TERRESTRIAL PROPULSION

TERRESTRIAL PROPULSION

Land propulsion is needed for land vehicles (on roads, railways, and off-road), and for the land-phase of other vehicles (e.g. aircraft rolling, take-off, landing, spacecraft rovers...).

Forecast is that by 2015 there will be 1000 million powered land vehicles, and growing, because the saturation observed in the most active countries tend to 0.5 cars per person (i.e. almost 4000 million cars for the 8000 million inhabitants predicted by the UN for 2025).

A propulsion system comprises an energy source, an engine, and an actuator (the thruster); see Fundamentals of propulsion, aside. Although it seems inherent to mobility carrying the energy source on board, vehicles may be dependent on direct solar energy to be propelled (and use some energy accumulator to be powered in the dark), or may take energy from some infrastructure (e.g. electrical contact with aerial lines).

In the design of wheeled vehicles, high traction between wheel and ground (road or rail) is more desirable than low traction, as it allows for higher acceleration (including cornering and braking) without wheel slippage, what is important for speed and safety.

We restrict the study to the propulsion system, leaving aside all the other vehicle engineering aspects, like payload accommodation (fixtures, comfort), vehicle dynamics (performances), electronics (sensors and controls), safety, reliability, environmental effects, aesthetics, maintenance, cost...

TYPE OF VEHICLES AND ENGINES

Land vehicles may be grouped according to their energy autonomy in:

- Autonomous, i.e. vehicles carrying their own energy store (e.g. a fuel tank, or a battery pack), or using unconstrained environmental energy (e.g. solar power). They can start and stop at will without being dependent on infrastructure (at least for a while, because they may need refuelling stations).

- Dependent, i.e. vehicles that get power through cables, usually electric, either by fixed flexible hooks, or by rigid sliding contacts. A model car fed through an umbilical is dependent, but a radio-controlled model, fed from on-board batteries or by solar cells, is autonomous (in this power sense; it may be not autonomous on navigation control).

Land vehicles may be grouped according to the infrastructure required in:

- Routed (or thoroughfare), i.e. vehicles moving along well-prepared routes, which can be subdivided in:
  - Roadways, i.e. wide lanes where the vehicle is not restricted to lateral displacements (within the width of the road), what is advantageously used for vehicles to take over and come across. Road vehicles may even leave their normal route and go astray (off-road).
  - Railways, i.e. narrow supports where vehicles are strictly confined on their route; vehicle cross-over requires multiple railways and switches).
- Rovers (or off-roading), i.e. vehicles able to move through all kind of terrain (more or less); e.g. military vehicles, farming vehicles, cross-country vehicles, mountain cycles, planetary rovers...

There are some vehicles able to move on land and at sea, called amphibian, like hovercraft (using propellers), but only over relatively smooth surfaces.

By far the most used propulsive force on land vehicles is the wheel, an ancient evolution of using round logs as rollers to move heavy loads. The wooden spoke wheel (lighter than solid ones), with rods radiating from the centre (the hub where the axle connects), was known in Central Asia at least 2000 BCE. The even lighter wire-spoked wheel was invented in 1808 by the aeronautical engineer George Cayley, but was only strong enough for bicycles. The pneumatic tyre was invented by the Scottish veterinarian J.B. Dunlop in 1887 to smooth his son's bicycle riding. Railway wheels (and railtracks) were initially on wood, later on iron, and since 1857 on steel.

Friction between wheel and ground has two opposite effects: low friction means low wear (good), but low traction (bad, because it means long braking distance and slow acceleration; the fact that low traction poses less propulsion-power requirements is outweighed by the worse controllability).

Wheels are used in conjunction with axles; either the wheel turns on the axle, or the axle turns in the object body. Bearings are used to help reduce friction at the interface, but even if the friction coefficient μ was the same, the wheel-and-axle mechanism reduces friction in comparison with sliding motion by reducing the radius, i.e. for a chart of weight W to advance a distance L, the work needed by a sliding motion would be $W_s=\mu WL$ (see Friction, below, for μ) whereas the work needed by a rolling motion would be $W_s=\mu WLr/R$, where r is the axle radius and R the wheel radius, plus some additional energy lost at the wheel-to-road interface.

By far the most used propulsion system on land is the reciprocating internal combustion engine (ICE), either in its gasoline version (spark ignition, SI), or in its diesel version (compression ignition, CI), in spite of its many handicaps (trying to be overcome by hybrid propulsion systems and new fuels, see Types of hybrid vehicles, aside):

- Idle running (frequent in city traffic) consumes fuel. Stop-start (or idle-stop) systems have been developed to cope with that, but demanding improved batteries and controls; the IC engine is turned off whenever the car is coasting, braking, or stopped, yet restart quickly and cleanly, saving up to 10 % of fuel in city traffic.
- ICE-efficiency is poor at partial load (and cars are working almost all the time at partial load).
- ICE-engines cannot recover kinetic energy on braking or descent (contrary to electric motors, which are easily reversed). Some ICE developments to storage energy as compressed air, are being investigated.
- Their regime (rpm) is very reduced (they must work within 1000..5000 rpm in general; racing cars and motorcycles may run at >10 000 rpm), requiring a variable speed transmission to cope with the much wider advance speed range at the wheels, although the rotating rate is smaller; e.g. rolling at 130 km/h (v_0=36 m/s), a wheel of D=0.40 m in diameter spins at $v_0/(\pi D)$=36/1.26=29 rps (1720 rpm); and the larger the wheel, the slower. Hence, at top speed a reduction of about 3 from engine-axle speed to wheel-axle speed is needed, but, without a variable transmission, the engine could not drive the wheels at less than 1000/3=300 rpm, i.e. 5 rps that at 1.26 m of wheel circumference means 6.3 m/s, or 23 km/h; and on top of that, at such low speeds the maximum deliverable power would be too low to accelerate or climb. Much larger axe speed reductions (>10) are needed for a car to advance at low speeds with enough power (increasing engine speed if needed); the transmission ratio may be varied manually with a few sets of gears (to have optimum power transmission at several advancing speