometry



Common types seen from behind. From top to bottom: live axle with Watt bar, suspension like on a bike fork, swing axle, double wishbone, MacPherson. Some types are missing because trailing arm links are not presentable in this view and some types use elements which flex to some movements and are stiff to others and flexible elements are omitted for clarity.

Suspension systems can be broadly classified into two subgroups — dependent and independent. These terms refer to the ability of opposite wheels to move independently of each other.

A *dependent suspension* normally has a beam (a simple 'cart' axle) or (driven) live axle that holds wheels parallel to each other and perpendicular to the axle. When the camber

of one wheel changes, the camber of the opposite wheel changes in the same way (by convention on one side this is a positive change in camber and on the other side this a negative change). De Dion suspensions are also in this category as they rigidly connect the wheels together.

An *independent suspension* allows wheels to rise and fall on their own without affecting the opposite wheel. Suspensions with other devices, such as sway bars that link the wheels in some way are still classed as independent.

A third type is a *semi-dependent* suspension. In this case, the motion of one wheel does affect the position of the other but they are not rigidly attached to each other. A twist-beam rear suspension is such a system.

Dependent suspensions

Dependent systems may be differentiated by the system of linkages used to locate them, both longitudinally and transversely. Often both functions are combined in a set of linkages.

Examples of location linkages include:

- Satchell link
- Panhard rod
- Watt's linkage
- WOBLink
- Mumford linkage
- Leaf springs used for location (transverse or longitudinal)
    - Fully elliptical springs usually need supplementary location links and are no longer in common use
    - Longitudinal semi-elliptical springs used to be common and still are used in heavy-duty trucks and aircraft. They have the advantage that the spring rate can easily be made progressive (non-linear).
    - A single transverse leaf spring for both front wheels and/or both back wheels, supporting solid axles, was used by Ford Motor Company, before and soon after World War II, even on expensive models. It had the advantages of simplicity and low unsprung weight (compared to other solid axle designs).

In a front engine, rear-drive vehicle, dependent rear suspension is either "live axle" or deDion axle, depending on whether or not the differential is carried on the axle. Live axle is simpler but the unsprung weight contributes to wheel bounce.

Because it assures constant camber, dependent (and semi-independent) suspension is most common on vehicles that need to carry large loads as a proportion of the vehicle weight, that have relatively soft springs and that do not (for cost and simplicity reasons)

use active suspensions. The use of dependent front suspension has become limited to heavier commercial vehicles.

A rear independent suspension on an AWD car.

Semi-independent suspension

In a semi-independent suspensions, the wheels of an axle are able to move relative to one another as in an independent suspension but the position of one wheel has an effect on the position and attitude of the other wheel. This effect is achieved via the twisting or deflecting of suspension parts under load. The most common type of semi-independent suspension is the twist beam.

- Twist beam

Independent suspension

The variety of independent systems is greater and includes:

- Swing axle
- Sliding pillar
- MacPherson strut/Chapman strut
- Upper and lower A-arm (double wishbone)

- multi-link suspension
- semi-trailing arm suspension
- swinging arm
- leaf springs
    - Transverse leaf springs when used as a suspension link, or four quarter elliptics on one end of a car are similar to wishbones in geometry, but are more compliant. Examples are the front of the original Fiat 500, the Panhard Dyna Z and the early examples of Peugeot 403 and the back of the AC Ace and AC Aceca.

Because the wheels are not constrained to remain perpendicular to a flat road surface in turning, braking and varying load conditions, control of the wheel camber is an important issue. Swinging arm was common in small cars that were sprung softly and could carry large loads, because the camber is independent of load. Some active and semi-active suspensions maintain the ride height, and therefore the camber, independent of load. In sports cars, optimal camber change when turning is more important.

Wishbone and multi-link allow the engineer more control over the geometry, to arrive at the best compromise, than swing axle, MacPherson strut or swinging arm do; however the cost and space requirements may be greater. Semi-trailing arm is in between, being a variable compromise between the geometries of swinging arm and swing axle.

Armoured fighting vehicle suspension

This Grant I tank's suspension has road wheels mounted on wheel trucks, or *bogies*.

Military AFVs, including tanks, have specialized suspension requirements. They can weigh more than seventy tons and are required to move at high speed over very rough ground. Their suspension components must be protected from land mines and antitank weapons. Tracked AFVs can have as many as nine road wheels on each side. Many wheeled AFVs have six or eight wheels, to help them ride over rough and soft ground.

The earliest tanks of World War I had fixed suspension with no movement whatsoever. This unsatisfactory situation was improved with leaf spring or coil spring suspensions adopted from agricultural, automotive or railway machinery, but even these had very limited travel.

Speeds increased due to more powerful engines, and the quality of ride had to be improved. In the 1930s, the Christie suspension was developed, which allowed the use of coil springs inside a vehicle's armored hull, by changing the direction of force deforming the spring, using a bell crank. Horstmann suspension was a variation which used a combination of bell crank and exterior coil springs, in use from the 1930s to the 1990s.

By World War II the other common type was torsion-bar suspension, getting spring force from twisting bars inside the hull — this had less travel than the Christie-type, but was significantly more compact, allowing more space inside the hull, with consequent possibility to install larger turret rings and thus a heavier main armament. The torsion-bar suspension, sometimes including shock absorbers, has been the dominant heavy armored vehicle suspension since World War II.

Chapter- 6

Understeer

Understeer and **oversteer** are vehicle dynamics terms used to describe the sensitivity of a vehicle to steering. Automotive engineers originally defined understeer and oversteer based on the gradient of the steering needed to make a turn in a steady-state condition (constant speed, constant radius) on a flat and level ground surface. Car and motorsport enthusiasts often use the terminology more generally in magazines and blogs to describe vehicle response to steering in all kinds of maneuvers, even on banked turns. Simply put, oversteer is what occurs when a car turns(steers) by more than(over) the amount commanded by the driver. Conversely, understeer is what occurs when a car steers under the amount commanded by the driver.

Vehicle Dynamics Terminology

In standard terminology defined by the Society of Automotive Engineers (SAE) J670 and the International Organization for Standardization (ISO) 8855, understeer and oversteer are based on differences in steady-state conditions where the vehicle is following a constant-radius path at a constant speed with a constant steering wheel angle, on a flat and level surface. If the speed is increased slightly for the same radius path and, after settling into steady state, the same steering is measured, then the vehicle is said to have neutral steer. If more steering is needed at the higher speed to maintain the same radius of curvature, then the vehicle is said to have understeer. If less steering is needed at the higher speed, then the vehicle is said to have oversteer.

Understeer and oversteer are defined by an understeer gradient U that is the difference between a reference steer angle gradient and the Ackerman steer angle gradient. The reference steer angle (δ_R) is the average steer of the front axle wheels minus the average steer of the rear axle wheels. The Ackerman steer angle (δ_A) is defined for a given radius of turn as the reference steer angle that would be used at a very low speed. For a four-wheeled vehicle with steering only at the front wheels, the Ackerman angle δ_A (at the wheels) is the arctangent of the wheelbase divided by the turn radius (at the center of the rear axle).

Understeer and oversteer are formally defined using the gradient U: if U is positive, the vehicle is understeer; if U is negative, the vehicle is oversteer; if U is zero, the vehicle is neutral.

Different companies and organizations have different test procedures for defining U. In all cases, the gradient is taken by comparing measures from steady state tests, and expressed with units of degrees of steer (at the road wheels) divided by lateral acceleration Ay expressed in g's. In steady-state conditions, $Ay = V^2/R/G$, where V is the vehicle speed, R is the radius of the turn, and G is the gravitational scaling factor.

SAE J670 describes three methods for measuring U:

1. Constant radius: tests are repeated at different speeds for a given constant-radius track. In this kind of procedure, the Ackerman steering is always the same, so the gradient is: $U = d(\delta_R)/d(Ay)$
2. Constant steer angle: tests are repeated at different speeds for a given reference steer angle. In this kind of procedure, the reference steer is always the same so the gradient is: $U = -d(\delta_A)/d(Ay)$
3. Constant speed: tests are repeated with different reference steer angles for a given speed. In this kind of procedure, the gradient is: $U = d(\delta_R)/d(Ay) - d(\delta_A)/d(Ay)$

Gillespie goes into more detail on applying the first and third measurement methods.

Results depend on the type of test, so just giving a deg/g value is not sufficient; it is also necessary to indicate the type of procedure used to measure the gradient.

Vehicles are inherently nonlinear systems, and it is normal for U to vary over the range of testing. It is possible for a vehicle to be understeer in some conditions and oversteer in others. Therefore, it is necessary to specify the speed and lateral acceleration whenever reporting understeer/oversteer characteristics.

Contributions to understeer

Many properties of the vehicle affect the understeer gradient, including tire cornering stiffness, camber thrust, lateral force compliance steer, aligning torque, lateral load transfer, and compliance in the steering system. These individual contributions can be identified analytically or by measurement in a Bundorf analysis.

Limit conditions

When an understeer vehicle is taken to frictional limits where it is no longer possible to increase lateral acceleration, the vehicle will follow a path with a radius larger than intended. Although the vehicle cannot increase lateral acceleration, it is dynamically stable.

When an oversteer vehicle is taken to frictional limits, it becomes dynamically unstable with a tendency to spin out. Although the vehicle is unstable in open-loop control, a skilled driver can maintain control a little past the point of instability with counter-steering. However, at some limit in lateral acceleration, it is not physically possible for even the most skilled driver to maintain a steady state and spinout will occur.

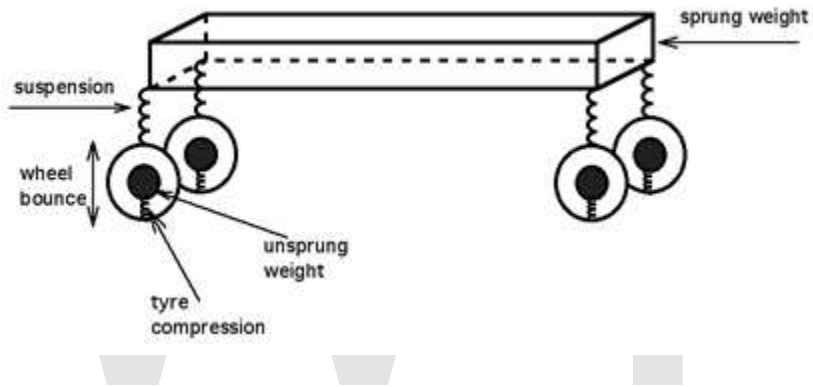
Related measures

Understeer gradient is one of the main measures for characterizing steady-state cornering behavior. It is involved in other properties such as characteristic speed (the speed for an understeer vehicle where the steer angle needed to negotiate a turn is twice the Ackerman angle), lateral acceleration gain (g's/deg), yaw velocity gain (1/s), and critical speed (the speed where an oversteer vehicle has infinite lateral acceleration gain).

Chapter- 7

Unsprung Mass and Vehicle Dynamics

Unsprung mass



In a ground vehicle with a suspension, the **unsprung weight** (or the **unsprung mass**) is the mass of the suspension, wheels or tracks (as applicable), and other components directly connected to them, rather than supported by the suspension. (The mass of the body and other components supported by the suspension is the sprung mass.) Unsprung weight includes the mass of components such as the wheel axles, wheel bearings, wheel hubs, tires, and a portion of the weight of driveshafts, springs, shock absorbers, and suspension links. Even if the vehicle's brakes are mounted outboard (i.e., within the wheel), their weight is still considered part of the unsprung weight.

Effects of Unsprung Weight

The unsprung weight of a wheel controls a trade-off between a wheel's bump-following ability and its vibration isolation. Bumps and surface imperfections in the road cause tire compression—which induces a force on the unsprung weight. The unsprung weight then responds to this force with movement of its own. The amount of movement, for short bumps, is inversely proportional to the weight - a lighter wheel which readily moves in response to road bumps will have more grip and more constant grip when tracking over an imperfect road. For this reason, lighter wheels are sought especially for high-

performance applications. In contrast, a heavier wheel which moves less will not absorb as much vibration; the irregularities of the road surface will transfer to the cabin through the geometry of the suspension and hence ride quality and road noise are deteriorated. For longer bumps that the wheels follow, greater unsprung mass causes more energy to be absorbed by the wheels and makes the ride worse.

Pneumatic or elastic tires help by providing some springing for most of the (otherwise) unsprung mass, but the damping that can be included in the tires is limited by considerations of fuel economy and overheating. The shock absorbers, if any, damp the spring motion also and must be less stiff than would optimally damp the wheel bounce. So the wheels execute some vibrations after each bump before coming to rest. On dirt roads and perhaps on some softly paved roads, these motions form small bumps, known as corrugations, washboarding or "corduroy" because they resemble smaller versions of the bumps in roads made of logs. These cause sustained wheel bounce in subsequent vehicles, enlarging the bumps.

High unsprung weight also exacerbates wheel control issues under hard acceleration or braking. If the vehicle does not have adequate wheel location in the vertical plane (such as a rear-wheel drive car with Hotchkiss drive, a live axle supported by simple leaf springs), vertical forces exerted by acceleration or hard braking combined with high unsprung mass can lead to severe wheel hop, compromising traction and steering control.

As mentioned above, there is a positive effect of unsprung mass. High frequency road irregularities, such as the gravel in an asphalt or concrete road surface, are isolated from the body more completely because the tires and springs act as separate filter stages, with the unsprung weight tending to uncouple them. Likewise, sound and vibration isolation is improved (at the expense of handling), in production automobiles, by the use of rubber bushings between the frame and suspension, by any flexibility in the frame or body work, and by the flexibility of the seats.

Unsprung Weight and Vehicle Design

Unsprung weight is largely a function of the design of a vehicle's suspension and the materials used in the construction of suspension components. Beam axle suspensions, in which wheels on opposite sides are connected as a rigid unit, generally have greater unsprung weight than independent suspension systems, in which the wheels are suspended and allowed to move separately. Heavy components such as the differential can be made part of the sprung weight by connecting them directly to the body (as in a de Dion tube rear suspension). Lightweight materials, such as aluminum, plastic, carbon fiber, and/or hollow components can provide further weight reductions at the expense of greater cost and/or fragility.

Inboard brakes can significantly reduce unsprung weight, but put more load on half axles and (constant velocity) universal joints, and require space that may not be easily accommodated. If located next to a differential or transaxle, waste heat from the brakes may overheat the differential or vice versa, particularly in hard use, such as motor racing.

They also make anti-dive suspension characteristics harder to achieve as the moment created by the act of braking is not reacted on the suspension arms.

Scooter-type motorcycles use an integrated engine-gearbox-final drive system that pivots as part of the rear suspension and hence is partly unsprung. This arrangement is linked to the use of quite small wheels, further impacting the reputation for road-holding.

Vehicle dynamics

Vehicle dynamics refers to the dynamics of vehicles, here assumed to be ground vehicles. Vehicle dynamics is a part of engineering primarily based on classical mechanics but it may also involve chemistry, solid state physics, electrical engineering, communications, psychology, control theory, etc.

Components

Components, attributes or aspects of vehicle dynamics include:

- Electronic Stability Control (ESC)
- Steering
- Suspension
- Traction control system (TCS)

Aerodynamic specific

Some attributes or aspects of vehicle dynamics are purely aerodynamic. These include:

- Automobile drag coefficient
- Automotive aerodynamics
- Center of pressure
- Downforce
- Ground effect in cars

Geometry specific

Some attributes or aspects of vehicle dynamics are purely geometric. These include:

- Ackermann steering geometry
- Axle track
- Camber angle
- Caster angle
- Ride height

- Roll center
- Toe
- Wheelbase

Mass specific

Some attributes or aspects of vehicle dynamics are purely due to mass and its distribution. These include:

- Center of mass
- Moment of inertia
- Sprung mass
- Unsprung mass
- Weight distribution

Motion specific

Some attributes or aspects of vehicle dynamics are purely dynamic. These include:

- Body flex
- Bump Steer
- Directional stability
- Critical speed
- Load transfer
- Noise, vibration, and harshness
- Oversteer
- Ride quality
- Speed wobble
- Understeer
- Weight transfer

Tire specific

Some attributes or aspects of vehicle dynamics can be attributed directly to the tires. These include:

- Camber thrust
- Circle of forces
- Contact patch
- Cornering force
- Ground pressure
- Pacejka's Magic Formula
- Pneumatic trail
- Relaxation length
- Rolling resistance
- Self aligning torque

- Slip angle
- Slip (vehicle dynamics)
- Steering ratio
- Tire load sensitivity

Driving techniques

Driving techniques which relate to, or improve the stability of vehicle dynamics include:

- Cadence braking
- Threshold braking
- Double declutching
- Drifting (motorsport)
- Handbrake turn
- Heel-and-Toe
- Left-foot braking
- Opposite lock
- Scandinavian flick

Analysis and simulation

The dynamic behavior of vehicles can be analysed in several different ways. This can be a straightforward as a simple spring mass system, through a 3 degree of freedom (DoF) bicycle model, to a large degree of complexity using a multibody system simulation package such as MSC ADAMS or Modelica. As computers have gotten faster, and software user interfaces have improved, commercial packages such as CarSim have become widely used in industry for rapidly evaluating hundreds of test conditions much faster than real time. Vehicle models are often simulated with advanced controller designs provided as software in the loop (SIL) with controller design software such as Simulink, or with physical hardware in the loop (HIL).

Vehicle motions are largely due to the shear forces generated between the tires and road, and therefore the tire model is an essential part of the math model. The tire model must produce realistic shear forces during braking, acceleration, cornering, and combinations, on a range of surface conditions. Many models are in use. Most are semi-empirical, such as the Pacejka Magic Formula model.

Racing car games or simulators are also a form of vehicle dynamics simulation. In early versions many simplifications were necessary in order to get real-time performance with reasonable graphics. However, improvements in computer speed have combined with interest in realistic physics, leading to driving simulators that are used for vehicle engineering using detailed models such as CarSim.

It is important that the models should agree with real world test results, hence many of the following tests are correlated against results from instrumented test vehicles.

Techniques include:

- Linear range constant radius understeer
- Fishhook
- Frequency response
- Lane change
- Moose test
- Sinusoidal steering
- Swept path analysis

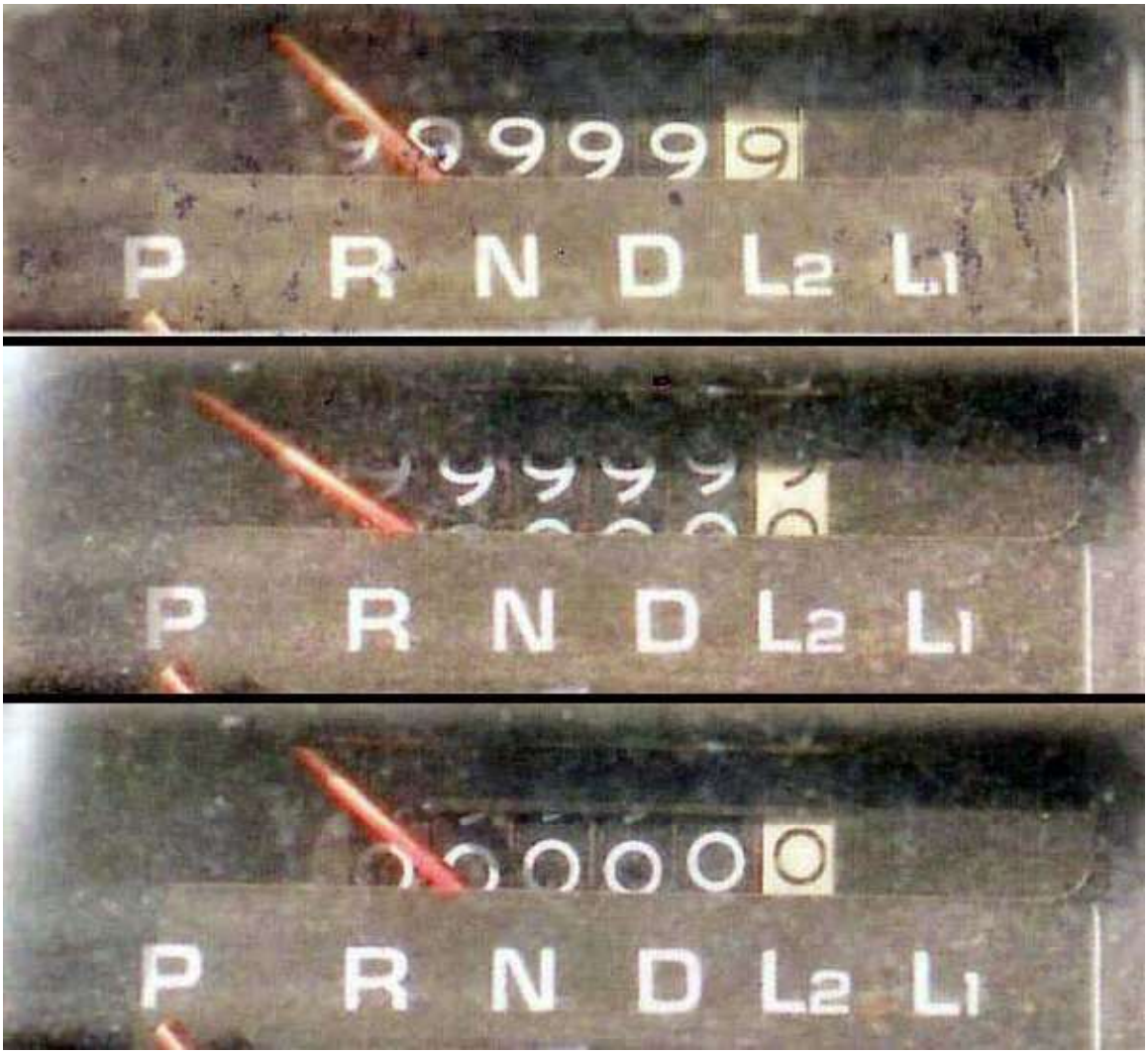
Chapter- 8

Odometer

A mechanical odometer with trip meter below.

An **odometer** (**mileometer**, **milometer**) indicates distance travelled by a bicycle, automobile, or other vehicle. The device may be electronic, mechanical, or a combination of the two. The word derives from the Greek words *hodós*, meaning "path" or gateway and "métron", "measure".

Description

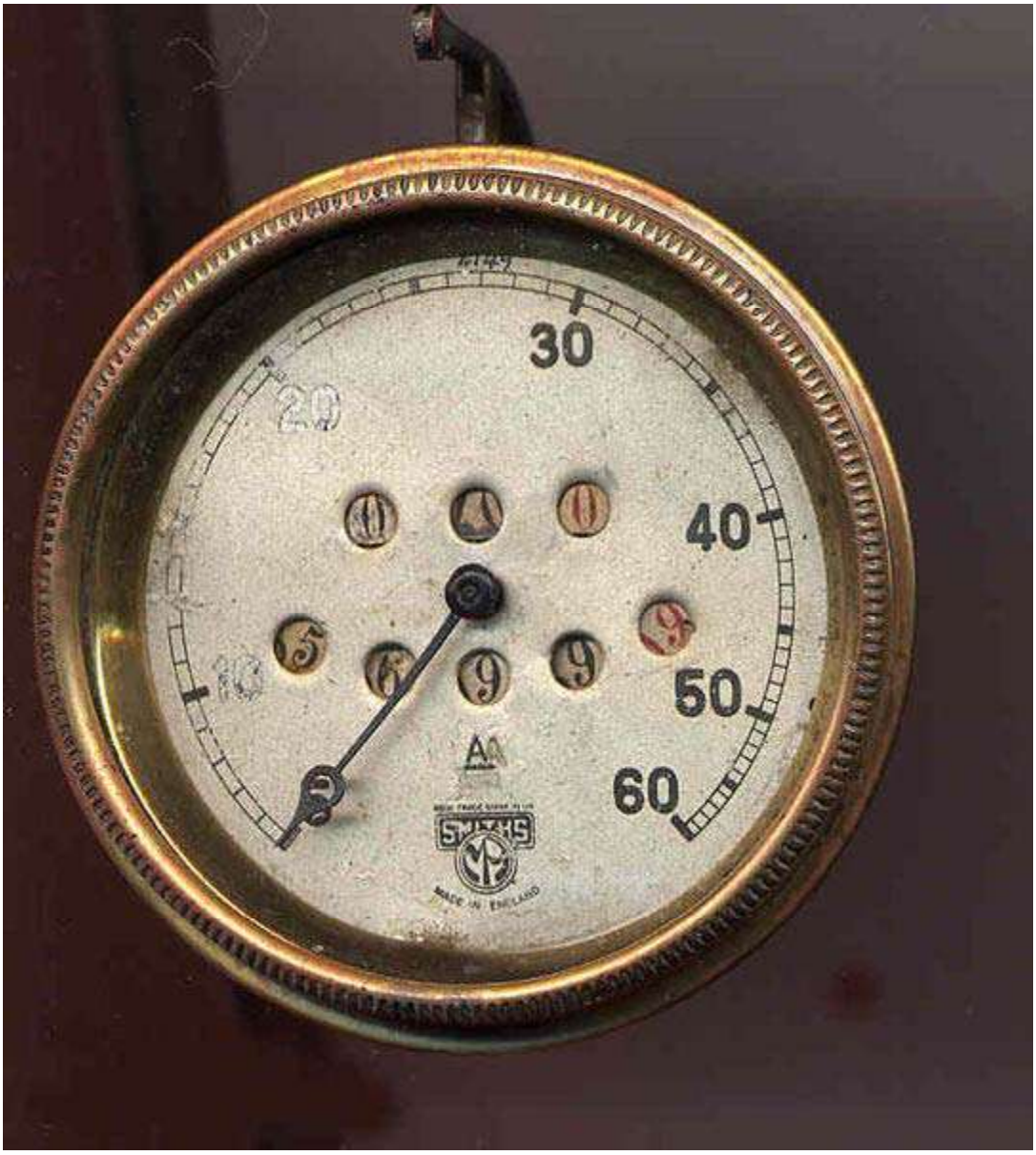


Odometer rollover.

In the early cars a top reading of 99,999 was enough. With improvements, modern vehicles need an extra digit. At the top reading, an odometer restarts from zero (odometer rollover).

Most modern cars include a **trip meter** (**trip odometer**). Unlike the odometer, a trip meter is reset at any point in a journey, making it possible to record the distance travelled in any particular journey or part of a journey. It was traditionally a purely mechanical device but, in most modern vehicles, it is now electronic. Luxury vehicles often have multiple trip meters. Most trip meters will show a maximum value of 999.9. The trip meter may be used to record the distance travelled on each tank of fuel, making it very easy to accurately track the energy efficiency of the vehicle; another common use is resetting it to zero at each instruction in a sequence of driving directions, to be sure when one has arrived at the next turn.

History



A Smiths speedometer from the 1920s showing odometer and trip meter

An electronic odometer with digital display

Classical Era

Possibly the first evidence for the use of an odometer can be found in the works of Pliny (NH 6. 61-62) and Strabo (11.8.9). Both authors list the distances of routes travelled by Alexander the Great (r. 336-323 BC) as measured by his bematists Diognetus and Baeton. However, the high precision of the bematists's measurements rather indicates the use of a mechanical device. For example, the section between the cities Hecatompylos and Alexandria Areion, which later became a part of the silk road, was given by Alexander's bematists as 529 English miles long, that is with a deviation of 0.4% from the actual distance (531 English miles). From the nine surviving bematists' measurements in Pliny's *Naturalis Historia* eight show a deviation of less than 5% from the actual distance, three of them being within 1%. Since these minor discrepancies can be adequately explained by slight changes in the tracks of roads during the last 2300 years, the overall accuracy of the measurements implies that the bematists already must have used a sophisticated device for measuring distances, although there is no direct mentioning of such a device.

An odometer for measuring distance was first described by Vitruvius around 27 and 23 BC, although the actual inventor may have been Archimedes of Syracuse during the First Punic War. Hero of Alexandria describes a similar device in chapter 34 of his *Dioptra*.

The machine was also used in the time of Roman Emperor Commodus (c. 192 AD), although after this point in time there seems to be a gap between its use in Roman times and that of the 15th century in Western Europe. Some researchers have speculated that the device might have included technology similar to that of the Greek Antikythera mechanism.

The odometer of Vitruvius was based on chariot wheels of 4 feet (1.2 m) diameter turning 400 times in one Roman mile (about 1400 m). For each revolution a pin on the axle engaged a 400 tooth cogwheel thus turning it one complete revolution per mile. This engaged another gear with holes along the circumference, where pebbles (*calculus*) were located, that were to drop one by one into a box. The distance travelled would thus be given simply by counting the number of pebbles. Whether this instrument was ever built at the time is disputed. Leonardo da Vinci later tried to build it himself according to the description, but failed. However, in 1981 engineer Andre Sleewyck built his own replica, replacing the square-toothed gear designs of da Vinci with the triangular, pointed teeth found in the *Antikythera mechanism*. With this modification, the Vitruvius odometer functioned perfectly.

Ancient China

The odometer was also later invented in ancient China, possibly by the profuse inventor and early scientist Zhang Heng (78–139 AD) of the Han Dynasty (202 BC–220 AD). Zhang Heng is often accredited with the invention of the first odometer device in China, an achievement alongside earlier contemporaries Archimedes and Heron of Alexandria from the Hellenized West. By the 3rd century (during the Three Kingdoms Period), the Chinese had termed the device as the 'jì lǐ gǔ chē' (記里鼓車) , or 'li-recording drum carriage' (Note: the modern measurement of li = 500 m/1640 ft). Chinese texts of the 3rd century tell of the mechanical carriage's functions, and as one li is traversed, a mechanical-driven wooden figure strikes a drum, and when ten li is traversed, another wooden figure would strike a gong or a bell with its mechanical-operated arm.

Despite its association with Zhang Heng or even the later Ma Jun (c. 200–265), there is evidence to suggest that the invention of the odometer was a gradual process in Han Dynasty China that centered around the *huang men* court people (i.e. eunuchs, palace officials, attendants and familiars, actors, acrobats, etc.) that would follow the musical procession of the royal 'drum-chariot'. The historian Joseph Needham asserts that it is no surprise this social group would have been responsible for such a device, since there is already other evidence of their craftsmenship with mechanical toys to delight the emperor and the court. There is speculation that some time in the 1st century BC (during the Western Han Dynasty), the beating of drums and gongs were mechanically-driven by working automatically off the rotation of the road-wheels. This might have actually been the design of one Loxia Hong (c. 110 BC), yet by 125 AD the mechanical odometer carriage in China was already known (depicted in a mural of the Xiao Tang Shan Tomb).

The odometer was used also in subsequent periods of Chinese history. In the historical text of the *Jin Shu* (635 AD), the oldest part of the compiled text, the book known as the

Cui Bao (c. 300 AD), recorded the use of the odometer, providing description (and interestingly enough attributing it to the Western Han era, from 202 BC–9 AD). The passage in the *Jin Shu* expanded upon this, explaining that it took a similar form to the mechanical device of the South Pointing Chariot invented by Ma Jun (200–265). As recorded in the *Song Shi* of the Song Dynasty (960-1279 AD), the odometer and South Pointing Chariot were combined into one wheeled device by engineers of the 9th century, 11th century, and 12th century (refer to South Pointing Chariot). The *Sun Tzu Suan Ching* (Master Sun's Mathematical Manual), dated from the 3rd century to 5th century, presented a mathematical problem for students involving the odometer. It involved a given distance between two cities, the small distance needed for one rotation of the carriage's wheel, and the posed question of how many rotations the wheels would have in all if the carriage was to travel between point A and B.

Song Dynasty odometer

The historical text of the *Song Shi* (1345 AD), recording the people and events of the Chinese Song Dynasty (960–1279), also mentioned the odometer used in that period. However, unlike written sources of earlier periods, it provided a much more thoroughly detailed description of the device that harkens back to its ancient form (Wade-Giles spelling):

The odometer. [The mile-measuring carriage] is painted red, with pictures of flowers and birds on the four sides, and constructed in two storeys, handsomely adorned with carvings. At the completion of every li, the wooden figure of a man in the lower storey strikes a drum; at the completion of every ten li, the wooden figure in the upper storey strikes a bell. The carriage-pole ends in a phoenix-head, and the carriage is drawn by four horses. The escort was formerly of 18 men, but in the 4th year of the Yung-Hsi reign-period (987 AD) the emperor Thai Tsung increased it to 30. In the 5th year of the Thien-Sheng reign-period (1027 AD) the Chief Chamberlain Lu Tao-lung presented specifications for the construction of odometers as follows:

What follows is a long dissertation made by the Chief Chamberlain Lu Daolong on the ranging measurements and sizes of wheels and gears, along with a concluding description at the end of how the device ultimately functions:

The vehicle should have a single pole and two wheels. On the body are two storeys, each containing a carved wooden figure holding a drumstick. The road-wheels are each 6 ft in diameter, and 18 ft in circumference, one evolution covering 3 paces. According to ancient standards the pace was equal to 6 ft and 300 paces to a li; but now the li is reckoned as 360 paces of 5 ft each.

The vehicle wheel (li lun) is attached to the left road-wheel; it has a diameter of 1.38 ft with a circumference of 4.14 ft, and has 18 cogs (chhih) 2.3 inches apart. There is also a lower horizontal wheel (hsia phing lun), of diameter 4.14 ft and circumference 12.42 ft, with 54 cogs, the same distance apart as those on the vertical wheel (2.3 inches). (This engages with the former.)

Upon a vertical shaft turning with this wheel, there is fixed a bronze "turning-like-the-wind wheel" (hsuan feng lun) which has (only) 3 cogs, the distance between these being 1.2 inches. (This turns the following one.) In the middle is a horizontal wheel, 4 ft in diameter, and 12 ft circumference, with 100 cogs, the distance between these cogs being the same as on the "turning-like-the-wind wheel" (1.2 inches).

Next, there is fixed (on the same shaft) a small horizontal wheel (hsiao phing lun) 3.3 inches in diameter and 1 ft in circumference, having 10 cogs 1.5 inches apart. (Engaging with this) there is an upper horizontal wheel (shang phing lun) having a diameter of 3.3 ft and a circumference of 10 ft, with 100 cogs, the same distance apart as those of the small horizontal wheel (1.5 inches).

When the middle horizontal wheel has made 1 revolution, the carriage will have gone 1 li and the wooden figure in the lower story will strike the drum. When the upper horizontal wheel has made 1 revolution, the carriage will have gone 10 li and the figure in the upper storey will strike the bell. The number of wheels used, great and small, is 8 inches in all, with a total of 285 teeth. Thus the motion is transmitted as if by the links of a chain, the "dog-teeth" mutually engaging with each other, so that by due revolution everything comes back to its original starting point (ti hsiang kou so, chhuan ya hsiang chih, chou erh fu shih).

Subsequent Developments

Odometers were first developed in the 1600s for wagons and other horse-drawn vehicles in order to measure distances traveled. In 1645 Blaise Pascal invented the *pascaline*. The *pascaline* utilized gears to compute measurements. Each gear contained 10 teeth. The first gear advanced the next gear one position when moved one complete revolution, the same principle employed on modern mechanical odometers.

Odometers were developed for ships in 1698 with the odometer invented by the Englishman Thomas Savery. Benjamin Franklin, U.S. statesman and the first Postmaster General, built a prototype odometer in 1775 that he attached to his carriage to help measure the mileage of postal routes. In 1847, William Clayton, a Mormon traveller, invented the *Roadometer*, which he attached to a wagon used by American settlers heading west. The *Roadometer* recorded the distance travelled each day by the wagon trains.

In 1895 Curtis Hussey Veeder invented the *Cyclometer*. The *Cyclometer* was a mechanical device that counted the number of rotations of a bicycle wheel. A flexible cable transmitted the number of rotations of the wheel to an analog odometer visible to the rider, which converted the wheel rotations into the number of miles traveled according to a predetermined formula.

In 1903 Arthur P. and Charles H. Warner, two brothers from Beloit, Wisconsin, introduced their patented *Auto-meter*. The *Auto-Meter* used a magnet attached to a rotating shaft to induce a magnetic pull upon a thin metal disk. Measuring this pull

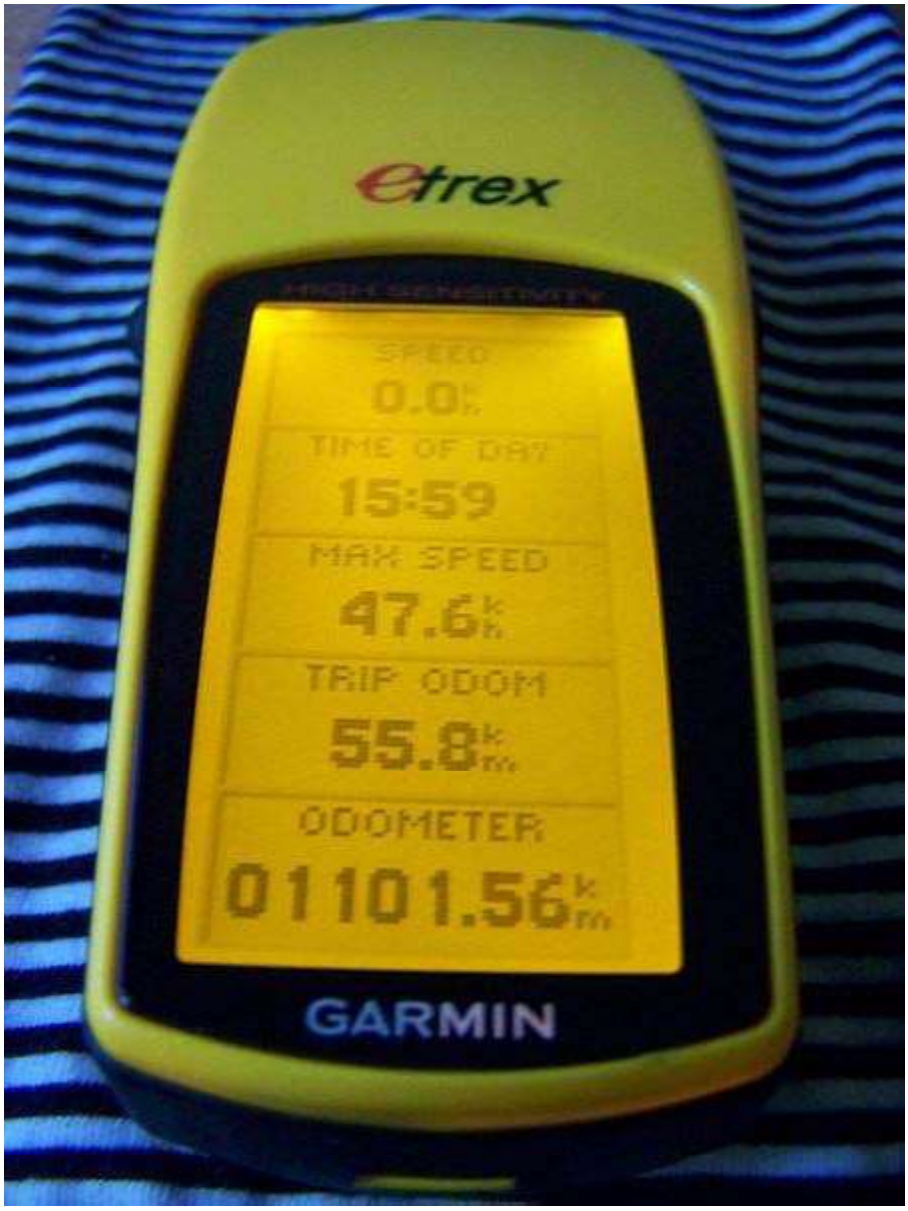
provided accurate measurements of both distance and speed information to automobile drivers in a single instrument. The Warners sold their company in 1912 to the Stewart & Clark Company of Chicago. The new firm was renamed the Stewart-Warner Corporation. By 1925, Stewart-Warner magnetic speedometers were standard equipment on the vast majority of automobiles and motorcycles manufactured in the United States.

Clocking and legality

A common form of fraud is to tamper with the reading on an odometer; this is often referred to as clocking. This is done to make a car appear to have been driven less than it really has been, and thus increase its apparent market value. Many new cars sold today use digital odometers that store the mileage in the vehicle's engine control module making it difficult (but not impossible) to manipulate the mileage electronically. With mechanical odometers, the speedometer can be removed from the car dash board and the digits wound back, or the drive cable can be disconnected and connected to another odometer/speedometer pair while on the road. Modern odometers now add mileage driven in reverse to the total as if driven forward, to accurately reflect the true total wear and tear on the vehicle (older vehicles could be driven in reverse to subtract mileage).

The resale value of a vehicle is often strongly influenced by the number of miles or kilometres a passenger vehicle has on the odometer, yet odometers are inherently insecure because they are under the control of their owners. Many jurisdictions have chosen to enact laws which penalize people who are found to commit odometer fraud. In the US (and many other countries), vehicle mechanics are also required to keep records of the odometer any time a vehicle is serviced. Companies such as Carfax then use this data to help potential car buyers detect whether odometer rollback has occurred.

GPS used as odometer



A Garmin etrex H GPS receiver, showing an odometer, trip odometer as well as speed related information

Recently, exercise enthusiasts have observed that an advanced Global Positioning System receiver (GPSr) with an odometer mode serves as a very accurate pedometer for outdoor activities. While not truly counting steps (no pendulum is involved) an advanced GPS odometer can accurately reveal the distance traveled to within 1/100 of a mile (depending on the model, perhaps 1/1000 of a mile). 1/1000 of a mile is approximately the distance of a single pace or 2 steps (1.609 m). Precise metric odometers have a precision of 1/100 or 1/1000 km, 10 or 1 metre(s) respectively.

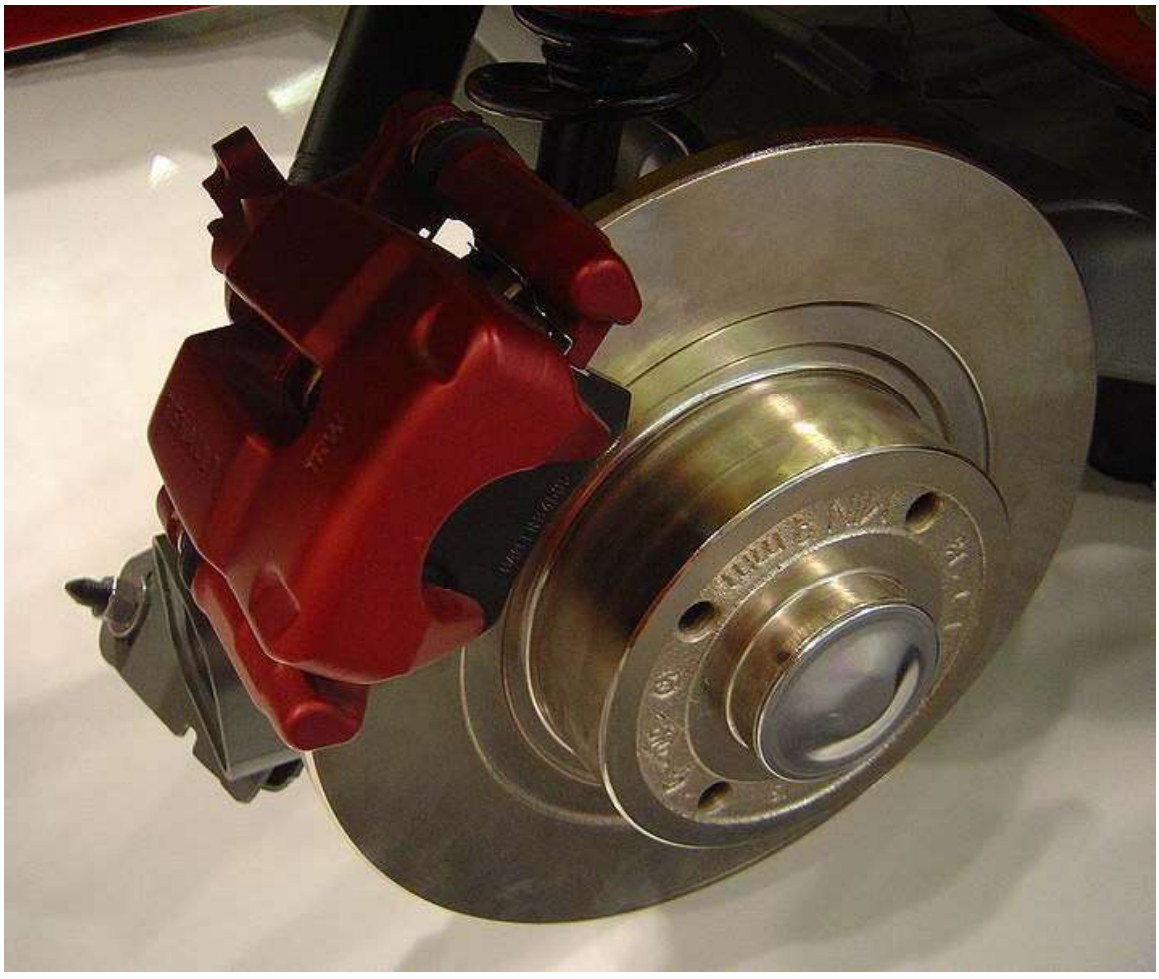
A GPS with odometer mode is also an excellent and inexpensive means to verify proper operation of both the speedometer and odometer mounted in a vehicle.

Odometer tax

This would be a road vehicle tax collected per distance unit of travel. This may become increasingly important with the increasing presence of gas burning vehicles & very fuel efficient electric models.

Chapter- 9

Disc Brake



Close-up of a disc brake on a car

On automobiles, disc brakes are often located within the wheel

The **disc brake** or **disk brake** is a device for slowing or stopping the rotation of a wheel while it is in motion. A brake disc (or *rotor* in U.S. English) is usually made of cast iron, but may in some cases be made of composites such as reinforced carbon-carbon or ceramic-matrix composites. This is connected to the wheel and/or the axle. To stop the wheel, friction material in the form of brake pads (mounted on a device called a **brake caliper**) is forced mechanically, hydraulically, pneumatically or electromagnetically against both sides of the disc. Friction causes the disc and attached wheel to slow or stop. Brakes convert motion to heat, and if the brakes get too hot, they become less effective, a phenomenon known as brake fade.

History

Disc-style brakes development and use began in England in the 1890s. The first caliper-type automobile disc brake was patented by Frederick William Lanchester in his Birmingham, UK factory in 1902 and used successfully on Lanchester cars. However, the limited choice of metals in this period, meant that he had to use copper as the braking medium acting on the disc. The poor state of the roads at this time, no more than dusty, rough tracks, meant that the copper wore quickly making the disc brake system non-viable (as recorded in The Lanchester Legacy). It took another half century for his innovation to be widely adopted.

Modern-style disc brakes first appeared on the low-volume Crosley Hotshot in 1949, although they had to be discontinued in 1950 due to design problems. Chrysler's Imperial also offered a type of disc brake from 1949 through 1953, though in this instance they were enclosed with dual internal-expanding, full-circle pressure plates. Reliable modern disc brakes were developed in the UK by Dunlop and first appeared in 1953 on the Jaguar C-Type racing car. The Citroën DS of 1955, with powered inboard front disc brakes was the first foreign, whilst the 1956 Triumph TR3 was the first English production car to feature modern disc brakes. The first production car to feature disc brakes at all 4 wheels was the Austin-Healey 100S in 1954. The first British company to market a production saloon fitted with disc brakes to all four wheels was Jensen Motors with the introduction of a Deluxe version of the Jensen 541 with Dunlop disc brakes. The first German production car with disc brakes was the 1961 Mercedes-Benz 220SE coupe featuring British-built Girling units on the front. The next American production automobile equipped with caliper-type disc brakes was the 1963 model year Studebaker Avanti (the Bendix system optional on some of the other Studebaker models). Front disk brakes became standard equipment in 1965 on the Rambler Marlin (the Bendix units were optional on all American Motors "senior" platform models), the Ford Thunderbird, and the Lincoln Continental. A four-wheel disc brake system was also introduced in 1965 on the Chevrolet Corvette Stingray.

Compared to drum brakes, disc brakes offer better stopping performance, because the disc is more readily cooled. As a consequence discs are less prone to the "brake fade" caused when brake components overheat; and disc brakes recover more quickly from immersion (wet brakes are less effective). A drum brake will have at least one leading shoe, which gives a servo-effect. By contrast, a disc brake has no self-servo effect and its braking force is always proportional to the pressure placed on the brake pad by the braking system via any brake servo, braking pedal or lever.

Many early implementations for automobiles located the brakes on the inboard side of the driveshaft, near the differential, but most brakes today are located inside the road wheels. (An inboard location reduces the unsprung weight and eliminates a source of heat transfer to the tires.)

Disc brakes were most popular on sports cars when they were first introduced, since these vehicles are more demanding about brake performance. Discs have now become the more

common form in most passenger vehicles, although many (particularly light weight vehicles) use drum brakes on the rear wheels to keep costs and weight down as well as to simplify the provisions for a parking brake. As the front brakes perform most of the braking effort, this can be a reasonable compromise.

Discs

A cross-drilled disc on a modern motorcycle

The design of the disc varies somewhat. Some are simply solid cast iron, but others are hollowed out with fins or vanes joining together the disc's two contact surfaces (usually included as part of a casting process). The weight and power of the vehicle will determine the need for ventilated discs. The "ventilated" disc design helps to dissipate the generated heat and is commonly used on the more-heavily-loaded front discs. The front brakes provide most of the stopping power.

Many higher performance brakes have holes drilled through them. This is known as cross-drilling and was originally done in the 1960s on racing cars. For heat dissipation purposes, cross drilling is still used on some braking components, but is not favored for racing or other hard use as the holes are a source of stress cracks under severe conditions.

Discs may also be slotted, where shallow channels are machined into the disc to aid in removing dust and gas. Slotting is the preferred method in most racing environments to remove gas, water, and de-glaze brake pads. Some discs are both drilled and slotted. Slotted discs are generally not used on standard vehicles because they quickly wear down

brake pads; however, this removal of material is beneficial to race vehicles since it keeps the pads soft and avoids vitrification of their surfaces. As a way of avoiding thermal stress, cracking and warping of the disc these are sometimes mounted in a half loose way to the hub with coarse splines. This allows the disc to expand in a controlled symmetrical way and with less unwanted heat transfer to the hub.

A mountain bike disc brake

On the road, drilled or slotted discs still have a positive effect in wet conditions because the holes or slots prevent a film of water building up between the disc and the pads. Crossdrilled discs may eventually crack at the holes due to metal fatigue. Cross-drilled brakes that are manufactured poorly or subjected to high stresses will crack much sooner and more severely.

The first motorcycles to use disc brakes were racing machines. The first mass-produced road-going motorcycle to sport a disc-brake was the 1969 Honda CB750. Disc brakes are now common on motorcycles, mopeds and even bicycles. Motorcycle disc brakes have become increasingly sophisticated, partly through marketing. Their discs are usually drilled and occasionally slotted. Calipers have evolved from simple "single-pot" units to 2-, 4- and even 6-pot items. It is debatable whether the modern fashions of "radially-mounted calipers" and "wavy discs" significantly improve braking. Since (compared to

cars) motorcycles have a higher centre of gravity:wheelbase ratio, they experience more weight transference when braking. A modern sports bike will typically have twin front discs of large diameter, but only a single rear disc that is very much smaller (or even a small rear drum brake). The front brake(s) provide most of the required deceleration; the rear brake serves mainly as to "balance" the motorcycle during braking. If too much braking force is applied to the rear brake, the rear wheel is liable to lock up; so motorcycles should not have oversize rear brakes.

Mountain bike disc brakes range from simple, mechanical (cable) systems, to expensive and powerful, 6-pot (piston) hydraulic disc systems, commonly used on downhill racing bikes. Improved technology has seen the creation of the first vented discs for use on mountain bikes, similar to those on cars, introduced to help avoid heat fade on fast alpine descents. Although less common, discs are also used on road bicycles for all-weather cycling with predictable braking, although drums are sometimes preferred as harder to damage in crowded parking, where discs are sometimes bent. Most bicycle brake discs are made of stainless steel, although some lightweight discs are made of titanium or aluminium. Discs are thin, often about 2 mm. Some use a two-piece floating disc style, others use a floating caliper, others use pads that float in the caliper, and some use one moving pad that makes the caliper slide on its mounts, pulling the other pad into contact with the disc. Because the "motor" is small, an uncommon feature of bicycle brakes is pads that retract to eliminate residual drag when the brake is released. In contrast, most other brakes drag the pads lightly when released.

Disc brakes are increasingly used on very large and heavy road vehicles, where previously large drum brakes were nearly universal. One reason is the disc's lack of self-assist makes brake force much more predictable, so peak brake force can be raised without more risk of braking-induced steering or jackknife on articulated vehicles. Another is disk brakes fade less when hot, and in a heavy vehicle air and rolling drag and engine braking are small parts of total braking force, so brakes are used harder than on lighter vehicles, and drum brake fade can occur in a single stop. For these reasons, a heavy truck with disc brakes can stop in about 120% the distance of a passenger car, but with drums stopping takes about 150% the distance. In Europe, stopping distance regulations essentially require disc brakes for heavy vehicles. In the U.S., drums are allowed and are typically preferred for their lower purchase price, despite higher total lifetime cost and more frequent service intervals.

A railroad bogie and disc brakes

Yet larger discs are used for railroads and some airplanes. Passenger rail cars and light rail often use disc brakes outboard of the wheels, which helps ensure a free flow of cooling air. In contrast, some airplanes have the brake mounted with very little cooling and the brake gets quite hot in a stop, but this is acceptable as the maximum braking energy is very predictable.

For auto use, disc brake discs are commonly manufactured out of a material called grey iron. The SAE maintains a specification for the manufacture of grey iron for various applications. For normal car and light truck applications, the SAE specification is J431 G3000 (superseded to G10). This specification dictates the correct range of hardness, chemical composition, tensile strength, and other properties necessary for the intended use. Some racing cars and airplanes use brakes with carbon fiber discs and carbon fiber pads to reduce weight. Wear rates tend to be high, and braking may be poor or grabby until the brake is hot.

Historically, brake discs were manufactured throughout the world with a strong concentration in Europe, and America. Between 1989 and 2005, manufacturing of brake discs is migrating predominantly to China.

Racing

A reinforced carbon brake disc installed on a Ferrari F430 Challenge race car

In racing and very high performance road cars, other disc materials have been employed. Reinforced carbon discs and pads inspired by aircraft braking systems were introduced in Formula One by Brabham in conjunction with Dunlop in 1976. Carbon-Carbon braking is now used in most top-level motorsport worldwide, reducing unsprung weight, giving better frictional performance and improved structural properties at high temperatures, compared to cast iron. Carbon brakes have occasionally been applied to road cars, by the French Venturi sports car manufacturer in the mid 1990s for example, but need to reach a very high operating temperature before becoming truly effective and so are not well suited to road use. The extreme heat generated in these systems is easily visible during night racing, especially at shorter tracks. It is not uncommon to be able to look at the cars, either live in person or on television and see the brake discs glowing red during application.

Ceramic composites

Ceramic discs are used in some high-performance cars and heavy vehicles.

The first development of the modern ceramic brake was made by British Engineers working in the railway industry for TGV applications in 1988. The objective was to reduce weight, the number of brakes per axle, as well as provide stable friction from very high speeds and all temperatures. The result was a carbon fibre reinforced ceramic process which is now used in various forms for automotive, railway, and aircraft brake applications.

The requirement for a large section of ceramic composite material having very high heat tolerance and mechanical strength often relegates ceramic discs to exotic vehicles where the cost is not prohibitive to the application, and industrial use where the ceramic disc's light weight and low maintenance properties justify the cost relative to alternatives. Composite brakes can withstand temperatures that would make steel discs bendable.

Mercedes Benz AMG Carbon Ceramic brake

Porsche's ceramic composite brakes, known as PCCB (Porsche Composite Ceramic Brakes), are siliconized carbon fiber, with very high temperature capability, a 50% weight reduction over iron rotors (therefore reducing the unsprung weight of the vehicle), a significant reduction in dust generation, substantially increased maintenance intervals, and enhanced durability in corrosive environments over conventional iron rotors. Found on some of their more expensive models, e.g., the Carrera GT, 911 GT2, etc. it is also an

optional brake for all street Porsches at added expense. It is generally recognized by the bright yellow paintwork on the aluminum 6-piston calipers that are matched with the rotors. The rotors are internally vented much like cast iron rotors, and also cross-drilled.

Disc damage modes

Discs are usually damaged in one of four ways: scarring, cracking, warping or excessive rusting. Service shops will sometimes respond to any disc problem by changing out the discs entirely, This is done mainly where the cost of a new disc may actually be lower than the cost of labour to resurface the original disc. Mechanically this is unnecessary unless the discs have reached manufacturer's minimum recommended thickness, which would make it unsafe to use them, or vane rusting is severe (ventilated discs only). Most leading vehicle manufacturers recommend brake disc skimming (US: rotor turning) as a solution for lateral run-out, vibration issues and brake noises. The machining process is performed in a brake lathe, which removes a very thin layer off the disc surface to clean off minor damage and restore uniform thickness. Machining the disc as necessary will maximise the mileage out of the current discs on the vehicle.

Excessive lateral run-out (warping)

Measuring this is accomplished using a dial indicator on a fixed rigid base, with the tip perpendicular to the brake disc's face. It is typically measured about 1/2" (12 mm) from the outside diameter of the disc. The disc is spun. The difference between minimum and maximum value on the dial is called lateral runout. Typical hub/disc assembly runout specifications for passenger vehicles are around 0.0020" or 50 micrometers. Runout can be caused either by deformation of the disc itself or by runout in the underlying wheel hub face or by contamination between the disc surface and the underlying hub mounting surface. Determining the root cause of the indicator displacement (lateral runout) requires disassembly of the disc from the hub. Disc face runout due to hub face runout or contamination will typically have a period of 1 minimum and 1 maximum per revolution of the brake disc.

Discs can be machined to eliminate thickness variation and lateral runout. Machining can be done in-situ (on-car) or off-car (bench lathe). Both methods will eliminate thickness variation. Machining on-car with proper equipment can also eliminate lateral runout due to hub-face non-perpendicularity.

Incorrect fitting can distort (warp) discs; the disc's retaining bolts (or the wheel/lug nuts, if the disc is simply sandwiched in place by the wheel, as on many cars) must be tightened progressively and evenly. The use of air tools to fasten lug nuts is extremely bad practice, unless a torque tube is also used. The vehicle manual will indicate the proper pattern for tightening as well as a torque rating for the bolts. Lug nuts should never be tightened in a circle. Some vehicles are sensitive to the force the bolts apply and tightening should be done with a torque wrench.

Often uneven pad transfer is confused for disc warping. In reality, the majority of brake discs which are diagnosed as "warped" are actually simply the product of uneven transfer of pad material.

Uneven pad transfer will often lead to a thickness variation of the disc. When the thicker section of the disc passes between the pads, the pads will move apart and the brake pedal will raise slightly; this is pedal pulsation. The thickness variation can be felt by the driver when it is approximately 0.17 mm or greater (on automobile discs).

This type of thickness variation has many causes, but there are three primary mechanisms which contribute the most to the propagation of disc thickness variations connected to uneven pad transfer. The first is improper selection of brake pads for a given application. Pads which are effective at low temperatures, such as when braking for the first time in cold weather, often are made of materials which decompose unevenly at higher temperatures. This uneven decomposition results in uneven deposition of material onto the brake disc. Another cause of uneven material transfer is improper break in of a pad/disc combination. For proper break in, the disc surface should be refreshed (either by machining the contact surface or by replacing the disc as a whole) every time the pads are changed on a vehicle. Once this is done, the brakes are heavily applied multiple times in succession. This creates a smooth, even interface between the pad and the disc. When this is not done properly the brake pads will see an uneven distribution of stress and heat, resulting in an uneven, seemingly random, deposition of pad material. The third primary mechanism of uneven pad material transfer is known as "pad imprinting." This occurs when the brake pads are heated to the point that the material begins to break-down and transfer to the disc. In a properly broken in brake system (with properly selected pads), this transfer is natural and actually is a major contributor to the braking force generated by the brake pads. However, if the vehicle comes to a stop and the driver continues to apply the brakes, the pads will deposit a layer of material in the shape of the brake pad. This small thickness variation can begin the cycle of uneven pad transfer.

Once the disc has some level of variation in thickness, uneven pad deposition can accelerate, sometimes resulting in changes to the crystal structure of the metal that composes the disc in extreme situations. As the brakes are applied, the pads slide over the varying disc surface. As the pads pass by the thicker section of the disc, they are forced outwards. The foot of the driver applied to the brake pedal naturally resists this change, and thus more force is applied to the pads. The result is that the thicker sections see higher levels of stress. This causes an uneven heating of the surface of the disc, which causes two major issues. As the brake disc heats unevenly it also expands unevenly. The thicker sections of the disc expand more than the thinner sections due to seeing more heat, and thus the difference in thickness is magnified. Also, the uneven distribution of heat results in further uneven transfer of pad material. The result is that the thicker-hotter sections receive even more pad material than the thinner-cooler sections, contributing to a further increase in the variation in the disk's thickness. In extreme situations, this uneven heating can actually cause the crystal structure of the disc material to change. When the hotter sections of the discs reach extremely high temperatures(1200-1300 degrees Fahrenheit), the carbon within the cast iron of the disc will react with the iron molecules

to form a carbide known as cementite. This iron carbide is very different from the cast iron the rest of the disc is composed of. It is extremely hard, very brittle, and does not absorb heat well. After cementite is formed, the integrity of the disc is compromised. Even if the disc surface is machined, the cementite within the disc will not wear or absorb heat at the same rate as the cast iron surrounding it, causing the uneven thickness and uneven heating characteristics of the disc to return.

Scarring

Brake discs being polished after scarring occurred

Scarring (US: Scoring) can occur if brake pads are not changed promptly when they reach the end of their service life and are considered worn out. Once enough of the friction material has worn away, the pad's steel backing plate (for glued pads) or the pad

retainer rivets (for riveted pads) will bear directly upon the disc's wear surface, reducing braking power and making scratches on the disc. Generally a moderately scarred / scored disc, which operated satisfactorily with existing brake pads, will be equally usable with new pads. If the scarring is deeper but not excessive, it can be repaired by machining off a layer of the disc's surface. This can only be done a limited number of times as the disc has a minimum rated safe thickness. The minimum thickness value is typically cast into the disc during manufacturing on the hub or the edge of the disc. In Pennsylvania, which has one of the most rigorous auto safety inspection programs in North America, an automotive disc cannot pass safety inspection if any scoring is deeper than .015 inches (0.38 mm), and must be replaced if machining will reduce the disc below its minimum safe thickness.

To prevent scarring, it is prudent to periodically inspect the brake pads for wear. A tire rotation is a logical time for inspection, since rotation must be performed regularly based on vehicle operation time and all wheels must be removed, allowing ready visual access to the brake pads. Some types of alloy wheels and brake arrangements will provide enough open space to view the pads without removing the wheel. When practical, pads that are near the wear-out point should be replaced immediately, as complete wear out leads to scarring damage and unsafe braking. Many disc brake pads will include some sort of soft steel spring or drag tab as part of the pad assembly, which is designed to start dragging on the disc when the pad is nearly worn out. The result is a moderately loud metallic squealing noise, alerting the vehicle user that service is required, and this will not normally scar the disc if the brakes are serviced promptly. A set of pads can be considered for replacement if the thickness of the pad material is the same or less than the thickness of the backing steel. In Pennsylvania, the standard is 1/32".

Cracking

Cracking is limited mostly to drilled discs, which may develop small cracks around edges of holes drilled near the edge of the disc due to the disc's uneven rate of expansion in severe duty environments. Manufacturers that use drilled discs as OEM typically do so for two reasons: appearance, if they determine that the average owner of the vehicle model will prefer the look while not overly stressing the hardware; or as a function of reducing the unsprung weight of the brake assembly, with the engineering assumption that enough brake disc mass remains to absorb racing temperatures and stresses. A brake disc is a heat sink, but the loss of heat sink mass may be balanced by increased surface area to radiate away heat. Small hairline cracks may appear in any cross drilled metal disc as a normal wear mechanism, but in the severe case the disc will fail catastrophically. No repair is possible for the cracks, and if cracking becomes severe, the disc must be replaced.

Rusting

The discs are commonly made from cast iron and a certain amount of what is known as "surface rust" is normal. The disc contact area for the brake pads will be kept clean by regular use, but a vehicle that is stored for an extended period can develop significant rust

in the contact area that may reduce braking power for a time until the rusted layer is worn off again. Over time, vented brake discs may develop severe rust corrosion inside the ventilation slots, compromising the strength of the structure and needing replacement.

Calipers

Disc brake caliper (twin-pot, floating) removed from brake pad for changing pads

The **brake caliper** is the assembly which houses the brake pads and pistons. The pistons are usually made of aluminium or chrome-plated steel. There are two types of calipers: floating or fixed. A fixed caliper does not move relative to the disc and is, thus, less tolerant of disc imperfections. It uses one or more single or pairs of opposing pistons to clamp from each side of the disc, and is more complex and expensive than a floating caliper. A floating caliper (also called a "sliding caliper") moves with respect to the disc, along a line parallel to the axis of rotation of the disc; a piston on one side of the disc pushes the inner brake pad until it makes contact with the braking surface, then pulls the caliper body with the outer brake pad so pressure is applied to both sides of the disc.

Floating caliper (single piston) designs are subject to sticking failure, which can occur due to dirt or corrosion entering at least one mounting mechanism and stopping its normal movement. This can cause the pad attached to the caliper to rub on the disc when the brake is not engaged, or cause it to engage at an angle. Sticking can occur due to infrequent vehicle use, failure of a seal or rubber protection boot allowing debris entry, dry-out of the grease in the mounting mechanism and subsequent moisture incursion leading to corrosion, or some combination of these factors. Consequences may include reduced fuel efficiency and excessive wear on the affected pad.

Various types of brake calipers are also used on bicycle rim brakes.

Pistons and cylinders

The most common caliper design uses a single hydraulically actuated piston within a cylinder, although high performance brakes use as many as twelve. Modern cars use different hydraulic circuits to actuate the brakes on each set of wheels as a safety measure. The hydraulic design also helps multiply braking force. The number of pistons in a caliper is often referred to as the number of 'pots', so if a vehicle has 'six pot' calipers it means that each caliper houses six pistons.

Brake failure can occur due to failure of the piston to retract - this is usually a consequence of not operating the vehicle during a time that it is stored outdoors in adverse conditions. On high mileage vehicles the piston seals may leak, which must be promptly corrected. The brake disc must have enough surface to perform well and the **coefficient of friction** is the most important factor to be considered when designing a brake system.

Brake pads

The brake pads are designed for high friction with brake pad material embedded in the disc in the process of bedding while wearing evenly. Although it is commonly thought that the pad material contacts the metal of the disc to stop the car, the pads work with a very thin layer of their own material and generate a semi-liquid friction boundary that creates the actual braking force. Friction can be divided into two parts: Adhesive and abrasive. Of course, depending on the properties of the material of both the pad and the disc and the configuration and the usage, pad and disc wear rates will vary considerably. The properties that determine material wear involve trade-offs between performance and longevity. The friction coefficient for most standard pads will be in the region of .40 when used with cast iron discs. Racing pads with high iron content designed for use with cast iron brake discs reach .55 to .60 which gives a very significant increase in braking power and high temperature performance. High iron content racing pads wear down discs very quickly and usually when the pads are worn out so are the discs.

The brake pads must usually be replaced regularly (depending on pad material), and some are equipped with a mechanism that alerts drivers that replacement is needed. Some have a thin piece of soft metal that rubs against the disc when the pads are too thin, causing the brakes to squeal, while others have a soft metal tab embedded in the pad material that closes an electric circuit and lights a warning light when the brake pad gets thin. More expensive cars may use an electronic sensor.

Generally road-going vehicles have two brake pads per caliper, while up to six are installed on each racing caliper, with varying frictional properties in a staggered pattern for optimum performance.

Early brake pads (and linings) contained asbestos. When working on an older car's brakes, care must be taken not to inhale any dust present on the caliper (or drum).

Although newer pads can be made of exotic materials like ceramics, kevlar and other plastics, inhalation of brake dust should still be avoided regardless of material.

Brake squeal

Sometimes a loud noise or high pitch squeal occurs when the brakes are applied. Most brake squeal is produced by vibration (resonance instability) of the brake components, especially the pads and discs (known as *force-coupled excitation*). This type of squeal should not negatively affect brake stopping performance. Simple techniques like adding chamfers to linings, greasing or gluing the contact between caliper and the pads (finger to backplate, piston to backplate), bonding insulators (damping material) to pad backplate, inclusion of a brake shim between the brake pad and back plate, etc. may help to reduce squeal. Cold weather combined with high early morning humidity (dew) often makes brake-squeal worse, although the squeal stops when the lining reaches regular operating temperatures. Dust on the brakes may also cause squeal; there are many commercial brake cleaning products that can be used to remove dust and contaminants. Finally, some lining wear indicators, located either as a semi-metallic layer within the brake pad material or with an external squealer "sensor", are also designed to squeal when the lining is due for replacement. The typical external sensor is fundamentally different because it occurs when the brakes are off, and goes away when the brakes are on.

Overall brake squeal can be annoying to the vehicle passengers, passers-by, pedestrians, etc. especially as vehicle designs become quieter. Noise, vibration, and harshness (NVH) are among the most important priorities for today's vehicle manufacturers.

Apart from noise generated from squeal, brakes may also develop a phenomenon called *brake judder* or *shudder*.

Brake judder

Brake judder is usually perceived by the driver as minor to severe vibrations transferred through the chassis during braking.

The judder phenomenon can be classified into two distinct subgroups: *hot* (or *thermal*), or *cold* judder.

Hot judder is usually produced as a result of longer, more moderate braking from high speed where the vehicle does not come to a complete stop. It commonly occurs when a motorist decelerates from speeds of around 120 km/h (74.6 MPH) to about 60 km/h (37.3 MPH), which results in severe vibrations being transmitted to the driver. These vibrations are the result of uneven thermal distributions, or *hot spots*. Hot spots are classified as concentrated thermal regions that alternate between both sides of a disc that distort it in such a way that produces a sinusoidal waviness around its edges. Once the brake pads (friction material/brake lining) comes in contact with the sinusoidal surface during braking, severe vibrations are induced, and can produce hazardous conditions for the person driving the vehicle.

Cold judder, on the other hand, is the result of uneven disc wear patterns or disc thickness variation (DTV). These variations in the disc surface are usually the result of extensive vehicle road usage. DTV is usually attributed to the following causes: waviness of disc surface, misalignment of axis (runout), elastic deflection, wear and friction material transfers.

Brake dust

When braking force is applied, small amounts of material are gradually ground off the brake pads. This material is known as "brake dust" and a fair amount of it usually deposits itself on the braking system and the surrounding wheel. Brake dust can badly damage the finish of most wheels if not washed off. Airborne brake dust is known to be a health hazard, so most repair manuals recommend the use of a chemical 'brake cleaner' instead of compressed air to remove the dust. Different brake pad formulations create different amounts of dust, and some formulations, particularly metallic brake pads, are much more damaging than others. Ceramic brake pads contain significantly fewer metal particles, and therefore produce less corrosion of surrounding metal parts.

Chapter- 10

Windscreen Wiper

A common windscreen wiper arm and blade

Windscreen wiper on a parked car. In this common design, the force from the arm is distributed evenly with a series of linkages known as a whippletree.

A common design for a "wiper" (also called a clear view screen) on a ship. A round portion of the windscreen has two layers, the outer one of which is spun at high speed.

A **windscreen wiper** or **windshield wiper** is a device used to remove rain and debris from a windscreen or windshield. Almost all motor vehicles, including trains, aircraft and watercraft, are equipped with such wipers, which are usually a legal requirement.

A wiper generally consists of an arm, pivoting at one end and with a long rubber blade attached to the other. The blade is swung back and forth over the glass, pushing water from its surface. The speed is normally adjustable, with several continuous speeds and often one or more "intermittent" settings. Most automobiles use two synchronized radial type arms, while many commercial vehicles use one or more pantograph arms.

History

In 1903, Mary Anderson is credited with inventing the first operational windshield wiper. In Anderson's patent, she called her invention a window cleaning device for electric cars and other vehicles. Operated via a lever from inside a vehicle, her version of windshield wipers closely resembles the windshield wiper found on many early car models. Anderson had a model of her design manufactured. She then filed a patent (U.S. patent number 743,801) on June 18, 1903 that was issued to her by the U.S. Patent Office on November 10, 1903.

Irish born Inventor James Henry Apjohn (1845–1914) devised a method of moving two brushes up and down on a vertical plate glass windscreen in 1903. This was patented in the UK.

In April 1911, a patent for windscreen wipers was registered by Sloan & Lloyd Barnes, patent agents of Liverpool, England, for Gladstone Adams of Whitley Bay. The first designs for the windscreen wiper are also credited to concert pianist Józef Hofmann, and Mills Munitions, Birmingham who also claimed to have been the first to patent windscreen wipers in England.

The company Oishei formed, the Tri-Continental Corporation, introduced the first windshield wiper, Rain Rubber, for the slotted, two-piece windshields found on many of the automobiles of the time. Today Trico Products is one of the world's leading manufacturers of windshield wiping systems, windshield wiper blades and refills, with wiper plants on five continents.

Inventor William M. Folberth, in 1919, applied for a patent for an automatic windscreen wiper apparatus. The patent was granted in 1922. It was the first automatic mechanism. Trico later settled a patent dispute with Folberth and purchased Folberth's Cleveland company, the Folberth Auto Specialty Co. The new vacuum-powered system quickly became standard equipment on automobiles, and the vacuum principle was in use until about 1960. In the late 1950s, a feature common on modern vehicles first appeared, operating the wipers automatically for two or three passes when the washer button was pressed, making it unnecessary to manually turn them on as well. Today, an electronic timer is used, but originally a small vacuum cylinder mechanically linked to a switch provided the delay as the vacuum leaked off.

In 1963, the first intermittent wipers were invented by Robert Kearns, an engineering professor at Wayne State University in Detroit. The road to intermittent wipers began on his wedding night earlier in 1953, when an errant champagne cork shot into Kearn's left

eye, which eventually went almost completely blind. Nearly a decade later Kearns was driving his Ford Galaxie through a light rain, and the constant movement of the wiper blades irritated his already troubled vision. He got to thinking about the human eye, which has its own kind of wiper, the eyelid, that automatically closes and opens every few seconds. FInally in 1963, Kearns put his idea into action, building the first intermittent wiper system using off-the-shelf electronic components and offered it to Ford. The interval between wipes was determined by the rate of current flow into a capacitor. When the charge in the capacitor reached a certain voltage, the capacitor discharged, activating the wiper motor for one cycle. After extensive testing, Ford executives decided to offer Kearns's intermittent wipers as an option on the company's Mercury line beginning with the 1969 models. Kearns may not, in fact, be the original inventor. John Amos, an engineer for the UK automative engineering company, Lucas Industries was the first to file a patent for an intermittent wiper. United States Patent #3,262,042 was filed two years before Kearns filed United States Patent #3,351,836, and was issued in 1966 before Kearns' patent was issued in 1967. The difference is that the Amos patent describes an electromechanical device, whereas Kearns proposed a solid-state circuit.

In 1970, Saab Automobile introduced headlight wipers across the product range. These operated on a horizontal reciprocating mechanism, with a single motor. They were later superseded by a radial spindle action wiper mechanism, with individual motors on each headlamp. In March 1970, Citroën introduced rain-sensitive intermittent windscreen wipers on the SM model. When the intermittent function was selected, the wiper would make one swipe. If the windscreen was relatively dry, the wiper motor drew high current, which set the control circuit timer to delay the next wipe longest. If the motor drew little current, it indicated that the glass was wet, setting the timer to minimize the delay.

Bosch has the world's biggest windscreen wiper factory in Tienen, Belgium, which produces 350,000 wiper blades every day.

Power

Wipers may be powered by a variety of means, although most in existence today are powered by an electric motor through a series of mechanical components, typically two 4-bar linkages in series or parallel. Vehicles with air operated brakes sometimes use air operated wipers, run by bleeding a small amount of air pressure from the brake system to a small air operated motor mounted just above the windscreen. These wipers are activated by opening a valve which allows pressurized air to enter the motor.

Early wipers were often driven by a vacuum motor powered by manifold vacuum. This had the drawback that manifold vacuum alters depending on throttle position and is almost non-existent under wide-open throttle; the wipers would slow down or even stop. This problem was overcome somewhat by using a combined fuel/vacuum booster pump. Some cars, mostly from the 1960s and 1970s, had hydraulically driven wipers. On the earlier Citroën 2CV, the windscreen wipers were powered by a purely mechanical system: a cable connected to the transmission, to reduce cost this cable also powered the

speedometer. The wipers' speed was therefore variable with car speed. When the car was waiting at a crossroad, the wipers were not powered, thus a handle under the speedometer allowed the driver to power them by hand.

Geometry

Most wipers are of the pivot (or radial) type: they are attached to a single arm, which in turn is attached to the motor. These are commonly found on many cars, trucks, trains, boats, airplanes, etc.

Another type of wipers are the pantograph-based (see Fig. 6), which are used on many commercial vehicles, especially buses with large windscreens. Pantograph wipers feature two arms for each blade, with the blade assembly itself supported on a horizontal bar connecting the two arms. One of the arms is attached to the motor, while the other is on an idle pivot. The pantograph mechanism, while being more complex, allows the blade to cover more of the windscreen on each wipe. However, it also usually requires the wiper to be "parked" in the middle of the windscreen, where it may partially obstruct the driver's view when not in use. Some larger cars in the late '70s and early '80s, especially LH driver American cars, had a pantograph wiper on the driver's side, with a conventional pivot on the passenger side.

Mercedes-Benz pioneered a system called the Monoblade, based on cantilevers (see Fig. 5), in which a single arm extends outward to reach the top corners of the windscreen, and pulls in at the ends and middle of the stroke, sweeping out a somewhat 'W'-shaped path. This way, a single blade is able to cover more of the windscreen, and displace the residual streaks away from the center of the windscreen.

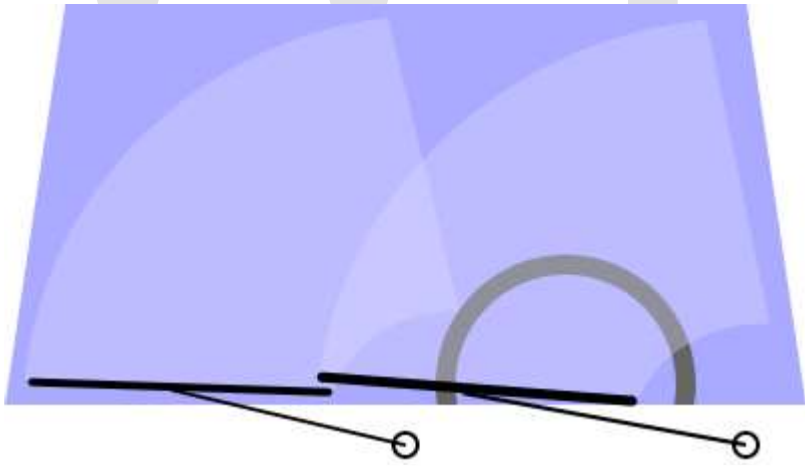


Fig. 1: most common geometry, found on vast majority of vehicles

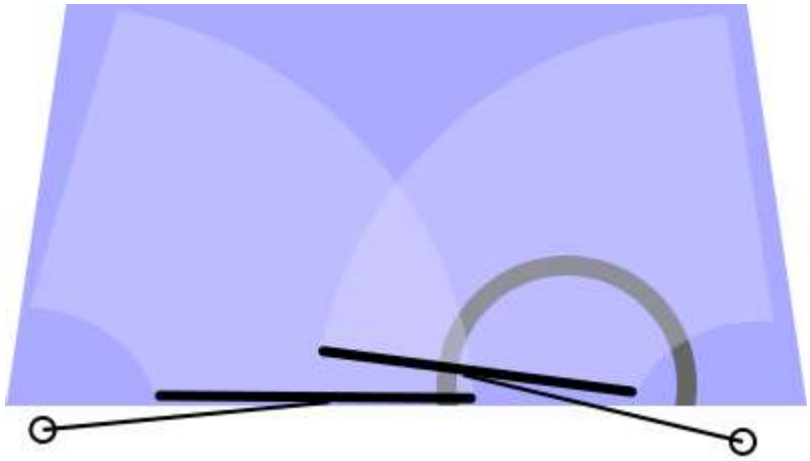


Fig. 2: Mercedes-Benz W114, W168, W169; VW Sharan, Honda Civic, some minivans, some school buses, Peugeot 307

Fig. 3: SEAT Altea, SEAT León Mk2, SEAT Toledo Mk3

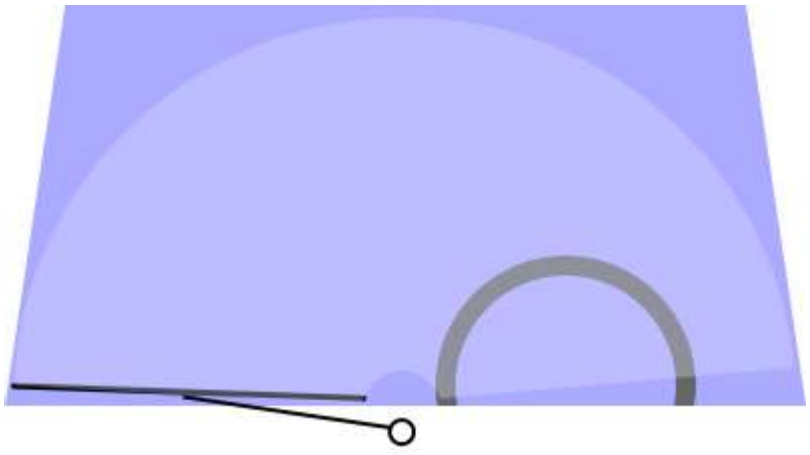


Fig. 4: VAZ-1111 Oka, 1990's Citroën Citroen AX

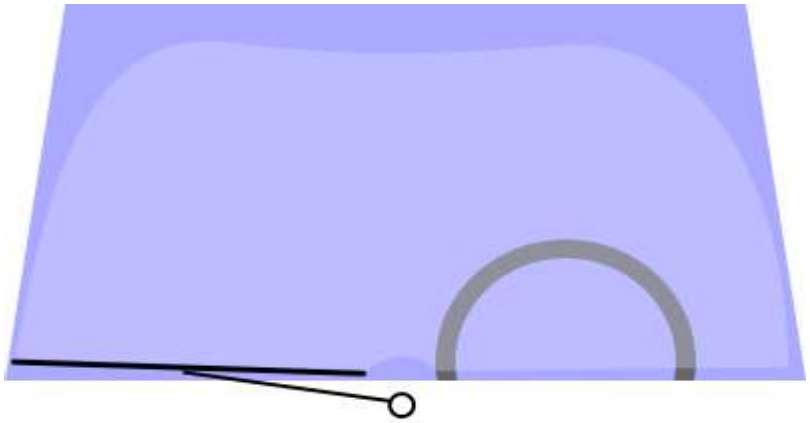


Fig. 5: Subaru XT, Mercedes-Benz W124, W201, W202, W210

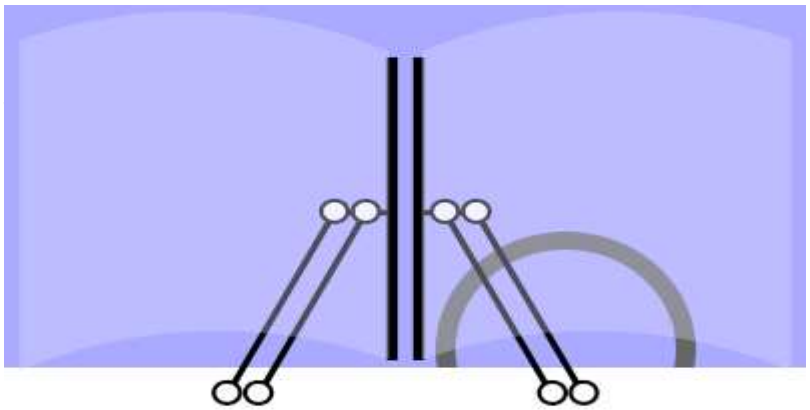


Fig. 6: Buses, some school buses, Mercedes-Benz O305

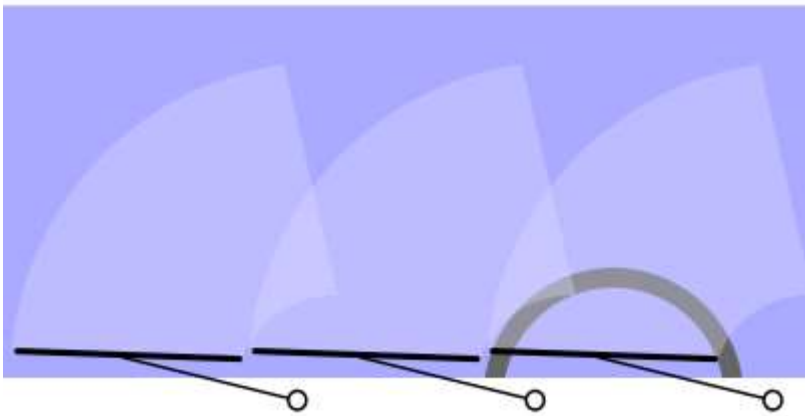


Fig. 7: MAN, Toyota FJ Cruiser, Jaguar E-type, MGB, MG Midget, Austin Healey Sprite, a 1968 US only ruling required a certain percentage of the windscreen to be wiped.

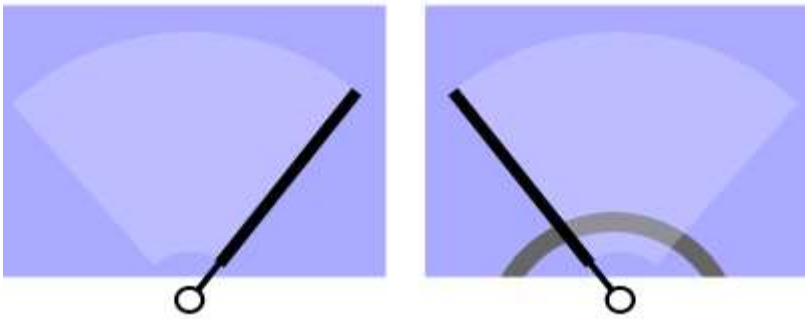


Fig. 8: obsolete, found on some older firetrucks

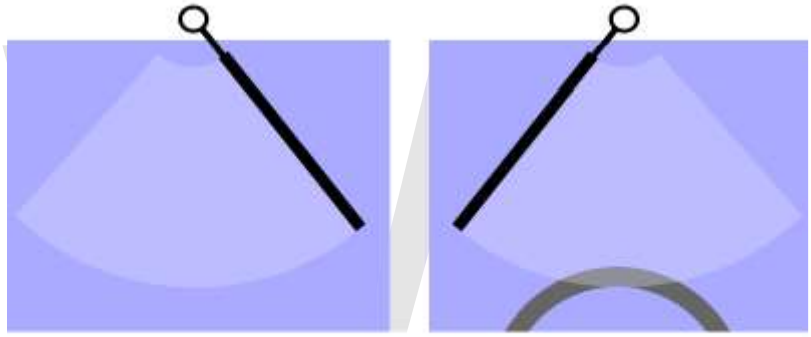


Fig. 9: US Military Wheeled Vehicles, some school buses

Other features

Some larger cars are equipped with "hidden" (or "depressed-park") wipers. When wipers are switched off, a "parking" mechanism or circuit moves the wipers to the lower extreme of the wiped area, near the bottom of the windscreen, but still in sight. To hide the wipers, the windscreen extends below the rear edge of the hood, and the wipers park themselves below the wiping range at the bottom of the windscreen, but out of sight.

Some vehicles have small wipers or washers on the headlights. In more modern vehicles, these have been replaced with a powerful jet spray, without wipers. Some vehicles are fitted with wipers (with or without washers) on the back window as well. Rear-window wipers are typically found on hatchbacks, station wagons, sport utility vehicles, minivans, and some sports cars. They were first implemented in the 1970s, but SUVs did not use them until the 1980s.

Windscreen washer

Most windscreen wipers operate together with a windscreen (or windshield) washer; a pump that supplies water and detergent (usually a blend called windshield washer fluid) from a tank to the windscreen through small nozzles, mounted on the hood or on the wipers, known as a 'wet-arm' system. In warmer climates, water may also work, but it can freeze in colder climates. Although automobile antifreeze is chemically similar to windscreen wiper fluid, it must not be used because it can damage paint. The earliest idea for having a windshield wiper unit hooked up to a windshield washer fluid reservoir was in 1931, Richland Auto Parts Co, Mansfield, Ohio

Rain-sensing wipers

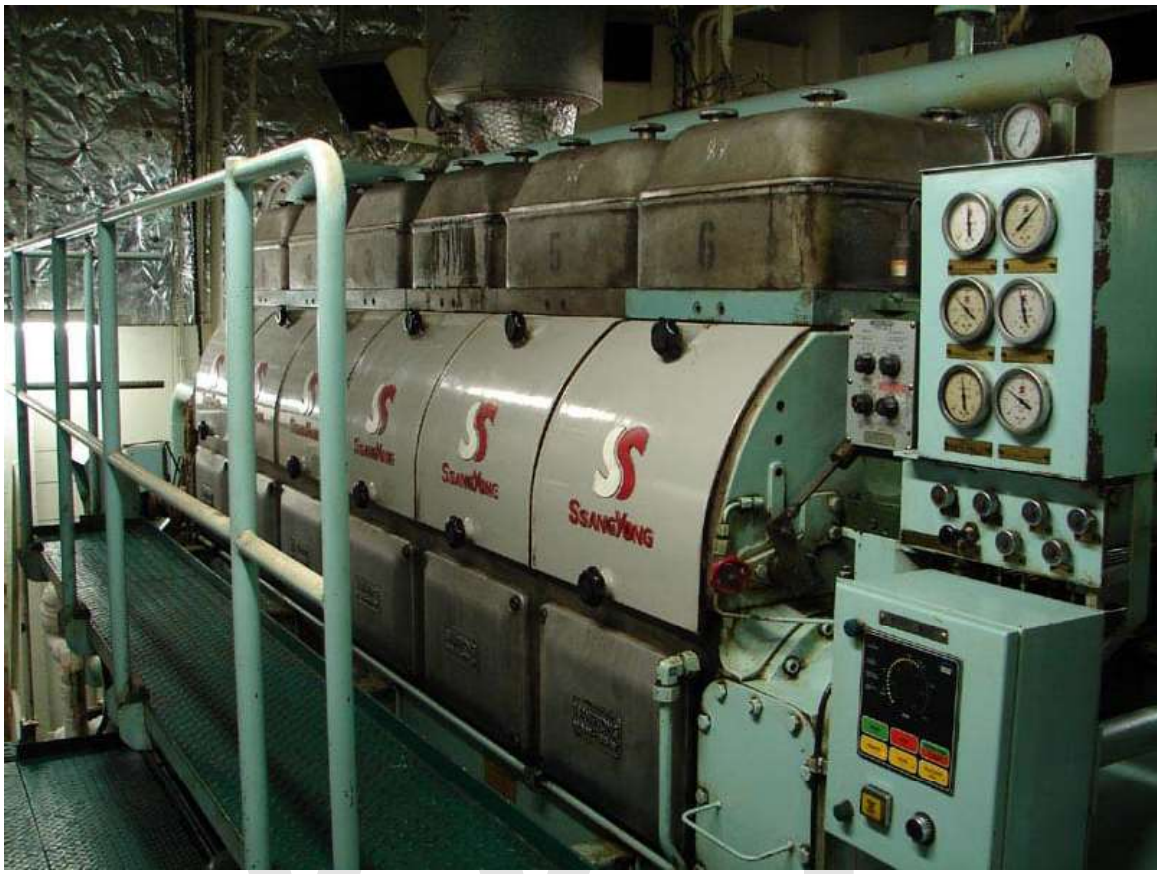
Vehicles are now available with driver-programmable *intelligent* (automatic) windscreen wipers that detect the presence and amount of rain using a rain sensor. The sensor automatically adjusts the speed of the blades according to the amount of rain detected. These usually have a manual override.

Rain-sensing windscreen wipers appeared on various models in the late 20th century, one of the first being Nissan's 200SX/Silvia. As of early 2006, rain-sensing wipers are optional or standard on all Cadillacs and most Volkswagen, and are available on many other main-stream manufacturers.

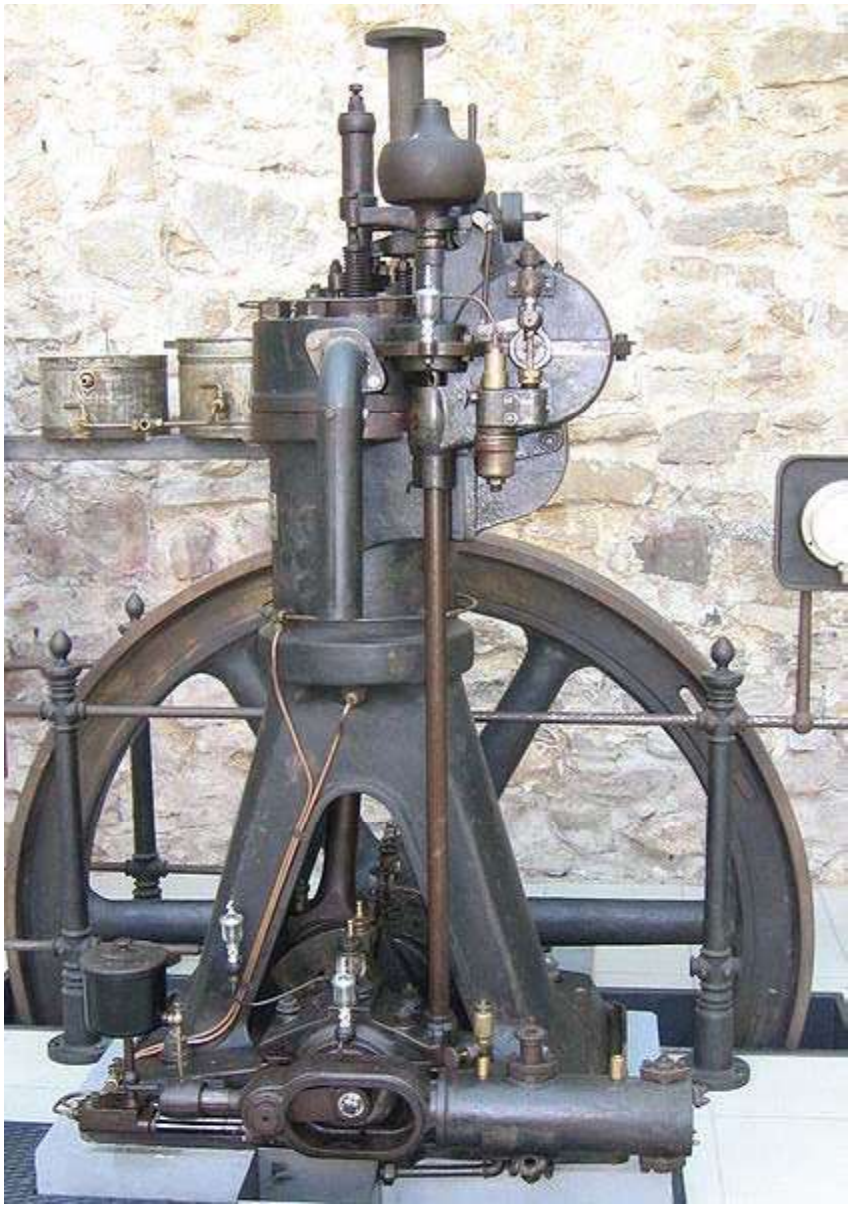
Chapter- 11

Diesel Engine

Diesel engines in a museum



Diesel generator on an oil tanker



A diesel engine built by MAN AG in 1906

A **diesel engine** (also known as a **compression-ignition engine** and sometimes capitalized as **Diesel engine**) is an internal combustion engine that uses the heat of compression to initiate ignition to burn the fuel, which is injected into the combustion chamber during the final stage of compression. This is in contrast to spark-ignition engines such as a petrol engine (gasoline engine) or gas engine (using a gaseous fuel as opposed to gasoline), which uses a spark plug to ignite an air-fuel mixture. The diesel engine is modeled on the Diesel cycle. The engine and thermodynamic cycle were both developed by Rudolf Diesel in 1897.

The diesel engine has the highest thermal efficiency of any regular internal or external combustion engine due to its very high compression ratio. Low-speed diesel engines (as

used in ships and other applications where overall engine weight is relatively unimportant) often have a thermal efficiency which exceeds 50 percent.

Diesel engines are manufactured in two stroke and four stroke versions. They were originally used as a more efficient replacement for stationary steam engines. Since the 1910s they have been used in submarines and ships. Use in locomotives, large trucks and electric generating plants followed later. In the 1930s, they slowly began to be used in a few automobiles. Since the 1970s, the use of diesel engines in larger on-road and off-road vehicles in the USA increased. As of 2007, about 50 percent of all new car sales in Europe are diesel.

The world's largest diesel engine is currently a Wärtsilä marine diesel of about 80 MW output.

History

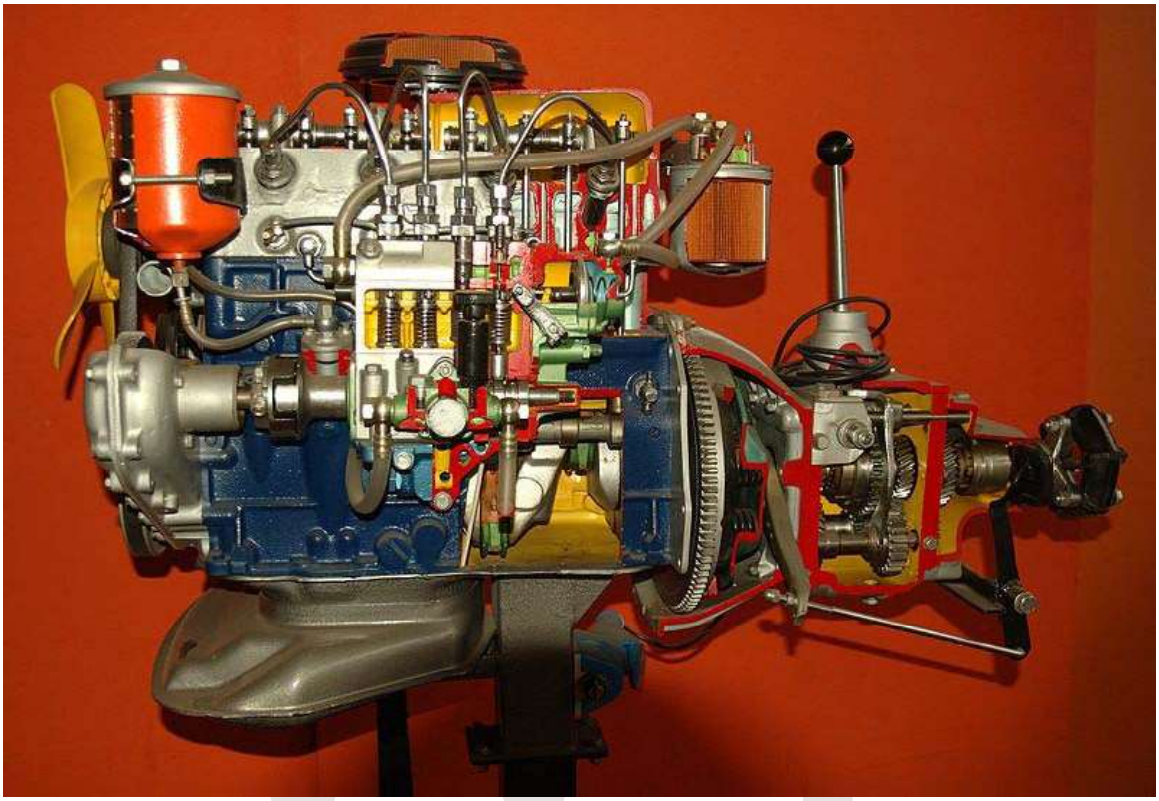
Rudolf Diesel, of German ethnicity, was born in 1858 in Paris where his parents were German immigrants. He was educated at Munich Polytechnic. After graduation he was employed as a refrigerator engineer, but his true love lay in engine design. Diesel designed many heat engines, including a solar-powered air engine. In 1892 he received patents in Germany, Switzerland, the United Kingdom and filed in the United States for "Method of and Apparatus for Converting Heat into Work". In 1893 he described a "slow-combustion engine" that first compressed air thereby raising its temperature above the igniting-point of the fuel, then gradually introducing fuel while letting the mixture expand "against resistance sufficiently to prevent an essential increase of temperature and pressure", then cutting off fuel and "expanding without transfer of heat". In 1894 and 1895 he filed patents and addenda in various countries for his Diesel engine; the first patents were issued in Spain (No.16,654), France (No.243,531) and Belgium (No.113,139) in December 1894, and in Germany (No.86,633) in 1895 and the United States (No.608,845) in 1898. He operated his first successful engine in 1897. His engine was the first to prove that fuel could be ignited without a spark.

Though best known for his invention of the pressure-ignited heat engine that bears his name, Rudolf Diesel was also a well-respected thermal engineer and a social theorist. Diesel's inventions have three points in common: they relate to heat transfer by natural physical processes or laws; they involve markedly creative mechanical design; and they were initially motivated by the inventor's concept of sociological needs. Rudolf Diesel originally conceived the diesel engine to enable independent craftsmen and artisans to compete with industry.

At Augsburg, on August 10, 1893, Rudolf Diesel's prime model, a single 10-foot (3.0 m) iron cylinder with a flywheel at its base, ran on its own power for the first time. Diesel spent two more years making improvements and in 1896 demonstrated another model with a theoretical efficiency of 75 percent, in contrast to the 10 percent efficiency of the steam engine. By 1898, Diesel had become a millionaire. His engines were used to power

pipelines, electric and water plants, automobiles and trucks, and marine craft. They were soon to be used in mines, oil fields, factories, and transoceanic shipping.

How diesel engines work



Diesel engine model, left side

Diesel engine model, right side

The diesel internal combustion engine differs from the gasoline powered Otto cycle by using highly compressed, hot air to ignite the fuel rather than using a spark plug (*compression ignition* rather than *spark ignition*).

In the true diesel engine, only air is initially introduced into the combustion chamber. The air is then compressed with a compression ratio typically between 15:1 and 22:1 resulting in 40-bar (4.0 MPa; 580 psi) pressure compared to 8 to 14 bars (0.80 to 1.4 MPa) (about 200 psi) in the petrol engine. This high compression heats the air to 550 °C (1,022 °F). At about the top of the compression stroke, fuel is injected directly into the compressed air in the combustion chamber. This may be into a (typically toroidal) void in the top of the piston or a *pre-chamber* depending upon the design of the engine. The fuel injector ensures that the fuel is broken down into small droplets, and that the fuel is distributed evenly. The heat of the compressed air vaporizes fuel from the surface of the droplets. The vapour is then ignited by the heat from the compressed air in the combustion chamber, the droplets continue to vaporise from their surfaces and burn, getting smaller, until all the fuel in the droplets has been burnt. The start of vaporisation causes a delay period during ignition, and the characteristic diesel knocking sound as the vapor reaches ignition temperature and causes an abrupt increase in pressure above the piston. The rapid expansion of combustion gases then drives the piston downward, supplying power to the

crankshaft. Engines for scale-model aeroplanes use a variant of the Diesel principle but premix fuel and air via a carburation system external to the combustion chambers.

As well as the high level of compression allowing combustion to take place without a separate ignition system, a high compression ratio greatly increases the engine's efficiency. Increasing the compression ratio in a spark-ignition engine where fuel and air are mixed before entry to the cylinder is limited by the need to prevent damaging pre-ignition. Since only air is compressed in a diesel engine, and fuel is not introduced into the cylinder until shortly before top dead centre (TDC), premature detonation is not an issue and compression ratios are much higher.

Early fuel injection systems

Diesel's original engine injected fuel with the assistance of compressed air, which atomized the fuel and forced it into the engine through a nozzle (a similar principle to an aerosol spray). The nozzle opening was closed by a pin valve lifted by the camshaft to initiate the fuel injection before top dead centre (TDC). This is called an air-blast injection. Driving the three stage compressor used some power but the efficiency and net power output was more than any other combustion engine at that time.

Diesel engines in service today raise the fuel to extreme pressures by mechanical pumps and deliver it to the combustion chamber by pressure-activated injectors without compressed air. With direct injected diesels, injectors spray fuel through 4 to 12 small orifices in its nozzle. The early air injection diesels always had a superior combustion without the sharp increase in pressure during combustion. Research is now being performed and patents are being taken out to again use some form of air injection to reduce the nitrogen oxides and pollution, reverting to Diesel's original implementation with its superior combustion and possibly quieter operation. In all major aspects, the modern diesel engine holds true to Rudolf Diesel's original design, that of igniting fuel by compression at an extremely high pressure within the cylinder. With much higher pressures and high technology injectors, present-day diesel engines use the so-called solid injection system applied by Herbert Akroyd Stuart for his hot bulb engine. The indirect injection engine could be considered the latest development of these low speed *hot bulb* ignition engines.

Fuel delivery

A vital component of all diesel engines is a mechanical or electronic governor which regulates the idling speed and maximum speed of the engine by controlling the rate of fuel delivery. Unlike Otto-cycle engines, incoming air is not throttled and a diesel engine without a governor cannot have a stable idling speed and can easily overspeed, resulting in its destruction. Mechanically governed fuel injection systems are driven by the engine's gear train. These systems use a combination of springs and weights to control fuel delivery relative to both load and speed. Modern electronically controlled diesel engines control fuel delivery by use of an electronic control module (ECM) or electronic control unit (ECU). The ECM/ECU receives an engine speed signal, as well as other

operating parameters such as intake manifold pressure and fuel temperature, from a sensor and controls the amount of fuel and start of injection timing through actuators to maximise power and efficiency and minimise emissions. Controlling the timing of the start of injection of fuel into the cylinder is a key to minimizing emissions, and maximizing fuel economy (efficiency), of the engine. The timing is measured in degrees of crank angle of the piston before top dead centre. For example, if the ECM/ECU initiates fuel injection when the piston is 10 degrees before TDC, the start of injection, or timing, is said to be 10° BTDC. Optimal timing will depend on the engine design as well as its speed and load.

Advancing the start of injection (injecting before the piston reaches to its SOI-TDC) results in higher in-cylinder pressure and temperature, and higher efficiency, but also results in elevated engine noise and increased oxides of nitrogen (NO_x) emissions due to higher combustion temperatures. Delaying start of injection causes incomplete combustion, reduced fuel efficiency and an increase in exhaust smoke, containing a considerable amount of particulate matter and unburned hydrocarbons.

Major advantages

Diesel engines have several advantages over other internal combustion engines:

- They burn less fuel than a petrol engine performing the same work, due to the engine's higher temperature of combustion and greater expansion ratio. Gasoline engines are typically 25 percent efficient while diesel engines can convert over 30 percent of the fuel energy into mechanical energy.
- They have no high-tension electrical ignition system to attend to, resulting in high reliability and easy adaptation to damp environments. The absence of coils, spark plug wires, etc., also eliminates a source of radio frequency emissions which can interfere with navigation and communication equipment, which is especially important in marine and aircraft applications.
- They can deliver much more of their rated power on a continuous basis than a petrol engine.
- The life of a diesel engine is generally about twice as long as that of a petrol engine due to the increased strength of parts used. Diesel fuel has better lubrication properties than petrol as well.

Bus powered by **biodiesel**

- Diesel fuel is considered safer than petrol in many applications. Although diesel fuel will burn in open air using a wick, it will not explode and does not release a large amount of flammable vapor. The low vapor pressure of diesel is especially advantageous in marine applications, where the accumulation of explosive fuel-air mixtures is a particular hazard. For the same reason, diesel engines are immune to vapor lock.
- For any given partial load the fuel efficiency (mass burned per energy produced) of a diesel engine remains nearly constant, as opposed to petrol and turbine engines which use proportionally more fuel with partial power outputs.
- They generate less waste heat in cooling and exhaust.
- Diesel engines can accept super- or turbo-charging pressure without any natural limit, constrained only by the strength of engine components. This is unlike petrol engines, which inevitably suffer detonation at higher pressure.
- The carbon monoxide content of the exhaust is minimal, therefore diesel engines are used in underground mines.
- Biodiesel is an easily synthesized, non-petroleum-based fuel (through transesterification) which can run directly in many diesel engines, while gasoline engines either need adaptation to run synthetic fuels or else use them as an additive to gasoline (e.g., ethanol added to gasohol), making diesel engines the clearly preferred choice for sustainability.

Mechanical and electronic injection

Many configurations of fuel injection have been used over the past century (1901–2000).

Most present day (2008) diesel engines make use of a camshaft, rotating at half crankshaft speed, lifted mechanical single plunger high pressure fuel pump driven by the engine crankshaft. For each cylinder, its plunger measures the amount of fuel and determines the timing of each injection. These engines use injectors that are very precise spring-loaded valves that open and close at a specific fuel pressure. For each cylinder a plunger pump is connected to an injector with a high pressure fuel line. Fuel volume for each single combustion is controlled by a slanted groove in the plunger which rotates only a few degrees releasing the pressure and is controlled by a mechanical governor, consisting of weights rotating at engine speed constrained by springs and a lever. The injectors are held open by the fuel pressure. On high speed engines the plunger pumps are together in one unit. Each fuel line should have the same length to obtain the same pressure delay.

A cheaper configuration on high speed engines with fewer than six cylinders is to use an axial-piston distributor pump, consisting of one rotating pump plunger delivering fuel to a valve and line for each cylinder (functionally analogous to points and distributor cap on an Otto engine). This contrasts with the more modern method of having a single fuel pump which supplies fuel constantly at high pressure with a common rail (single fuel line common) to each injector. Each injector has a solenoid operated by an electronic control unit, resulting in more accurate control of injector opening times that depend on other control conditions, such as engine speed and loading, and providing better engine performance and fuel economy. This design is also mechanically simpler than the combined pump and valve design, making it generally more reliable, and less noisy, than its mechanical counterpart.

Both mechanical and electronic injection systems can be used in either direct or indirect injection configurations.

Older diesel engines with mechanical injection pumps could be inadvertently run in reverse, albeit very inefficiently, as witnessed by massive amounts of soot being ejected from the air intake. This was often a consequence of push starting a vehicle using the wrong gear. Large ship diesels can run either way.

Indirect injection

An indirect injection diesel engine delivers fuel into a chamber off the combustion chamber, called a pre-chamber or ante-chamber, where combustion begins and then spreads into the main combustion chamber, assisted by turbulence created in the chamber. This system allows for a smoother, quieter running engine, and because combustion is assisted by turbulence, injector pressures can be lower, about 100 bar (10 MPa; 1,500 psi), using a single orifice tapered jet injector. Mechanical injection systems allowed high-speed running suitable for road vehicles (typically up to speeds of around 4,000 rpm). The pre-chamber had the disadvantage of increasing heat loss to the engine's cooling system, and restricting the combustion burn, which reduced the efficiency by 5–10 percent. Indirect injection engines were used in small-capacity, high-speed diesel engines in automotive, marine and construction uses from the 1950s, until

direct injection technology advanced in the 1980s. Indirect injection engines are cheaper to build and it is easier to produce smooth, quiet-running vehicles with a simple mechanical system. In road-going vehicles most prefer the greater efficiency and better controlled emission levels of direct injection. Indirect injection diesels can still be found in the many ATV diesel applications.

Direct injection

Modern diesel engines make use of one of the following direct injection methods:

Direct injection injectors are mounted in the top of the combustion chamber. The problem with these vehicles was the harsh noise they produced. Fuel consumption was about 15 to 20 percent lower than indirect injection diesels, which for some buyers was enough to compensate for the extra noise.

This type of engine was transformed by electronic control of the injection pump, pioneered by Fiat in 1986 (Croma). The injection pressure was still only around 300 bar (30 MPa; 4,400 psi), but the injection timing, fuel quantity, EGR and turbo boost were all electronically controlled. This gave more precise control of these parameters which eased refinement and lowered emissions.

Unit direct injection

Unit direct injection also injects fuel directly into the cylinder of the engine. In this system the injector and the pump are combined into one unit positioned over each cylinder controlled by the camshaft. Each cylinder has its own unit eliminating the high pressure fuel lines, achieving a more consistent injection. This type of injection system, also developed by Bosch, is used by Volkswagen AG in cars (where it is called a *Pumpe-Düse-System*—literally *pump-nozzle system*) and by Mercedes Benz ("PLD") and most major diesel engine manufacturers in large commercial engines (CAT, Cummins, Detroit Diesel, Volvo). With recent advancements, the pump pressure has been raised to 2,400 bar (240 MPa; 35,000 psi), allowing injection parameters similar to common rail systems.

Common rail direct injection

In common rail systems, the separate pulsing high pressure fuel line to each cylinder's injector is also eliminated. Instead, a high-pressure pump pressurizes fuel at up to 2,500 bar (250 MPa; 36,000 psi), in a "common rail". The common rail is a tube that supplies each computer-controlled injector containing a precision-machined nozzle and a plunger driven by a solenoid or piezoelectric actuator.

Cold weather

Starting

In cold weather, high speed diesel engines that are pre-chambered can be difficult to start because the mass of the cylinder block and cylinder head absorb the heat of compression, preventing ignition due to the higher surface-to-volume ratio. Pre-chambered engines therefore make use of small electric heaters inside the pre-chambers called glowplugs. These engines also generally have a higher compression ratio of 19:1 to 21:1. Low-speed and compressed-air-started larger and intermediate-speed diesels do not have glowplugs and compression ratios are around 16:1.

Some engines (e.g., some Cummins models) use resistive grid heaters in the intake manifold to warm the inlet air until the engine reaches operating temperature. Engine block heaters (electric resistive heaters in the engine block) connected to the utility grid are often used when an engine is turned off for extended periods (more than an hour) in cold weather to reduce startup time and engine wear. In the past, a wider variety of cold-start methods were used. Some engines, such as Detroit Diesel engines and Lister-Petter engines, used a system to introduce small amounts of ether into the inlet manifold to start combustion. Saab-Scania marine engines, Field Marshall tractors (among others) used slow-burning solid-fuel 'cigarettes' which were fitted into the cylinder head as a primitive glow plug.

Lucas developed the *Thermostart*, where an electrical heating element was combined with a small fuel valve in the inlet manifold. Diesel fuel slowly dripped from the valve onto the hot element and ignited. The flame heated the inlet manifold and when the engine was cranked, the flame was drawn into the cylinders to start combustion.

International Harvester developed a tractor in the 1930s that had a 7-litre 4-cylinder engine which started as a gasoline engine and ran on diesel after warming up. The cylinder head had valves which opened for a portion of the compression stroke to reduce the effective compression ratio, and a magneto produced the spark. An automatic ratchet system automatically disengaged the ignition system and closed the valves once the engine had run for 30 seconds. The operator then switched off the petrol fuel system and opened the throttle on the diesel injection system.

Recent direct-injection systems are advanced to the extent that pre-chambers systems are not needed by using a common rail fuel system with electronic fuel injection.

Gelling

Diesel fuel is also prone to *waxing* or *gelling* in cold weather; both are terms for the solidification of diesel oil into a partially crystalline state. The crystals build up in the fuel line (especially in fuel filters), eventually starving the engine of fuel and causing it to stop running. Low-output electric heaters in fuel tanks and around fuel lines are used to solve this problem. Also, most engines have a *spill return* system, by which any excess

fuel from the injector pump and injectors is returned to the fuel tank. Once the engine has warmed, returning warm fuel prevents waxing in the tank. Due to improvements in fuel technology with additives, waxing rarely occurs in all but the coldest weather when a mix of diesel and kerosene should be used to run a vehicle.

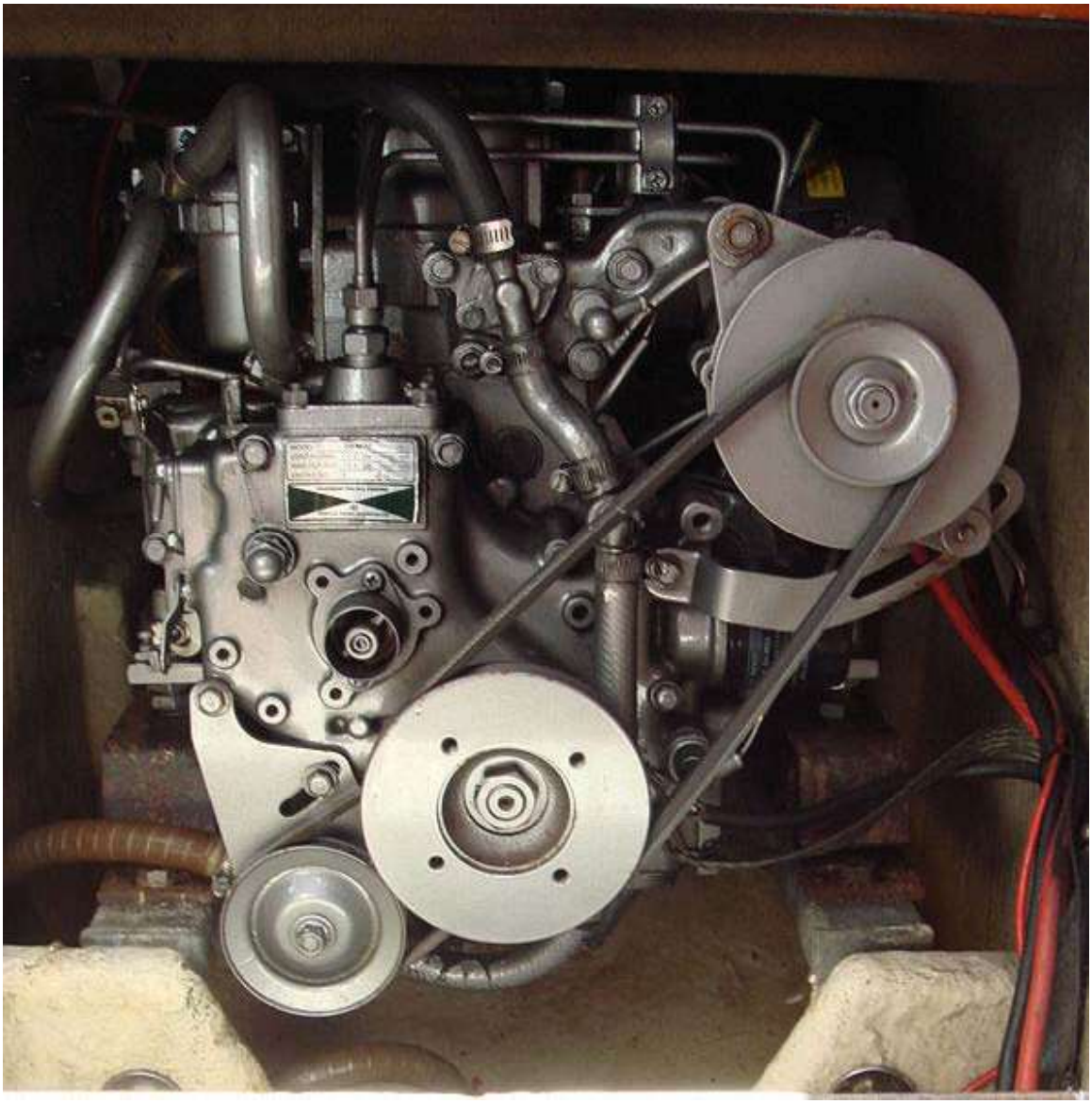
Types

Early

Rudolf Diesel intended his engine to replace the steam engine as the primary power source for industry. As such, diesel engines in the late 19th and early 20th centuries used the same basic layout and form as industrial steam engines, with long-bore cylinders, external valve gear, cross-head bearings and an open crankshaft connected to a large flywheel. Smaller engines would be built with vertical cylinders, while most medium- and large-sized industrial engines were built with horizontal cylinders, just as steam engines had been. Engines could be built with more than one cylinder in both cases. The largest early diesels resembled the triple-expansion steam reciprocating engine, being tens of feet high with vertical cylinders arranged in-line. These early engines ran at very slow speeds—partly due to the limitations of their air-blast injector equipment and partly so they would be compatible with the majority of industrial equipment designed for steam engines; maximum speeds of between 100 and 300 rpm were common. Engines were usually started by allowing compressed air into the cylinders to turn the engine, although smaller engines could be started by hand.

In the early decades of the 20th century, when large diesel engines were first being used, the engines took a form similar to the compound steam engines common at the time, with the piston being connected to the connecting rod by a crosshead bearing. Following steam engine practice some manufactures made double-acting two-stroke and four-stroke diesel engines to increase power output, with combustion taking place on both sides of the piston, with two sets of valve gear and fuel injection. While it produced large amounts of power and was very efficient, the double-acting diesel engine's main problem was producing a good seal where the piston rod passed through the bottom of the lower combustion chamber to the crosshead bearing, and no more were built. By the 1930s turbochargers were fitted to some engines. Crosshead bearings are still used to reduce the wear on the cylinders in large long-stroke main marine engines.

Modern



A Yanmar 2GM20 marine diesel engine, installed in a sailboat.

As with petrol engines, there are two classes of diesel engines in current use: two-stroke and four-stroke. The four-stroke type is the "classic" version, tracing its lineage back to Rudolf Diesel's prototype. It is also the most commonly used form, being the preferred power source for many motor vehicles, especially buses and trucks. Much larger engines, such as used for railroad locomotion and marine propulsion, are often two-stroke units, offering a more favourable power-to-weight ratio, as well as better fuel economy. The most powerful engines in the world are two-stroke diesels of mammoth dimensions.

Two-stroke diesel operation is similar to that of petrol counterparts, except that fuel is not mixed with air before induction, and the crankcase does not take an active role in the

cycle. The traditional two-stroke design relies upon a mechanically driven positive displacement blower to charge the cylinders with air before compression and ignition. The charging process also assists in expelling (scavenging) combustion gases remaining from the previous power stroke. The archetype of the modern form of the two-stroke diesel is the Detroit Diesel engine, in which the blower pressurizes a chamber in the engine block that is often referred to as the "air box". The (much larger) Electromotive prime mover used in EMD diesel-electric locomotives is built to the same principle.

In a two-stroke diesel engine, as the cylinder's piston approaches the bottom dead centre exhaust ports or valves are opened relieving most of the excess pressure after which a passage between the air box and the cylinder is opened, permitting air flow into the cylinder. The air flow blows the remaining combustion gases from the cylinder—this is the scavenging process. As the piston passes through bottom centre and starts upward, the passage is closed and compression commences, culminating in fuel injection and ignition. Refer to two-stroke diesel engines for more detailed coverage of aspiration types and supercharging of two-stroke diesel engines.

Normally, the number of cylinders are used in multiples of two, although any number of cylinders can be used as long as the load on the crankshaft is counterbalanced to prevent excessive vibration. The inline-six cylinder design is the most prolific in light to medium-duty engines, though small V8 and larger inline-four displacement engines are also common. Small-capacity engines (generally considered to be those below five litres in capacity) are generally four or six cylinder types, with the four cylinder being the most common type found in automotive uses. Five cylinder diesel engines have also been produced, being a compromise between the smooth running of the six cylinder and the space-efficient dimensions of the four cylinder. Diesel engines for smaller plant machinery, boats, tractors, generators and pumps may be four, three or two cylinder types, with the single cylinder diesel engine remaining for light stationary work. Direct reversible two-stroke marine diesels need at least three cylinders for reliable restarting forwards and reverse, while four-stroke diesels need at least six cylinders.

The desire to improve the diesel engine's power-to-weight ratio produced several novel cylinder arrangements to extract more power from a given capacity. The uniflow opposed-piston engine uses two pistons in one cylinder with the combustion cavity in the middle and gas in- and outlets at the ends. This makes a comparatively light, powerful, swiftly running and economic engine suitable for use in aviation. An example is the Junkers Jumo 204/205. The Napier Deltic engine, with three cylinders arranged in a triangular formation, each containing two opposed pistons, the whole engine having three crankshafts, is one of the better known.

Gas generator

Before 1950, Sulzer started experimenting with two-stroke engines with boost pressures as high as 6 atmospheres, in which all the output power was taken from an exhaust gas turbine. The two-stroke pistons directly drove air compressor pistons to make a positive displacement gas generator. Opposed pistons were connected by linkages instead of

crankshafts. Several of these units could be connected to provide power gas to one large output turbine. The overall thermal efficiency was roughly twice that of a simple gas turbine. This system was derived from Raúl Pateras Pescara's work on free-piston engines in the 1930s.

Advantages and disadvantages versus spark-ignition engines

Power and fuel economy

The MAN S80ME-C7 low speed diesel engines use 155 gram fuel per kWh for an overall energy conversion efficiency of 54.4 percent, which is the highest conversion of fuel into power by any internal or external combustion engine. Diesel engines are more efficient than gasoline (petrol) engines of the same power rating, resulting in lower fuel consumption. A common margin is 40 percent more miles per gallon for an efficient turbodiesel. For example, the current model Škoda Octavia, using Volkswagen Group engines, has a combined Euro rating of 6.2 L/100 km (38 miles per US gallon, 16 km/L) for the 102 bhp (76 kW) petrol engine and 4.4 L/100 km (54 mpg, 23 km/L) for the 105 bhp (78 kW) diesel engine.

However, such a comparison does not take into account that diesel fuel is denser and contains about 15 percent more energy by volume. Although the calorific value of the fuel is slightly lower at 45.3 MJ/kg (megajoules per kilogram) than petrol at 45.8 MJ/kg, liquid diesel fuel is significantly denser than liquid petrol. This is significant because volume of fuel, in addition to mass, is an important consideration in mobile applications. No vehicle has an unlimited volume available for fuel storage.

Adjusting the numbers to account for the energy density of diesel fuel, the overall energy efficiency is still about 20 percent greater for the diesel version.

While a higher compression ratio is helpful in raising efficiency, diesel engines are much more efficient than gasoline (petrol) engines when at low power and at engine idle. Unlike the petrol engine, diesels lack a butterfly valve (throttle) in the inlet system, which closes at idle. This creates parasitic loss and destruction of availability of the incoming air, reducing the efficiency of petrol engines at idle. In many applications, such as marine, agriculture, and railways, diesels are left idling and unattended for many hours, sometimes even days. These advantages are especially attractive in locomotives.

The average diesel engine has a poorer power-to-weight ratio than the petrol engine. This is because the diesel must operate at lower engine speeds and because it needs heavier, stronger parts to resist the operating pressure caused by the high compression ratio of the engine and the large amounts of torque generated to the crankshaft. In addition, diesels are often built with stronger parts to give them longer lives and better reliability, important considerations in industrial applications.

For most industrial or nautical applications, reliability is considered more important than light weight and high power. Diesel fuel is injected just before the power stroke. As a

result, the fuel cannot burn completely unless it has a sufficient amount of oxygen. This can result in incomplete combustion and black smoke in the exhaust if more fuel is injected than there is air available for the combustion process. Modern engines with electronic fuel delivery can adjust the timing and amount of fuel delivery (by changing the duration of the injection pulse), and so operate with less waste of fuel. In a mechanical system, the injection timing and duration must be set to be efficient at the anticipated operating rpm and load, and so the settings are less than ideal when the engine is running at any other RPM than what it is timed for. The electronic injection can "sense" engine revs, load, even boost and temperature, and continuously alter the timing to match the given situation. In the petrol engine, air and fuel are mixed for the entire compression stroke, ensuring complete mixing even at higher engine speeds.

Diesel engines usually have longer stroke lengths in order to achieve the necessary compression ratios. As a result piston and connecting rods are heavier and more force must be transmitted through the connecting rods and crankshaft to change the momentum of the piston. This is another reason that a diesel engine must be stronger for the same power output as a petrol engine.

Yet it is this characteristic that has allowed some enthusiasts to acquire significant power increases with turbocharged engines by making fairly simple and inexpensive modifications. A petrol engine of similar size cannot put out a comparable power increase without extensive alterations because the stock components cannot withstand the higher stresses placed upon them. Since a diesel engine is already built to withstand higher levels of stress, it makes an ideal candidate for performance tuning at little expense. However, it should be said that any modification that raises the amount of fuel and air put through a diesel engine will increase its operating temperature, which will reduce its life and increase service requirements. These are issues with newer, lighter, *high performance* diesel engines which are not "overbuilt" to the degree of older engines and they are being pushed to provide greater power in smaller engines. The addition of a turbocharger or supercharger to the engine greatly assists in increasing fuel economy and power output, mitigating the fuel-air intake speed limit mentioned above for a given engine displacement. Boost pressures can be higher on diesels than on petrol engines, due to the latter's susceptibility to knock, and the higher compression ratio allows a diesel engine to be more efficient than a comparable spark ignition engine. Because the burned gases are expanded further in a diesel engine cylinder, the exhaust gas is cooler, meaning turbochargers require less cooling, and can be more reliable, than with spark-ignition engines.

With a diesel, boost pressure is essentially unlimited. It is literally possible to run as much boost as the engine will physically stand before breaking apart.

The increased fuel economy of the diesel engine over the petrol engine means that the diesel produces less carbon dioxide (CO_2) per unit distance. Recent advances in production and changes in the political climate have increased the availability and awareness of biodiesel, an alternative to petroleum-derived diesel fuel with a much lower net-sum emission of CO_2, due to the absorption of CO_2 by plants used to produce the

fuel. Although concerns are now being raised as to the negative effect this is having on the world food supply, as the growing of crops specifically for biofuels takes up land that could be used for food crops and uses water that could be used by both humans and animals. The use of waste vegetable oil, sawmill waste from managed forests in Finland, and advances in the production of vegetable oil from algae demonstrate great promise in providing feed stocks for sustainable biodiesel that are not in competition with food production.

Diesel engines have a lower rotational speed than an equivalent size petrol engine because the diesel-air mixture burns slower than the petrol-air mixture. A combination of improved mechanical technology (such as multi-stage injectors which fire a short "pilot charge" of fuel into the cylinder to warm the combustion chamber before delivering the main fuel charge), higher injection pressures that have improved the atomisation of fuel into smaller droplets, and electronic control (which can adjust the timing and length of the injection process to optimise it for all speeds and temperatures) have mitigated most of these problems in the latest generation of common-rail designs, while greatly improving engine efficiency. Poor power and narrow torque bands have been addressed by superchargers, turbochargers, (especially variable geometry turbochargers), intercoolers, and a large efficiency increase from about 35 percent for IDI to 45 percent for the latest engines in the last 15 years.

Even though diesel engines have a theoretical fuel efficiency of 75 percent, in practice it is lower. Engines in large diesel trucks, buses, and newer diesel cars can achieve peak efficiencies around 45 percent, and could reach 55 percent efficiency in the near future. However, average efficiency over a driving cycle is lower than peak efficiency. For example, it might be 37 percent for an engine with a peak efficiency of 44 percent.

Emissions

Diesel exhaust has been found to contain a long list of toxic air contaminants. Among these pollutants, fine particle pollution is perhaps the most important as a cause of diesel's deleterious health effects.

Diesel engines produce very little carbon monoxide as they burn the fuel in excess air even at full load, at which point the quantity of fuel injected per cycle is still about 50 percent lean of stoichiometric. However, they can produce black soot (or more specifically diesel particulate matter) from their exhaust. The black smoke consists of carbon compounds that were not combusted, because of local low temperatures where the fuel is not fully atomized. These local low temperatures occur at the cylinder walls, and at the outside of large droplets of fuel. At these areas where it is relatively cold, the mixture is rich (contrary to the overall mixture which is lean). The rich mixture has less air to burn and some of the fuel turns into a carbon deposit. Modern car engines use a diesel particulate filter (DPF) to capture carbon particles and then intermittently burn them using extra fuel injected directly into the filter. This prevents carbon buildup at the expense of wasting a small quantity of fuel.

The full load limit of a diesel engine in normal service is defined by the "black smoke limit", beyond which point the fuel cannot be completely combusted. As the "black smoke limit" is still considerably lean of stoichiometric, it is possible to obtain more power by exceeding it, but the resultant inefficient combustion means that the extra power comes at the price of reduced combustion efficiency, high fuel consumption and dense clouds of smoke. This is only done in specialized applications (such as tractor pulling competitions) where these disadvantages are of little concern.

Likewise, when starting from cold, the engine's combustion efficiency is reduced because the cold engine block draws heat out of the cylinder in the compression stroke. The result is that fuel is not combusted fully, resulting in blue and white smoke and lower power outputs until the engine has warmed. This is especially the case with indirect injection engines, which are less thermally efficient. With electronic injection, the timing and length of the injection sequence can be altered to compensate for this. Older engines with mechanical injection can have mechanical and hydraulic governor control to alter the timing, and multi-phase electrically controlled glow plugs, that stay on for a period after start-up to ensure clean combustion—the plugs are automatically switched to a lower power to prevent their burning out.

Particles of the size normally called PM10 (particles of 10 micrometres or smaller) have been implicated in health problems, especially in cities. Some modern diesel engines feature diesel particulate filters, which catch the black soot and when saturated are automatically regenerated by burning the particles. Other problems associated with the exhaust gases (nitrogen oxides, sulfur oxides) can be mitigated with further investment and equipment; some diesel cars now have catalytic converters in the exhaust.

All diesel engine exhaust emissions can be significantly reduced by using biodiesel fuel. Oxides of nitrogen do increase from a vehicle using biodiesel, but they too can be reduced to levels below that of fossil fuel diesel, by changing fuel injection timing.

Power and torque

For commercial uses requiring towing, load carrying and other tractive tasks, diesel engines tend to have better torque characteristics. Diesel engines tend to have their torque peak quite low in their speed range (usually between 1600 and 2000 rpm for a small-capacity unit, lower for a larger engine used in a truck). This provides smoother control over heavy loads when starting from rest, and, crucially, allows the diesel engine to be given higher loads at low speeds than a petrol engine, making them much more economical for these applications. This characteristic is not so desirable in private cars, so most modern diesels used in such vehicles use electronic control, variable geometry turbochargers and shorter piston strokes to achieve a wider spread of torque over the engine's speed range, typically peaking at around 2500–3000 rpm.

While diesel engines tend to have more torque at lower engine speeds than petrol engines, diesel engines tend to have a narrower power band than petrol engines. Naturally aspirated diesels tend to lack power and torque at the top of their speed range. This

narrow band is a reason why a vehicle such as a truck may have a gearbox with as many as 18 or more gears, to allow the engine's power to be used effectively at all speeds. Turbochargers tend to improve power at high engine speeds; superchargers improve power at lower speeds; and variable geometry turbochargers improve the engine's performance equally by flattening the torque curve.

Noise

The characteristic noise of a diesel engine is variably called diesel clatter, diesel nailing, or diesel knock. Diesel clatter is caused largely by the diesel combustion process, the sudden ignition of the diesel fuel when injected into the combustion chamber causes a pressure wave. Engine designers can reduce diesel clatter through: indirect injection; pilot or pre-injection; injection timing; injection rate; compression ratio; turbo boost; and exhaust gas recirculation (EGR). Common rail diesel injection systems permit multiple injection events as an aid to noise reduction. Diesel fuels with a higher cetane rating modify the combustion process and reduce diesel clatter. CN (Cetane number) can be raised by distilling higher quality crude oil, by catalyzing a higher quality product or by using a cetane improving additive. Some oil companies market high cetane or premium diesel. Biodiesel has a higher cetane number than petrodiesel, typically 55CN for 100% biodiesel.

A combination of improved mechanical technology such as multi-stage injectors which fire a short "pilot charge" of fuel into the cylinder to initiate combustion before delivering the main fuel charge, higher injection pressures that have improved the atomisation of fuel into smaller droplets, and electronic control (which can adjust the timing and length of the injection process to optimise it for all speeds and temperatures), have mostly mitigated these problems in the latest generation of common-rail designs, while improving engine efficiency.

Reliability

The lack of an electrical ignition system greatly improves the reliability. The high durability of a diesel engine is also due to its overbuilt nature, a benefit that is magnified by the lower rotating speeds in diesels. Diesel fuel is a better lubricant than petrol so is less harmful to the oil film on piston rings and cylinder bores; it is routine for diesel engines to cover 250,000 miles (400,000 km) or more without a rebuild.

Due to the greater compression force required and the increased weight of the stronger components, starting a diesel engine is harder. More torque is required to push the engine through compression.

Either an electrical starter or an air-start system is used to start the engine turning. On large engines, pre-lubrication and slow turning of an engine, as well as heating, are required to minimise the amount of engine damage during initial start-up and running. Some smaller military diesels can be started with an explosive cartridge, called a Coffman starter, which provides the extra power required to get the machine turning. In

the past, Caterpillar and John Deere used a small petrol *pony* engine in their tractors to start the primary diesel engine. The pony engine heated the diesel to aid in ignition and used a small clutch and transmission to spin up the diesel engine. Even more unusual was an International Harvester design in which the diesel engine had its own carburetor and ignition system, and started on petrol. Once warmed up, the operator moved two levers to switch the engine to diesel operation, and work could begin. These engines had very complex cylinder heads, with their own petrol combustion chambers, and were vulnerable to expensive damage if special care was not taken (especially in letting the engine cool before turning it off).

Quality and variety of fuels

Petrol/gasoline engines are limited in the variety and quality of the fuels they can burn. Older petrol engines fitted with a carburetor required a volatile fuel that would vaporise easily to create the necessary air-fuel ratio for combustion. Because both air and fuel are admitted to the cylinder, if the compression ratio of the engine is too high or the fuel too volatile (with too low an octane rating), the fuel will ignite under compression, as in a diesel engine, before the piston reaches the top of its stroke. This pre-ignition causes a power loss and over time major damage to the piston and cylinder. The need for a fuel that is volatile enough to vaporise but not too volatile (to avoid pre-ignition) means that petrol engines will only run on a narrow range of fuels. There has been some success at dual-fuel engines that use petrol and ethanol, petrol and propane, and petrol and methane.

In diesel engines, a mechanical injector system vaporizes the fuel directly into the combustion chamber or a pre-combustion chamber (as opposed to a Venturi jet in a carburetor, or a Fuel injector in a fuel injection system vaporising fuel into the intake manifold or intake runners as in a petrol engine). This *forced vaporisation* means that less-volatile fuels can be used. More crucially, because only air is inducted into the cylinder in a diesel engine, the compression ratio can be much higher as there is no risk of pre-ignition provided the injection process is accurately timed. This means that cylinder temperatures are much higher in a diesel engine than a petrol engine, allowing less volatile fuels to be used.

Diesel fuel is a form of light fuel oil, very similar to kerosene/paraffin, but diesel engines, especially older or simple designs that lack precision electronic injection systems, can run on a wide variety of other fuels. Some of the most common alternatives are Jet A-1 type jet fuel or vegetable oil from a very wide variety of plants. Some engines can be run on vegetable oil without modification, and most others require fairly basic alterations. Biodiesel is a pure diesel-like fuel refined from vegetable oil and can be used in nearly all diesel engines. Requirements for fuels to be used in diesel engines are the ability of the fuel to flow along the fuel lines, the ability of the fuel to lubricate the injector pump and injectors adequately, and its ignition qualities (ignition delay, cetane number). Inline mechanical injector pumps generally tolerate poor-quality or bio-fuels better than distributor-type pumps. Also, indirect injection engines generally run more satisfactorily on bio-fuels than direct injection engines. This is partly because an indirect injection engine has a much greater 'swirl' effect, improving vaporisation and combustion of fuel,

and because (in the case of vegetable oil-type fuels) lipid depositions can condense on the cylinder walls of a direct-injection engine if combustion temperatures are too low (such as starting the engine from cold).

It is often reported that Diesel designed his engine to run on peanut oil, but this is not the case. Diesel stated in his published papers, "at the Paris Exhibition in 1900 (*Exposition Universelle*) there was shown by the Otto Company a small diesel engine, which, at the request of the French Government ran on Arachide (earth-nut or pea-nut) oil, and worked so smoothly that only a few people were aware of it. The engine was constructed for using mineral oil, and was then worked on vegetable oil without any alterations being made. The French Government at the time thought of testing the applicability to power production of the Arachide, or earth-nut, which grows in considerable quantities in their African colonies, and can easily be cultivated there." Diesel himself later conducted related tests and appeared supportive of the idea.

Most large marine diesels (often called *cathedral engines* due to their size) run on heavy fuel oil (sometimes called "bunker oil"), which is a thick, viscous and almost flameproof fuel which is very safe to store and cheap to buy in bulk as it is a waste product from the petroleum refining industry. The fuel must be heated to thin it out (often by the exhaust header) and is often passed through multiple injection stages to vaporise it.

Fuel and fluid characteristics

Diesel engines can operate on a variety of different fuels, depending on configuration, though the eponymous diesel fuel derived from crude oil is most common. The engines can work with the full spectrum of crude oil distillates, from natural gas, alcohols, petrol, wood gas to the *fuel oils* from diesel oil to residual fuels.

The type of fuel used is a combination of service requirements, and fuel costs. Good-quality diesel fuel can be synthesised from vegetable oil and alcohol. Diesel fuel can be made from coal or other carbon base using the Fischer-Tropsch process. Biodiesel is growing in popularity since it can frequently be used in unmodified engines, though production remains limited. Recently, biodiesel from coconut, which can produce a very promising coco methyl ester (CME), has characteristics which enhance lubricity and combustion giving a regular diesel engine without any modification more power, less particulate matter or black smoke, and smoother engine performance. The Philippines pioneers in the research on Coconut based CME with the help of German and American scientists. Petroleum-derived diesel is often called *petrodiesel* if there is need to distinguish the source of the fuel.

Pure plant oils are increasingly being used as a fuel for cars, trucks and remote combined heat and power generation especially in Germany where hundreds of decentralised small- and medium-sized oil presses cold press oilseed, mainly rapeseed, for fuel. There is a Deutsches Institut für Normung fuel standard for rapeseed oil fuel.

Residual fuels are the "dregs" of the distillation process and are a thicker, heavier oil, or oil with higher viscosity, which are so thick that they are not readily pumpable unless heated. Residual fuel oils are cheaper than clean, refined diesel oil, although they are dirtier. Their main considerations are for use in ships and very large generation sets, due to the cost of the large volume of fuel consumed, frequently amounting to many tonnes per hour. The poorly refined biofuels straight vegetable oil (SVO) and waste vegetable oil (WVO) can fall into this category, but can be viable fuels on non common rail or TDI PD diesels with the simple conversion of fuel heating to 80 to 100 degrees Celsius to reduce viscosity, and adequate filtration to OEM standards. Engines using these heavy oils have to start and shut down on standard diesel fuel, as these fuels will not flow through fuel lines at low temperatures. Moving beyond that, use of low-grade fuels can lead to serious maintenance problems because of their high sulphur and lower lubrication properties. Most diesel engines that power ships like supertankers are built so that the engine can safely use low-grade fuels due to their separate cylinder and crankcase lubrication.

Normal diesel fuel is more difficult to ignite and slower in developing fire than petrol because of its higher flash point, but once burning, a diesel fire can be fierce.

Fuel contaminants such as dirt and water are often more problematic in diesel engines than in petrol engines. Water can cause serious damage, due to corrosion, to the injection pump and injectors; and dirt, even very fine particulate matter, can damage the injection pumps due to the close tolerances that the pumps are machined to. All diesel engines will have a fuel filter (usually much finer than a filter on a petrol engine), and a water trap. The water trap (which is sometimes part of the fuel filter) often has a float connected to a warning light, which warns when there is too much water in the trap, and must be drained before damage to the engine can result. The fuel filter must be replaced much more often on a diesel engine than on a petrol engine, changing the fuel filter every 2-4 oil changes is not uncommon for some vehicles.

Safety

Fuel flammability

Diesel fuel has low flammability, leading to a low risk of fire caused by fuel in a vehicle equipped with a diesel engine.

In yachts diesels are used because petrol engines generate combustible vapors, which can accumulate in the bottom of the vessel, sometimes causing explosions. Therefore ventilation systems on petrol powered vessels are required.

The United States Army and NATO use only diesel engines and turbines because of fire hazard. Although neither Gasoline nor Diesel is explosive in liquid form, both can create an explosive air/vapor mix under the right conditions. However, Diesel fuel is less prone due to its lower vapor pressure, which is an indication of evaporation rate. The Material Safety Data Sheet for Ultra-Low Sulfur Diesel fuel indicates a vapor explosion hazard for Diesel indoors, outdoors, or in sewers.

US Army gasoline-engined tanks during World War II were nicknamed Ronsons, because of their greater likelihood of catching fire when damaged by enemy fire. (Although tank fires were usually caused by detonation of the ammunition rather than fuel.)

Maintenance hazards

Fuel injection introduces potential hazards in engine maintenance due to the high fuel pressures used. Residual pressure can remain in the fuel lines long after an injection-equipped engine has been shut down. This residual pressure must be relieved, and if it is done so by external bleed-off, the fuel must be safely contained. If a high-pressure diesel fuel injector is removed from its seat and operated in open air, there is a risk to the operator of injury by hypodermic jet-injection, even with only 100 psi pressure. The first known such injury occurred in 1937 during a diesel engine maintenance operation.

Diesel applications

The characteristics of diesel have different advantages for different applications.

Passenger cars

Diesel engines have long been popular in bigger cars and this is spreading to smaller cars. Diesel engines tend to be more economical at regular driving speeds and are much better at city speeds. Their reliability and life-span tend to be better (as detailed). Some 40% or more of all cars sold in Europe are diesel-powered where they are considered a low CO_2 option. Mercedes-Benz in conjunction with Robert Bosch GmbH produced diesel-powered passenger cars starting in 1936 and very large numbers are used all over the world (often as "Grande Taxis" in the Third World).

Railroad rolling stock

Diesel engines have eclipsed steam engines as the prime mover on all non-electrified railroads in the industrialized world. The first diesel locomotives appeared in the early 20th century, and diesel multiple units soon after.

While electric locomotives have now replaced the diesel locomotive almost completely on passenger traffic in Europe and Asia, diesel is still today very popular for cargo-hauling freight trains and on tracks where electrification is not feasible.

Most modern diesel locomotives are actually diesel-electric locomotives: the diesel engine is used to power an electric generator that in turn powers electric traction engines with no mechanical connection between diesel engine and traction.

Other transport uses

Larger transport applications (trucks, buses, etc.) also benefit from the diesel's reliability and high torque output. Diesel displaced paraffin (or tractor vaporising oil, TVO) in most parts of the world by the end of the 1950s with the U.S. following some 20 years later.

In merchant ships and boats, the same advantages apply with the relative safety of diesel fuel an additional benefit. The German pocket battleships were the largest diesel warships, but the German torpedo-boats known as E-boats (*Schnellboot*) of the Second World War were also diesel craft. Conventional submarines have used them since before the First World War, relying on the almost total absence of carbon monoxide in the exhaust. American World War II diesel-electric submarines operated on two-stroke cycle as opposed to the four-stroke cycle that other navies used.

Military fuel standardisation

NATO has a single vehicle fuel policy and has selected diesel for this purpose. The use of a single fuel simplifies wartime logistics. NATO and the United States Marine Corps have even been developing a diesel military motorcycle based on a Kawasaki off road motorcycle, with a purpose designed naturally aspirated direct injection diesel at Cranfield University in England, to be produced in the USA, because motorcycles were the last remaining gasoline-powered vehicle in their inventory. Before this, a few civilian motorcycles had been built using adapted stationary diesel engines, but the weight and cost disadvantages generally outweighed the efficiency gains.

Engine speeds

Within the diesel engine industry, engines are often categorized by their rotational speeds into three unofficial groups:

- High speed engines,
- medium speed engines, and
- slow speed engines

High and medium speed engines are predominantly four stroke engines. Medium speed engines are physically larger than high speed engines and can burn lower grade (slower burning) fuel than high speed engines. Slow speed engines are predominantly large two stroke crosshead engines, hence very different from high and medium speed engines. Due to the lower rotational speed of slow and medium speed engines, there is more time for combustion during the power stroke of the cycle, and these engine are capable of utilising lower fuel grades (slower burning) fuels than high speed engines.

High-speed engines

High-speed (approximately 1,000 rpm and greater) engines are used to power trucks (lorries), buses, tractors, cars, yachts, compressors, pumps and small electrical generators.

As of 2008, most high-speed engines have direct injection. Many modern engines, particularly in on-highway applications, have common rail direct injection, which is cleaner burning.

Medium-speed engines

Medium speed engines are used in large electrical generators, ship propulsion and mechanical drive applications such as large compressors or pumps. Medium speed diesel engines operate on either diesel fuel or heavy fuel oil by direct injection in the same manner as low speed engines.

Engines used in electrical generators run at approximately 300 to 1000 rpm and are optimized to run at a set synchronous speed depending on the generation frequency (50 or 60 hertz) and provide a rapid response to load changes. Typical synchronous speeds for modern medium speed engines are 500/514 rpm (50/60 Hz), 600 rpm (both 50 and 60 Hz), 720/750 rpm, and 900/1000 rpm.

As of 2009, the largest medium speed engines in current production have outputs up to approximately 20 MW (27,000 hp). and are supplied by companies like MAN B&W, Wärtsilä, and Rolls-Royce (who acquired Ulstein Bergen Diesel in 1999). Most medium speed engines produced are four-stroke machines, however there are some two-stroke medium speed engines such as by EMD (Electro-Motive Diesel), and the Fairbanks Morse OP (Opposed-piston engine) type.

Typical cylinder bore size for medium speed engines ranges from 20 cm to 50 cm, and engine configurations typically are offered ranging from in-line 4 cylinder units to V configuration 20 cylinder units. Most larger medium speed engines are started with compressed air direct on pistons, using an air distributor, as opposed to a pneumatic starting motor acting on the flywheel, which tends to be used for smaller engines. There is no definitive engine size cut-off point for this.

It should also be noted that most major manufacturers of medium speed engines make natural gas fueled versions of their diesel engines, which in fact operate on the Otto cycle, and require spark ignition, typically provided with a spark plug. There are also dual (diesel/natural gas/coal gas) fuel versions of medium and low speed diesel engines using a lean fuel air mixture and a small injection of diesel fuel (so called "pilot fuel") for ignition. In case of a gas supply failure or maximum power demand these engines will instantly switch back to full diesel fuel operation.

Low-speed engines

The MAN B&W 5S50MC 5-cylinder, 2-stroke, low-speed marine diesel engine. This particular engine is found aboard a 29,000 tonne chemical carrier.

Also known as *slow-speed*, or traditionally *oil engines*, the largest diesel engines are primarily used to power ships, although there are a few land-based power generation units as well. These extremely large two-stroke engines have power outputs up to approximately 85 MW (114,000 hp), operate in the range from approximately 60 to 200 rpm and are up to 15 m (50 ft) tall, and can weigh over 2,000 short tons (1,800 t). They typically use direct injection running on cheap low-grade heavy fuel, also known as *Bunker C* fuel, which requires heating in the ship for tanking and before injection due to the fuel's high viscosity. The heat for fuel heating is often provided by waste heat recovery boilers located in the exhaust ducting of the engine, which produce the steam required for fuel heating. Provided the heavy fuel system is kept warm and circulating, engines can be started and stopped on heavy fuel.

Large and medium marine engines are started with compressed air directly applied to the pistons. Air is applied to cylinders to start the engine forwards or backwards because they are normally directly connected to the propeller without clutch or gearbox, and to provide reverse propulsion either the engine must be run backwards or the ship will utilise an

adjustable propeller. At least three cylinders are required with two-stroke engines and at least six cylinders with four-stroke engines to provide torque every 120 degrees.

Companies such as MAN B&W Diesel, (formerly Burmeister & Wain) and Wärtsilä (which acquired Sulzer Diesel) design such large low speed engines. They are unusually narrow and tall due to the addition of a crosshead bearing. As of 2007, the 14 cylinder Wärtsilä-Sulzer 14RTFLEX96-C turbocharged two-stroke diesel engine built by Wärtsilä licensee Doosan in Korea is the most powerful diesel engine put into service, with a cylinder bore of 960 mm (37.8 in) delivering 114,800 hp (85.6 MW). It was put into service in September 2006, aboard the world's largest container ship *Emma Maersk* which belongs to the A.P. Moller-Maersk Group. Typical bore size for low speed engines ranges from approximately 35 to 98 cm (14 to 39 in). As of 2008, all produced low speed engines with crosshead bearings are in-line configurations; no Vee versions have been produced.

Supercharging and turbocharging

Most diesels are now turbocharged and some are both turbo charged and supercharged. Because diesels do not have fuel in the cylinder before combustion is initiated, more than one bar (100 kPa) of air can be loaded in the cylinder without preignition. A turbocharged engine can produce significantly more power than a naturally aspirated engine of the same configuration, as having more air in the cylinders allows more fuel to be burned and thus more power to be produced. A supercharger is powered mechanically by the engine's crankshaft, while a turbocharger is powered by the engine exhaust, not requiring any mechanical power. Turbocharging can improve the fuel economy of diesel engines by recovering waste heat from the exhaust, increasing the excess air factor, and increasing the ratio of engine output to friction losses. A two-stroke engine does not have an exhaust and intake stroke. These are performed when the piston is at the bottom of the cylinder. Therefore large two-stroke engines have a piston pump, or electrical driven turbo at startup. Smaller two stroke engines (for example, Detroit 71 series) are fitted with turbochargers and a mechanically driven supercharger. Because turbocharged or supercharged engines produce more power for a given engine size as compared to naturally aspirated engines, attention must be paid to the mechanical design of components, lubrication, and cooling to handle the power. Pistons are usually cooled with lubrication oil sprayed on the bottom of the piston. Large diesels may use water, sea water, or oil supplied through telescoping pipes attached to the cross head.

Other applications

- Aircraft diesel engine
- Motorcycles

Current and future developments

As of 2008, many common rail and unit injection systems already employ new injectors using stacked piezoelectric wafers in lieu of a solenoid, giving finer control of the injection event.

Variable geometry turbochargers have flexible vanes, which move and let more air into the engine depending on load. This technology increases both performance and fuel economy. Boost lag is reduced as turbo impeller inertia is compensated for.

Accelerometer pilot control (APC) uses an accelerometer to provide feedback on the engine's level of noise and vibration and thus instruct the ECU to inject the minimum amount of fuel that will produce quiet combustion and still provide the required power (especially while idling).

The next generation of common rail diesels is expected to use variable injection geometry, which allows the amount of fuel injected to be varied over a wider range, and variable valve timing similar to that on petrol engines. Particularly in the United States, coming tougher emissions regulations present a considerable challenge to diesel engine manufacturers. Ford's HyTrans Project has developed a system which starts the ignition in 400 ms, saving a significant amount of fuel on city routes, and there are other methods to achieve even more efficient combustion, such as homogeneous charge compression ignition, being studied.

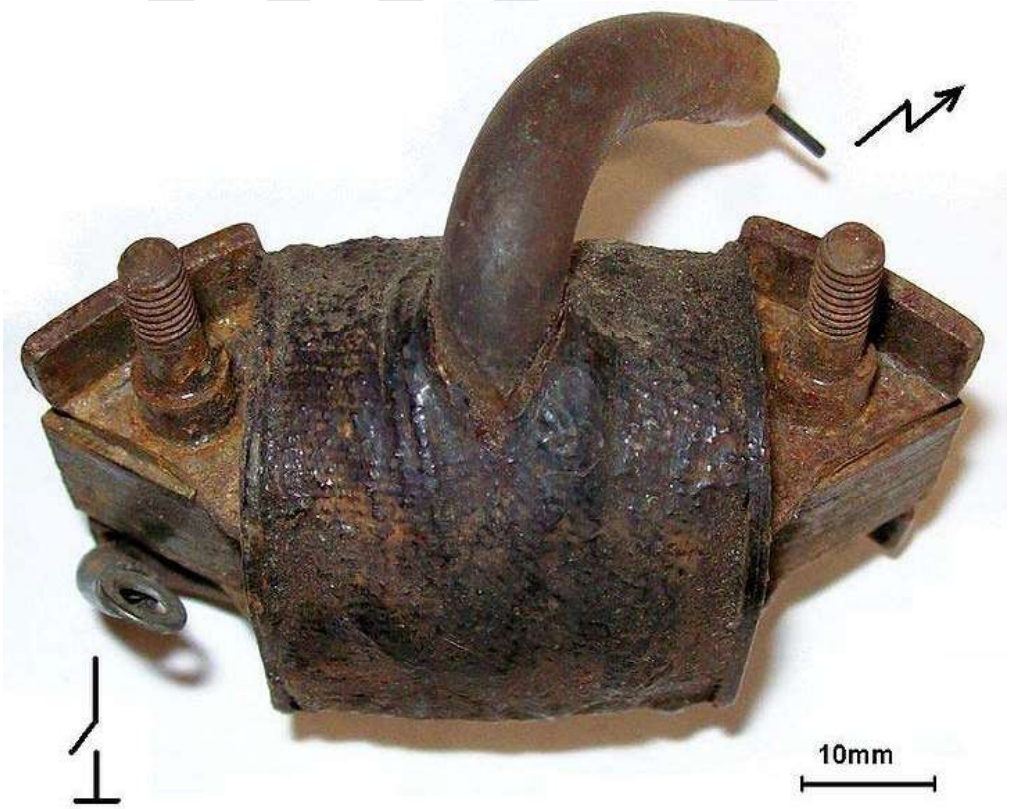
Chapter- 12

Ignition System

An **ignition system** is a system for igniting a fuel-air mixture. It is best known in the field of internal combustion engines but also has other applications, e.g. in oil-fired and gas-fired boilers. The earliest internal combustion engines used a flame, or a heated tube, for ignition but these were quickly replaced by systems using an electric spark.

History

Magneto systems



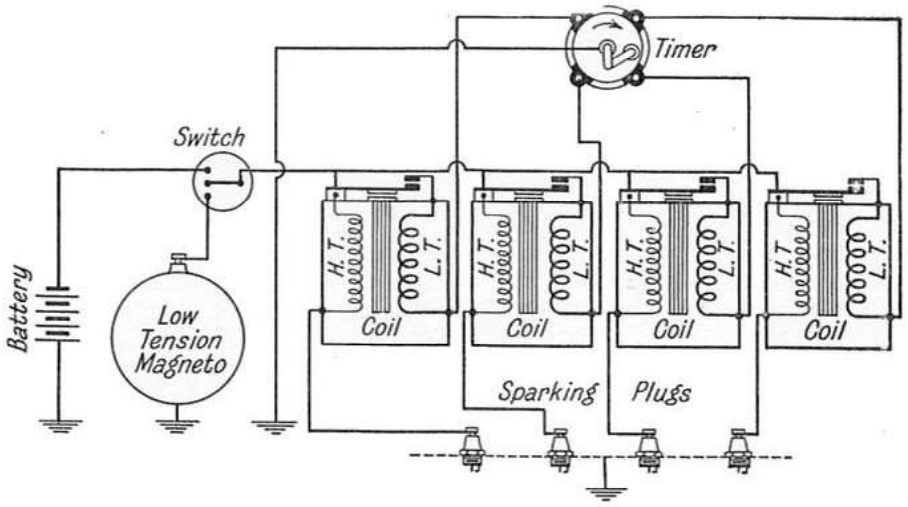
Magneto ignition coil

The simplest form of spark ignition is that using a magnet. The engine spins a magnet inside a coil, or, in the earlier designs, a coil inside a fixed magnet, and also operates a contact breaker, interrupting the current and causing the voltage to be increased sufficiently to jump a small gap. The spark plugs are connected directly from the magneto output. Early magnetos had one coil, with the contact breaker (sparking plug) inside the combustion chamber. In about 1902, Bosch introduced a double-coil magneto, with a fixed sparking plug, and the contact breaker outside the cylinder. Magnetos are not used in modern cars, but because they generate their own electricity they are often found on piston-engined aircraft engines and small engines such as those found in mopeds, lawnmowers, snowblowers, chainsaws, etc. where a battery-based electrical system is not present for any combination of necessity, weight, cost, and reliability reasons.

Magnetos were used on the small engine's ancestor, the stationary "hit or miss" engine which was used in the early twentieth century, on older gasoline or distillate farm tractors before battery starting and lighting became common, and on aircraft piston engines. Magnetos were used in these engines because their simplicity and self-contained operation was more reliable, and because magnetos weighed less than having a battery and dynamo or alternator.

Aircraft engines usually have multiple magnetos to provide redundancy in the event of a failure. Some older automobiles had both a magneto system and a battery actuated system running simultaneously to ensure proper ignition under all conditions with the limited performance each system provided at the time.This gave the benefits of easy starting (from the battery system) with reliable sparking at speed (from the magneto).

Switchable systems



Ford Model T ignition circuit

The output of a magneto depends on the speed of the engine, and therefore starting can be problematic. Some magnetos include an impulse system, which spins the magnet quickly at the proper moment, making easier starting at slow cranking speeds. Some engines, such as aircraft but also the Ford Model T, used a system which relied on non rechargeable dry cells, (similar to a large flashlight battery, and which was not maintained by a charging system as on modern automobiles) to start the engine or for starting and running at low speed. The operator would manually switch the ignition over to magneto operation for high speed operation.

In order to provide high voltage for the spark from the low voltage batteries, a 'tickler' was used, which was essentially a larger version of the once widespread electric buzzer. With this apparatus, the direct current passes through an electromagnetic coil which pulls open a pair of contact points, interrupting the current; the magnetic field collapses, the spring-loaded points close again, the circuit is reestablished, and the cycle repeats rapidly. The rapidly collapsing magnetic field, however, induces a high voltage across the coil which can only relieve itself by arcing across the contact points; while in the case of the buzzer this is a problem as it causes the points to oxidize and/or weld together, in the case of the ignition system this becomes the source of the high voltage to operate the spark plugs.

In this mode of operation, the coil would "buzz" continuously, producing a constant train of sparks. The entire apparatus was known as the 'Model T spark coil' (in contrast to the modern ignition coil which is *only* the actual coil component of the system). Long after the demise of the Model T as transportation they remained a popular self-contained source of high voltage for electrical home experimenters, appearing in articles in magazines such as *Popular Mechanics* and projects for school science fairs as late as the early 1960s. In the UK these devices were commonly known as trembler coils and were popular in cars pre-1910, and also in commercial vehicles with large engines until around 1925 to ease starting.

The Model T (built into the flywheel) differed from modern implementations by not providing high voltage directly at the output; the maximum voltage produced was about 30 volts, and therefore also had to be run through the spark coil to provide high enough voltage for ignition, as described above, although the coil would not "buzz" continuously in this case, only going through one cycle per spark. In either case, the low voltage was switched to the appropriate spark plug by the *'timer'* mounted on the front of the engine. This performed the equivalent function to the modern distributor, although by directing the low voltage, not the high voltage as for the distributor. The timing of the spark was adjustable by rotating this mechanism through a lever mounted on the steering column. As the precise timing of the spark depends on *both* the 'timer' and the trembler contacts within the coil, this is less consistent than the breaker points of the later distributor. However for the low speed and the low compression of such early engines, this imprecise timing was acceptable.

Battery-operated ignition

With the universal adaptation of electrical starting for automobiles, and the concomitant availability of a large battery to provide a constant source of electricity, magneto systems were abandoned for systems which interrupted current at battery voltage, used an ignition coil (a type of autotransformer) to step the voltage up to the needs of the ignition, and a distributor to route the ensuing pulse to the correct spark plug at the correct time.

The first reliable battery operated ignition was developed by the Dayton Engineering Laboratories Co. (Delco) and introduced in the 1910 Cadillac. This ignition was developed by Charles Kettering and was a wonder in its day. It consisted of a single coil, points (the switch), a capacitor and a distributor set up to allocate the spark from the ignition coil timed to the correct cylinder. The coil was basically an autotransformer set up to step up the low (6 or 12 V) voltage supply to the high ignition voltage required to jump a spark plug gap.

The points allow the coil to charge magnetically and then, when they are opened by a cam arrangement, the magnetic field collapses and a large (20 kV or greater) voltage is produced. The capacitor is used to absorb the back EMF from the magnetic field in the coil to minimize point contact burning and maximize point life. The Kettering system became the primary ignition system for many years in the automotive industry due to its lower cost, higher reliability and relative simplicity.

Modern ignition systems

The ignition system is typically controlled by a key operated Ignition switch.

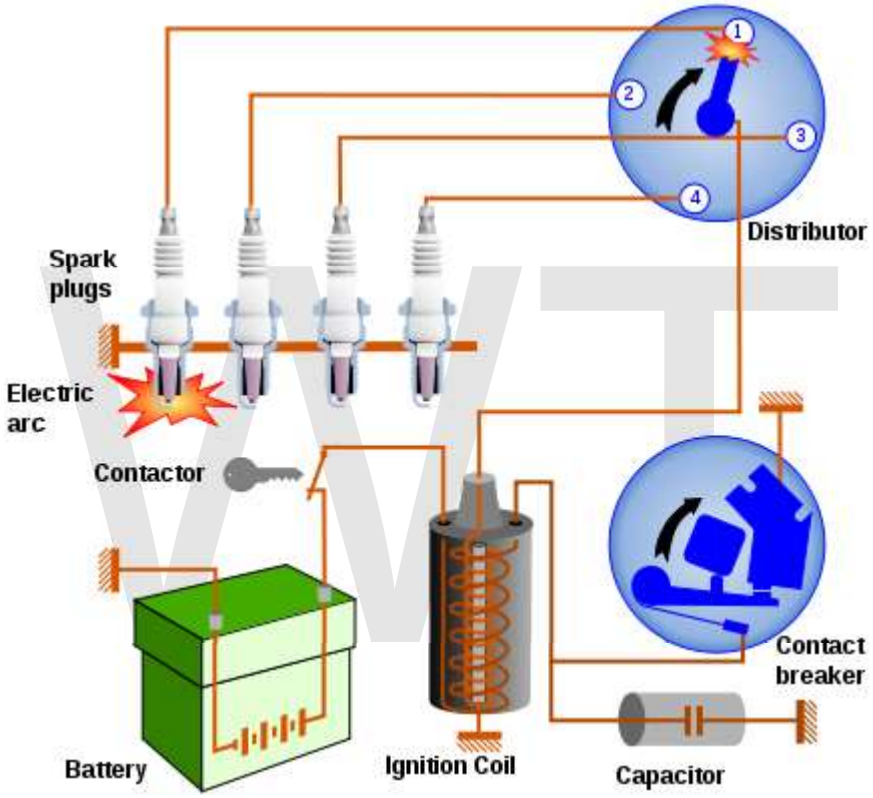
Mechanically timed ignition

Distributor cap

Most four-stroke engines have used a mechanically timed electrical ignition system. The heart of the system is the distributor. The distributor contains a rotating cam driven by the engine's drive, a set of breaker points, a condenser, a rotor and a distributor cap. External to the distributor is the ignition coil, the spark plugs and wires linking the distributor to the spark plugs and ignition coil.

The system is powered by a lead-acid battery, which is charged by the car's electrical system using a dynamo or alternator. The engine operates contact breaker points, which interrupt the current to an induction coil (known as the ignition coil).

The ignition coil consists of two transformer windings sharing a common magnetic core—the primary and secondary windings. An alternating current in the primary induces alternating magnetic field in the coil's core. Because the ignition coil's secondary has far more windings than the primary, the coil is a step-up transformer which induces a much higher voltage across the secondary windings. For an ignition coil, one end of windings of both the primary and secondary are connected together. This common point is connected to the battery (usually through a current-limiting ballast resistor). The other end of the primary is connected to the points within the distributor. The other end of the secondary is connected, via the distributor cap and rotor, to the spark plugs.



Ignition Circuit Diagram - Mechanically Timed Ignition

The ignition firing sequence begins with the points (or contact breaker) closed. A steady charge flows from the battery, through the current-limiting resistor, through the coil primary, across the closed breaker points and finally back to the battery. This steady current produces a magnetic field within the coil's core. This magnetic field forms the energy reservoir that will be used to drive the ignition spark.

As the engine turns, so does the cam inside the distributor. The points ride on the cam so that as the engine turns and reaches the top of the engine's compression cycle, a high point in the cam causes the breaker points to open. This breaks the primary winding's circuit and abruptly stops the current through the breaker points. Without the steady current through the points, the magnetic field generated in the coil immediately and

rapidly collapses. This change in the magnetic field induces a high voltage in the coil's secondary windings.

At the same time, current exits the coil's primary winding and begins to charge up the capacitor ("condenser") that lies across the now-open breaker points. This capacitor and the coil's primary windings form an oscillating LC circuit. This LC circuit produces a damped, oscillating current which bounces energy between the capacitor's electric field and the ignition coil's magnetic field. The oscillating current in the coil's primary, which produces an oscillating magnetic field in the coil, extends the high voltage pulse at the output of the secondary windings. This high voltage thus continues beyond the time of the initial field collapse pulse. The oscillation continues until the circuit's energy is consumed.

The ignition coil's secondary windings are connected to the distributor cap. A turning rotor, located on top of the breaker cam within the distributor cap, sequentially connects the coil's secondary windings to one of the several wires leading to each cylinder's spark plug. The extremely high voltage from the coil's secondary -– often higher than 1000 volts—causes a spark to form across the gap of the spark plug. This, in turn, ignites the compressed air-fuel mixture within the engine. It is the creation of this spark which consumes the energy that was originally stored in the ignition coil's magnetic field.

High performance engines with eight or more cylinders that operate at high r.p.m. (such as those used in motor racing) demand both a higher rate of spark and a higher spark energy than the simple ignition circuit can provide. This problem is overcome by using either of these adaptations:

- **Two complete sets of coils, breakers and condensers** can be provided - one set for each half of the engine, which is typically arranged in V-8 or V-12 configuration. Although the two ignition system halves are electrically independent, they typically share a single distributor which in this case contains two breakers driven by the rotating cam, and a rotor with two isolated conducting planes for the two high voltage inputs.
- A single breaker driven by a cam and a return spring is limited in spark rate by the onset of contact bounce or float at high rpm. This limit can be overcome by substituting for the breaker a **pair of breakers** that are connected electrically in series but spaced on opposite sides of the cam so they are driven out of phase. Each breaker then switches at half the rate of a single breaker and the "dwell" time for current buildup in the coil is maximized since it is shared between the breakers. The Lamborghini V-12 engine has both these adaptations and therefore uses two ignition coils and a single distributor that contains 4 contact breakers.

A distributor-based system is not greatly different from a magneto system except that more separate elements are involved. There are also advantages to this arrangement. For example, the position of the contact breaker points relative to the engine angle can be changed a small amount dynamically, allowing the ignition timing to be automatically

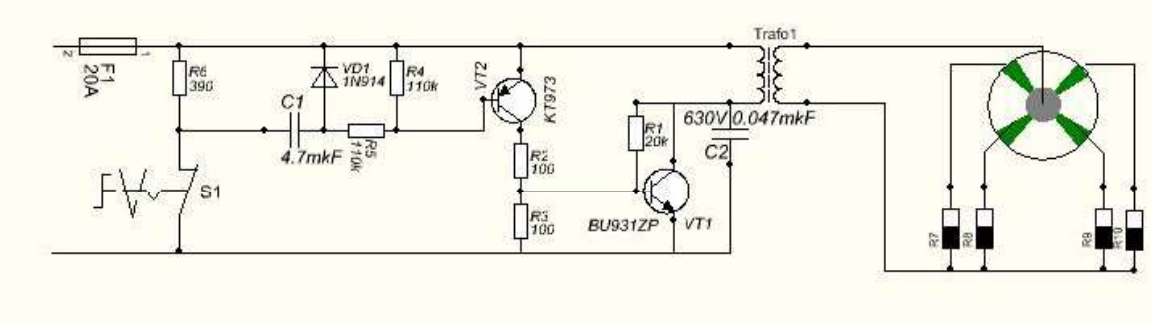
advanced with increasing revolutions per minute (RPM) and/or increased manifold vacuum, giving better efficiency and performance.

However it is necessary to check periodically the maximum opening gap of the breaker(s), using a feeler gauge, since this mechanical adjustment affects the "dwell" time during which the coil charges, and breakers should be re-dressed or replaced when they have become pitted by electric arcing. This system was used almost universally until the late 1970s, when electronic ignition systems started to appear.

Electronic ignition

The disadvantage of the mechanical system is the use of breaker points to interrupt the low-voltage high-current through the primary winding of the coil; the points are subject to mechanical wear where they ride the cam to open and shut, as well as oxidation and burning at the contact surfaces from the constant sparking. They require regular adjustment to compensate for wear, and the opening of the contact breakers, which is responsible for spark timing, is subject to mechanical variations.

In addition, the spark voltage is also dependent on contact effectiveness, and poor sparking can lead to lower engine efficiency. A mechanical contact breaker system cannot control an average ignition current of more than about 3 A while still giving a reasonable service life, and this may limit the power of the spark and ultimate engine speed.



Example of a basic electronic ignition system

Electronic ignition (EI) solves these problems. In the initial systems, points were still used but they handled only a low current which was used to control the high primary current through a solid state switching system. Soon, however, even these contact breaker points were replaced by an angular sensor of some kind - either optical, where a vaned rotor breaks a light beam, or more commonly using a Hall effect sensor, which responds to a rotating magnet mounted on the distributor shaft. The sensor output is shaped and processed by suitable circuitry, then used to trigger a switching device such as a thyristor, which switches a large current through the coil.

The first electronic ignition (a cold cathode type) was tested in 1948 by Delco-Remy, while Lucas introduced a transistorized ignition in 1955, which was used on BRM and Coventry Climax Formula One engines in 1962. The aftermarket began offering EI that year, with both the AutoLite Electric Transistor 201 and Tung-Sol EI-4 being available. Pontiac became the first automaker to offer an optional EI, the breakerless magnetic pulse-triggered Delcotronic, on some 1963 models; it was also available on some Corvettes. Ford fitted a Lucas system on the Lotus 25s entered at Indianapolis the next year, ran a fleet test in 1964, and began offering optional EI on some models in 1965. Beginning in 1958, Earl W. Meyer at Chrysler worked on EI, continuing until 1961 and resulting in use of EI on the company's NASCAR hemis in 1963 and 1964.

Prest-O-Lite's CD-65, which relied on capacitance discharge (CD), appeared in 1965, and had "an unprecedented 50,000 mile warranty." (This differs from the non-CD Prest-O-Lite system introduced on AMC products in 1972, and made standard equipment for the 1975 model year.) A similar CD unit was available from Delco in 1966, which was optional on Oldsmobile, Pontiac, and GMC vehicles in the 1967 model year. Also in 1967, Motorola debuted their breakerless CD system.

FIAT became the first company to offer standard EI, in 1968, followed by Chrysler (after a 1971 trial) in 1973 and by Ford and GM in 1975.

In 1967, Prest-O-Lite made a "Black Box" ignition amplifier, intended to take the load off of the distributor's breaker points during high r.p.m. runs, which was used by Dodge and Plymouth on their factory Super Stock Coronet and Belvedere and drag racers. This amp was installed on the interior-side of the cars' firewall, and had a duct which provided outside air to cool the amp. The rest of the system (distributor and spark plugs) remains as for the mechanical system. The lack of moving parts compared with the mechanical system leads to greater reliability and longer service intervals. Chrysler introduced breakerless ignition in mid-1971 as an option for its 340 V8 and the 426 Street Hemi. For the 1972 model year, the system became standard on its high-performance engines (the 340 cu in (5.6 l) and the four-barrel carburetor-equipped 400 hp (298 kW) 400 cu in (7 l)) and was an option on its 318 cu in (5.2 l), 360 cu in (5.9 l), two-barrel 400 cu in (6.6 l), and low-performance 440 cu in (7.2 l) . Breakerless Ignition was standardised across the model range for 1973. For older cars, it is usually possible to retrofit an EI system in place of the mechanical one. In some cases, a modern distributor will fit into the older engine with no other modifications, like the H.E.I. distributor made by General Motors, and the aforementioned Chrysler-built electronic ignition (with an "Orange Box" amplifier and a faster-advance curve distributor).

Other innovations are currently available on various cars. In some models, rather than one central coil, there are individual coils on each spark plug, sometimes known as direct ignition or coil on plug (COP). This allows the coil a longer time to accumulate a charge between sparks, and therefore a higher energy spark. A variation on this has each coil handle two plugs, on cylinders which are 360 degrees out of phase (and therefore reach TDC at the same time); in the four-cycle engine this means that one plug will be sparking during the end of the exhaust stroke while the other fires at the usual time, a so-called

"wasted spark" arrangement which has no drawbacks apart from faster spark plug erosion; the paired cylinders are 1/4 and 2/3. Other systems do away with the distributor as a timing apparatus and use a magnetic crank angle sensor mounted on the crankshaft to trigger the ignition at the proper time.

During the 1980s, electronic ignition systems were developed alongside other improvements such as fuel injection systems. After a while it became logical to combine the functions of fuel control and ignition into one electronic system known as an engine control unit. However on older vehicles this was not possible and now a common electronic ignition system for classic cars is the Powerspark electronic ignition.

Digital electronic ignitions

At the turn of the 21st century digital electronic ignition modules became available for small engines on such applications as chainsaws, string trimmers, leaf blowers, and lawn mowers. This was made possible by low cost, high speed, and small footprint microcontrollers. Digital electronic ignition modules can be designed as either capacitor discharge ignition (CDI) or inductive discharge ignition (IDI) systems. Capacitive discharge digital ignitions store charged energy for the spark in a capacitor within the module that can be released to the spark plug at virtually any time throughout the engine cycle via a control signal from the microprocessor. This allows for greater timing flexibility, and engine performance; especially when designed hand-in-hand with the engine carburetor.

Engine management

In an Engine Management System (EMS), electronics control fuel delivery, ignition timing and firing order. Primary sensors on the system are engine angle (crank or Top Dead Center (TDC) position), airflow into the engine and throttle demand position. The circuitry determines which cylinder needs fuel and how much, opens the requisite injector to deliver it, then causes a spark at the right moment to burn it. Early EMS systems used analogue computer circuit designs to accomplish this, but as embedded systems became fast enough to keep up with the changing inputs at high revolutions, digital systems started to appear.

Some designs using EMS retain the original coil, distributor and spark plugs found on cars throughout history. Other systems dispense with the distributor and individual coils mounted directly atop each spark plug. This removes the need for both distributor and high-tension leads, both components with a poor record for long-term reliability.

Modern EMSs read in data from various sensors about the crank position, manifold temperature, manifold pressure (or air mass flow), throttle position, fuel mixture via the O2 sensor and sometimes the unit will read data from knock sensors and exhaust gas temperature sensors. The EMS then uses collected data to precisely determine how much fuel to deliver and when and thus how far to advance the ignition timing. With electronic ignition systems, individual cylinders can have their own individual ignition timing so

that timing can be as aggressive as possible per cylinder without fuel detonation. As a result, sophisticated electronic ignition systems can be both more fuel efficient, and produce better performance, over their counterparts.

Turbine and jet engines

Turbine engines have a capacitor discharge ignition system using one or more ignitor plugs, which are only used at startup or in case the combustor(s) flame goes out. Rocket engines have particularly demanding ignitions systems- if prompt ignition does not occur the chamber can fill with excess fuel and oxidiser and significant overpressure can occur (a 'hard start'). Rockets often employ pyrotechnic devices that place flames across the face of the injector plate, or, alternatively, self-ignition chemicals.

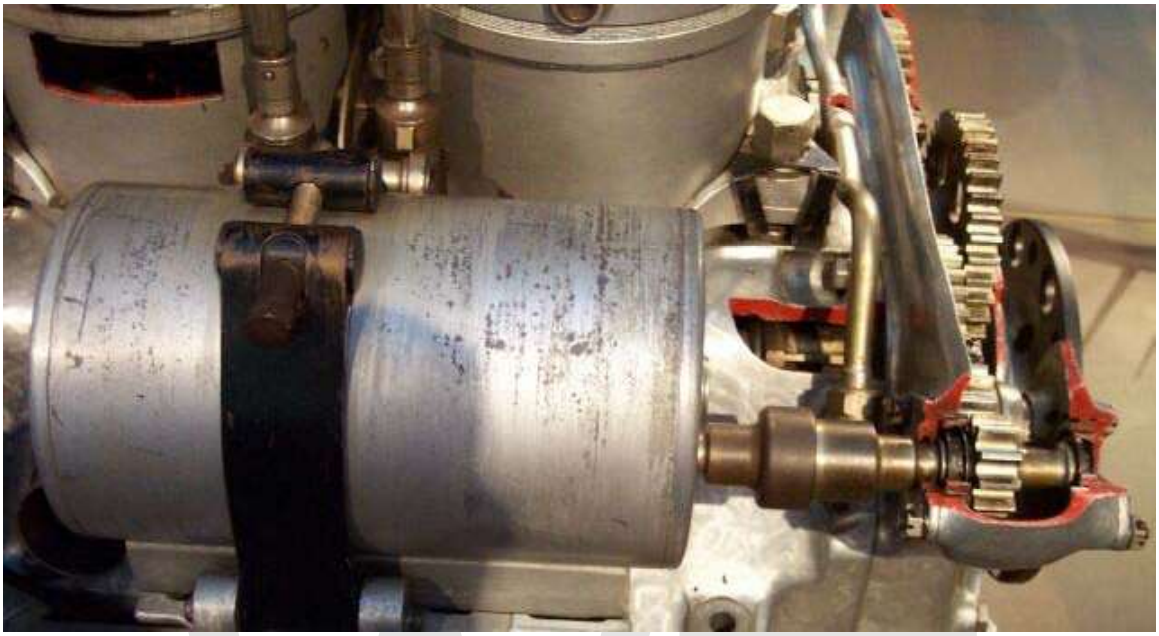
Chapter- 13

Starter Motor

An automobile starter motor

A **starter motor** (also **starting motor**, or **starter**) is an electric motor for rotating an internal-combustion engine so as to initiate the engine's operation under its own power.

History



A 1920s era self-starter

Typical starter installed underneath and toward the rear of an automobile engine

Both Otto cycle and Diesel cycle internal-combustion engines require the pistons to be moving before the ignition phase of the cycle. This means that the engine must be set in motion by an external force before it can power itself.

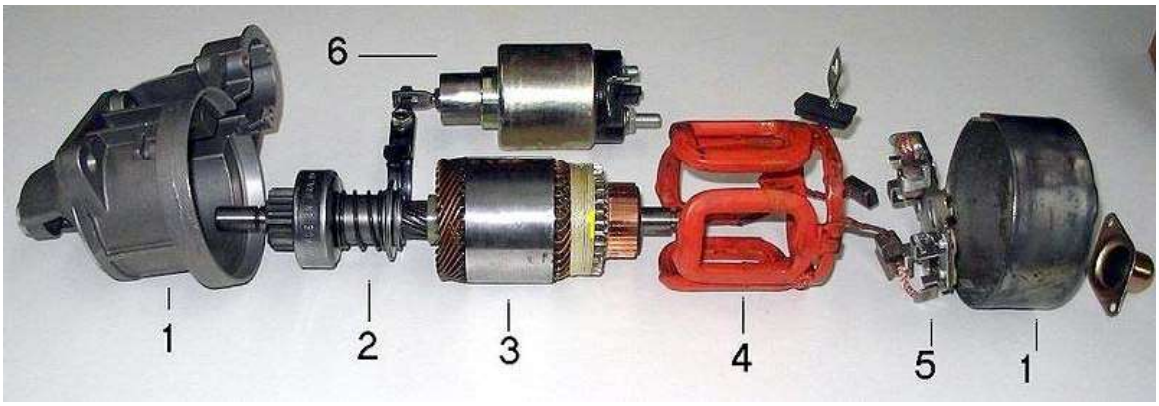
Originally, a hand crank was used to start engines, but it was inconvenient, difficult, and dangerous to crank-start an engine. Even though cranks had an overrun mechanism, when the engine started, the crank could begin to spin along with the crankshaft and potentially strike the person cranking the engine. Additionally, care had to be taken to retard the spark in order to prevent backfiring; with an advanced spark setting, the engine could *kick back* (run in reverse), pulling the crank with it, because the overrun safety mechanism works in one direction only.

Although users were advised to cup their fingers under the crank and pull up, it felt natural for operators to grasp the handle with the fingers on one side, the thumb on the other. Even a simple backfire could result in a broken thumb; it was possible to end up with a broken wrist, or worse. Moreover, increasingly larger engines with higher compression ratios made hand cranking a more physically demanding endeavour.

While the need was fairly obvious—as early as 1899, Clyde J. Coleman applied for U.S. Patent 745,157 for an electric automobile self-starter—inventing one that worked successfully in most conditions did not occur until 1911 when Charles F. Kettering of Dayton Engineering Laboratories Company (DELCO) invented and filed for U.S. Patent 1,150,523 for the first useful electric starter. (Kettering had replaced the hand crank on NCR's cash registers with an electric motor five years earlier.) One aspect of the invention lay in the realization that a relatively small motor, driven with higher voltage and current than would be feasible for continuous operation, could deliver enough power to crank the engine for starting. At the voltage and current levels required, such a motor would burn out in a few minutes of continuous operation, but not during the few seconds needed to start the engine. The starters were first installed by Cadillac on production models in 1912. These starters also worked as generators once the engine was running, a concept that is now being revived in hybrid vehicles. The Model T relied on hand cranks until 1919; by 1920 most manufacturers included self-starters, thus ensuring that anyone, regardless of strength or physical handicap, could easily start a car with an internal combustion engine.

Before Chrysler's 1949 innovation of the key-operated combination ignition-starter switch, the starter was operated by the driver pressing a button mounted on the floor or dashboard.

Electric starter



1. Main Housing (yoke)
2. Overrunning clutch
3. Armature
4. Field coils
5. Brushes
6. Solenoid

The modern starter motor is either a permanent-magnet or a series-parallel wound direct current electric motor with a starter solenoid (similar to a relay) mounted on it. When current from the starting battery is applied to the solenoid, usually through a key-operated switch, the solenoid engages a lever that pushes out the drive pinion on the starter driveshaft and meshes the pinion with the starter ring gear on the flywheel of the engine.

The solenoid also closes high-current contacts for the starter motor, which begins to turn. Once the engine starts, the key-operated switch is opened, a spring in the solenoid assembly pulls the pinion gear away from the ring gear, and the starter motor stops. The starter's pinion is clutched to its driveshaft through an overrunning sprag clutch which permits the pinion to transmit drive in only one direction. In this manner, drive is transmitted through the pinion to the flywheel ring gear, but if the pinion remains engaged (as for example because the operator fails to release the key as soon as the engine starts, or if there is a short and the solenoid remains engaged), the pinion will spin independently of its driveshaft. This prevents the engine driving the starter, for such backdrive would cause the starter to spin so fast as to fly apart. However, this sprag clutch arrangement would preclude the use of the starter as a generator if employed in hybrid scheme mentioned above, unless modifications are made. Also, a standard starter motor is only designed for intermittent use which would preclude its use as a generator; the electrical components are designed only to operate for typically under 30 seconds before overheating (by too-slow dissipation of heat from ohmic losses), to save weight and cost. This is the same reason why most automobile owner's manuals instruct the operator to pause for at least ten seconds after each ten or fifteen seconds of cranking the engine, when trying to start an engine that does not start immediately.

This overrunning-clutch pinion arrangement was phased into use beginning in the early 1960s; before that time, a Bendix drive was used. The Bendix system places the starter drive pinion on a helically-cut driveshaft. When the starter motor begins turning, the inertia of the drive pinion assembly causes it to ride forward on the helix and thus engage with the ring gear. When the engine starts, backdrive from the ring gear causes the drive pinion to exceed the rotative speed of the starter, at which point the drive pinion is forced back down the helical shaft and thus out of mesh with the ring gear.

An intermediate development between the Bendix drive developed in the 1930s and the overrunning-clutch designs introduced in the 1960s was the Bendix Folo-Thru drive. The standard Bendix drive would disengage from the ring gear as soon as the engine fired, even if it did not continue to run. The Folo-Thru drive contains a latching mechanism and a set of flyweights in the body of the drive unit. When the starter motor begins turning and the drive unit is forced forward on the helical shaft by inertia, it is latched into the engaged position. Only once the drive unit is spun at a speed higher than that attained by the starter motor itself (i.e., it is backdriven by the running engine) will the flyweights pull radially outward, releasing the latch and permitting the overdriven drive unit to be spun out of engagement. In this manner, unwanted starter disengagement is avoided before a successful engine start.

Gear-reduction starters

Chrysler Corporation contributed materially to the modern development of the starter motor. In 1962, Chrysler introduced a starter incorporating a geartrain between the motor and the driveshaft. Rolls Royce had introduced a conceptually similar starter in 1946, but Chrysler's was the first volume-production unit. The motor shaft has integrally-cut gear teeth forming a drive gear which mesh with a larger adjacent driven gear to provide a gear reduction ratio of 3.75:1. This permits the use of a higher-speed, lower-current, lighter and more compact motor assembly while increasing cranking torque. Variants of this starter design were used on most vehicles produced by Chrysler Corporation from 1962 through 1987. The Chrysler starter made a unique, readily identifiable sound when cranking the engine.

This starter formed the design basis for the offset gear reduction starters now employed by about half the vehicles on the road, and the conceptual basis for virtually all of them. Many Japanese automakers phased in gear reduction starters in the 1970s and 1980s. Light aircraft engines also made extensive use of this kind of starter, because its light weight offered an advantage.

Those starters not employing offset geartrains like the Chrysler unit generally employ planetary epicyclic geartrains instead. Direct-drive starters are almost entirely obsolete owing to their larger size, heavier weight and higher current requirements. Ford also issued a nonstandard starter, a direct-drive "movable pole shoe" design that provided cost reduction rather than electrical or mechanical benefits. This type of starter eliminated the solenoid, replacing it with a movable pole shoe and a separate starter relay. The Ford starter operated as follows:

1. The operator closed the key-operated starting switch.
2. A small electric current flowed through the starter relay coil, closing the contacts and sending a large current to the starter motor assembly.
3. One of the pole shoes, hinged at the front, linked to the starter drive, and spring-loaded away from its normal operating position, swung into position. This moved a pinion gear to engage the flywheel ring gear, and simultaneously closed a pair of heavy-duty contacts supplying current to the starter motor winding.
4. The starter motor cranked the engine until it started. An overrunning clutch in the pinion gear uncoupled the gear from the ring gear.
5. The operator released the key-operated starting switch, cutting power to the starter motor assembly.
6. A spring retracted the pole shoe, and with it, the pinion gear.

This starter was used on Ford vehicles from 1973 through 1990, when a gear-reduction unit conceptually similar to the Chrysler unit replaced it. Light motor vehicles have now adopted 9.6 volt to 10.4 volt starter motors for use with 12 volt systems to give increased power. The lower current starter will give increased torque, but will tend to overheat and burn out with prolonged use under load.

Pneumatic starter

Some gas turbine engines and Diesel engines, particularly on trucks, use a pneumatic self-starter. The system consists of a geared turbine, an air compressor and a pressure tank. Compressed air released from the tank is used to spin the turbine, and through a set of reduction gears, engages the ring gear on the flywheel, much like an electric starter. The engine, once running, powers the compressor to recharge the tank.

On larger diesel generators found in large shore installations and especially on ships, a pneumatic starting gear is used. The air motor is normally powered by compressed air at pressures of 10–30 bar. The air motor is made up of a center drum about the size of a soup can with four or more slots cut into it to allow for the vanes to be placed radially on the drum to form chambers around the drum. The drum is offset inside a round casing so that the inlet air for starting is admitted at the area where the drum and vanes form a small chamber compared to the others. The compressed air can only expand by rotating the drum which allows the small chamber to become larger and puts another one of the cambers in the air inlet. The air motor spins much too fast to be used directly on the flywheel of the engine, instead a large gearing reduction such as a planetary gear is used to lower the output speed. A Bendix gear is used to engage the flywheel.

On large diesel generators and almost all diesel engines used as the prime mover of ships will use compressed air acting directly on the cylinder head. This is not ideal for smaller diesels as it provides too much cooling on starting. Also the cylinder head needs to have enough space to support an extra valve for the air start system. The air start system operates very similar to a distributor in a car. There is an air distributor that is geared to the camshaft of the diesel engine, on the top of the air distributor is a single lobe similar to what is found on a camshaft. Arranged radially around this lobe are roller tip followers

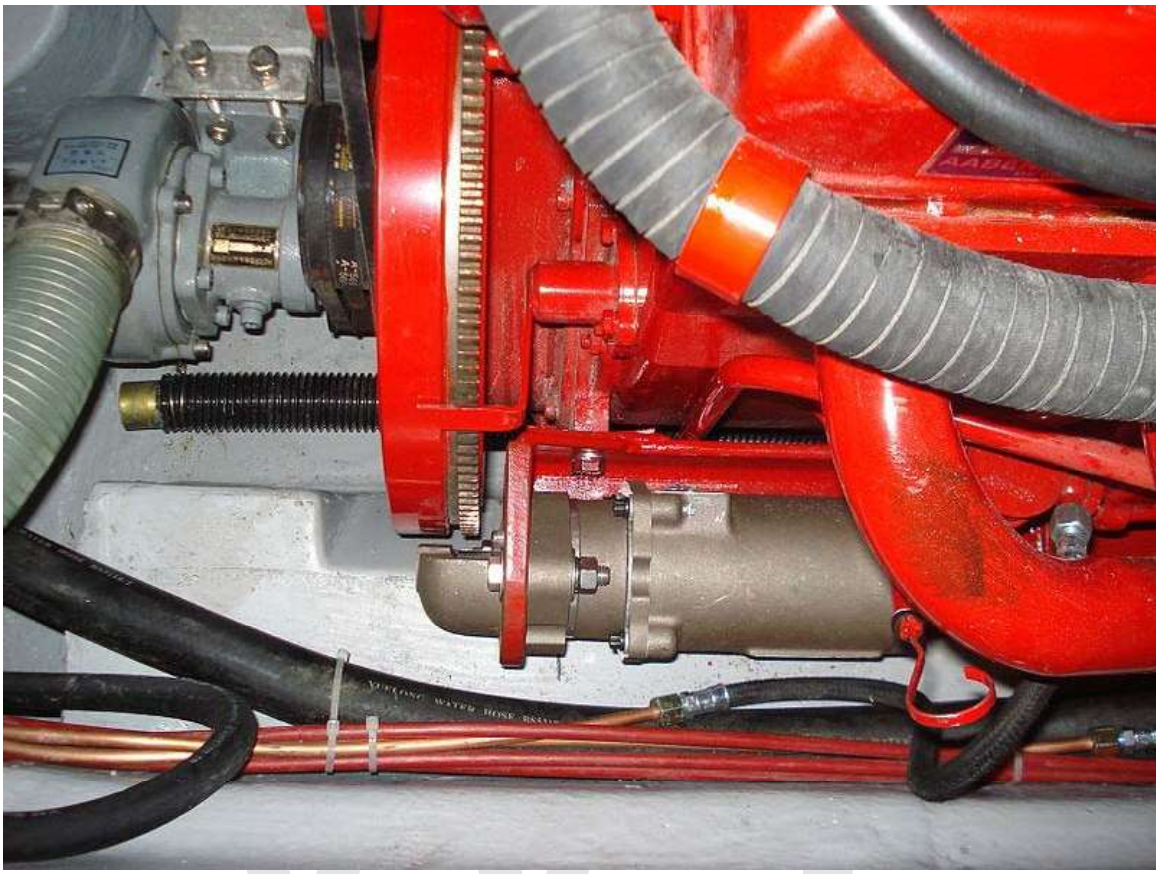
for every cylinder. When the lobe of the air distributor hits one of the followers it will send an air signal that acts upon the back of the air start valve located in the cylinder head causing it to open. The actual compressed air is provided from a large reservoir that feeds into a header located along the engine. As soon as the air start valve is opened the compressed air is admitted and the engine will begin turning. It can be used on 2-cycle and 4-cycle engines and on reversing engines. On large 2-stroke engines less than one revolution of the crankshaft is needed for starting.

Since large trucks typically use air brakes, the system does double duty, supplying compressed air to the brake system. Pneumatic starters have the advantages of delivering high torque, mechanical simplicity and reliability. They eliminate the need for oversized, heavy storage batteries in prime mover electrical systems.

Hydraulic starter

Some diesel engines from 6 to 16 cylinders are started by means of a hydraulic motor. Hydraulic starters and the associated systems provide a sparkless, reliable method of engine starting at a wide temperature range. Typically hydraulic starters are found in applications such as remote generators, lifeboat propulsion engines, offshore fire pumping engines, and hydraulic fracturing rigs. The system used to support the hydraulic starter includes valves, pumps, filters, a reservoir, and piston accumulators. The operator can manually recharge the hydraulic system; this cannot readily be done with air or electric starting systems, so hydraulic starting systems are favored in applications wherein emergency starting is a requirement.

Hydraulic Starter



Hydraulic Starter

Other methods

Other ways of starting have included a spring wound by regenerative braking, gun powder cylinders, windmilling the propeller of an engine while the airplane is flying and putting a car in gear when it is moving.

Manual starting

Early internal combustion engines were generally started with energy supplied by a human operator. The methods included cranking, pushing, flipping the propeller, pulling a cord, foot pedal starters, and indirect methods such as springs and fly wheels. Since these methods are inconvenient and sometimes dangerous, they have gradually been replaced by electric motors and compressed air, starting in the larger engines.

In case of failure of the battery or starter motor, a car with a manual transmission can be started by pushing it or rolling it down hill and then engaging the clutch while it is moving, usually in second gear.

Self starting

Some modern gasoline engines with twelve or more cylinders always have at least one piston at the beginning of its power stroke and are able to start by injecting fuel into that cylinder and igniting it.

Chapter- 14

Speedometer

An Aston Martin speedometer, showing how an eddy-current speedometer indicates the vehicle's speed

A **speedometer** is a gauge that measures and displays the instantaneous speed of a land vehicle.

Now universally fitted to motor vehicles, they started to be available as options in the 1900s, and as standard equipment from about 1910 onwards.

Speedometers for other vehicles have specific names and use other means of sensing speed. For a boat, this is a pit log. For an aircraft, this is an airspeed indicator.

The speedometer was invented by the Croatian Josip Belušić in 1888, and was originally called a velocimeter.

Operation

Eddy current

An eddy-current speedometer gauge on a car, showing the speed of the vehicle in kilometres per hour. Also shown is the tachometer, which displays the rate of rotation of the engine's crankshaft.

The eddy current speedometer has been used for over a century and is still in widespread use. Until the 1980s and the appearance of electronic speedometers it was the only type commonly used.

Originally patented by a German, Otto Schulze on 7 October 1902, it uses a rotating flexible cable usually driven by gearing linked to the tail shaft (output) of the vehicle's transmission. The early Volkswagen Beetle and many motorcycles, however, use a cable driven from a front wheel.

A small permanent magnet affixed to the rotating cable interacts with a small aluminum cup (called a *speedcup*) attached to the shaft of the pointer on the analogue instrument. As the magnet rotates near the cup, the changing magnetic field produces eddy currents in the cup, which themselves produce another magnetic field. The effect is that the magnet "drags" the cup, and thus the speedometer pointer, in the direction of its rotation with no mechanical connection between them.

The pointer shaft is held toward zero by a fine spring. The torque on the cup increases with the speed of rotation of the magnet (which is driven by the car's transmission.) Thus an increase in the speed of the car will twist the cup and speedometer pointer against the spring. When the torque due to the eddy currents in the cup equals that provided by the spring on the pointer shaft, the pointer will remain motionless and pointing to the appropriate number on the speedometer's dial.

The return spring is calibrated such that a given revolution speed of the cable corresponds to a specific speed indication on the speedometer. This calibration must take into account several factors, including ratios of the tailshaft gears that drive the flexible cable, the final drive ratio in the differential, and the diameter of the driven tires.

Electronic

Historic speedometers from the turn of the century

Speedometer currently displaying "Oil Service" necessity

Many modern speedometers are electronic. A rotation sensor, usually mounted on the rear of the transmission, delivers a series of electronic pulses whose frequency corresponds to the rotational speed of the driveshaft. The sensor is typically a toothed metal disk positioned between a coil and a magnetic field sensor. As the disk turns, the teeth pass between the two, each time producing a pulse in the sensor as they affect the strength of the magnetic field it is measuring. Alternatively, some manufactures rely on pulses coming from the ABS wheel sensors.

A computer converts the pulses to a speed and displays this speed on an electronically-controlled, analog-style needle or a digital display. Pulse counts may also be used to increment the odometer.

Another early form of electronic speedometer relies upon the interaction between a precision watch mechanism and a mechanical pulsator driven by the car's wheel or transmission. The watch mechanism endeavors to push the speedometer pointer toward zero, while the vehicle-driven pulsator tries to push it toward infinity. The position of the

speedometer pointer reflects the relative magnitudes of the outputs of the two mechanisms.

Changing a car's tire size can throw off a speedometer's accuracy.

Bicycle Speedometers

Some speedometers for bicycles measure the time between each wheel revolution. The sensor is mounted on the bike at a fixed location, pulsing when the spoke-mounted magnet passes by. These digital devices can be programmed by tire size or by wheel circumference in order to make accurate distance measurements. Others are cable driven as in the automotive speedometers described above.

Error

Most speedometers have tolerances of some 10% plus or minus mainly due variation in tires diameter. Sources of error due to tire diameter variations are wear, temperature, pressure, vehicle load, and nominal tire size.

Excessive speedometer error after manufacture can come from several causes but most commonly is due to nonstandard tire diameter, in which case the

percent error = 100x(1 - "new diameter"/"standard diameter").

Nearly all tires now have their size shown as "T/A_W" on the side of the tire

diameter in inches = TxA/1270 + W.

For example, a standard tire is "185/70R14" with diameter = 185x70/1270 + 14 = 24.20 in. Another is "195/50R15" with 195x50/1270 + 15 = 22.68 in. Replacing the first tire (and wheels) with the second (on 15" wheels), a speedometer reads 100 * (1 - 22.68/24.20) = 6.28% higher than the actual speed. At an actual speed of 60 mph, the speedometer will indicate 60 * 1.0628 = 63.77 mph, approximately.

In the case of wear, a new "185/70R14" tyre of 24.4 inch diameter will have ~8mm tread depth, at legal limit this reduces to 1.6mm, the difference being 12.8mm in diamter or 0.5 inches which is 2% in 24.4 inches.

MPH and KM/H framed somewhat artistically

International agreements

In many countries the legislated error in speedometer readings is ultimately governed by the United Nations Economic Commission for Europe (UNECE) Regulation 39 which covers those aspects of vehicle type approval which relate to speedometers. The main purpose of the UNECE regulations is to facilitate trade in motor vehicles by agreeing uniform type approval standards rather than requiring a vehicle model to undergo different approval processes in each country in which it is to be sold.

European Union member states must also grant type approval to vehicles meeting similar EU standards. The ones covering speedometers are similar to the UNECE regulation in that they specify that:

- The indicated speed must never be less than the actual speed, i.e. it should not be possible to inadvertently speed because of an incorrect speedometer reading.
- The indicated speed must not be more than 110 percent of the true speed plus 4 km/h at specified test speeds. For example, at 80 km/h, the indicated speed must be no more than 92 km/h.

The standards specify both the limits on accuracy and many of the details of how it should be measured during the approvals process, for example that the test measurements should be made (for most vehicles) at 40, 80 and 120 km/h, and at a particular ambient temperature. There are slight differences between the different standards, for example in the minimum accuracy of the equipment measuring the true speed of the vehicle.

The UNECE regulation relaxes the requirements for vehicles mass produced following type approval. At Conformity of Production Audits the upper limit on indicated speed is increased to 110 percent plus 6 km/h for cars, buses, trucks and similar vehicles, and 110 percent plus 8 km/h for two or three wheeled vehicles which have a maximum speed above 50 km/h (or a cylinder capacity, if powered by a heat engine, of more than 50 cc). European Union Directive 2000/7/EC, which relates to two and three wheeled vehicles, provides similar slightly relaxed limits in production.

Australia

There were no design rules in place for speedometers in Australia prior to July 1988. They had to be introduced when speed cameras were first used. This means there are no legally accurate speedometers for these older vehicles. All vehicles manufactured on or after 1 July 2007, and all models of vehicle introduced on or after 1 July 2006, must conform to UNECE Regulation 39.

The speedometers in vehicles manufactured before these dates but after 1 July 1995 (or 1 January 1995 for forward control passenger vehicles and off-road passenger vehicles) must conform to the previous Australian design rule. This specifies that they need only display the speed to an accuracy of +/- 10% at speeds above 40 km/h, and there is no specified accuracy at all for speeds below 40 km/h. All vehicles manufactured in Australia or imported for supply to the Australian market must comply with the Australian Design Rules.

The state and territory governments may set policies for the tolerance of speed over the posted speed limits that may be lower than the 10% in the earlier versions of the Australian Design Rules permitted, such as in Victoria. This has caused some controversy since it would be possible for a driver to be unaware that he is speeding should his vehicle be fitted with an under-reading speedometer.

United Kingdom

A speedometer showing mph and km/h along with an odometer and a separate 'trip' odometer (both showing distance traveled in miles).

The amended Road Vehicles (Construction and Use) Regulations 1986 permits the use of speedometers that meet either the requirements of EC Council Directive 75/443 (as amended by Directive 97/39) or UNECE Regulation 39.

The Motor Vehicles (Approval) Regulations 2001 permits single vehicles to be approved. As with the UNECE regulation and the EC Directives, the speedometer must never show an indicated speed less than the actual speed. However it differs slightly from them in specifying that for all actual speeds between 25 mph and 70 mph (or the vehicles'

maximum speed if it is lower than this), the indicated speed must not exceed 110% of the actual speed, plus 6.25 mph.

For example, if the vehicle is actually travelling at 50 mph, the speedometer must not show more than 61.25 mph or less than 50 mph.

United States

As of 1997, Federal standards in the United States allowed a maximum 5% error on speedometer readings. Aftermarket modifications, such as different tire and wheel sizes or different differential gearing, can cause speedometer inaccuracy.

GPS

GPS devices are positional speedometers, based on how far the receiver has moved since the last measurement. Its speed calculations are not subject to the same sources of error as the vehicle's speedometer (wheel size, transmission/drive ratios). Instead, the GPS's positional accuracy, and therefore the accuracy of its calculated speed, is dependent on the satellite signal quality at the time. Speed calculations will be more accurate at higher speeds, when the ratio of positional error to positional change is lower. The GPS software may also use a moving average calculation to reduce error.

As mentioned in the satnav article, GPS data has been used to overturn a speeding ticket; the GPS logs showed the defendant traveling below the speed limit when they were ticketed. That the data came from a GPS device was likely less important than the fact that it was logged; logs from the vehicle's speedometer could likely have been used instead, had they existed.

* some satnav devices may also use data from the car's systems to improve accuracy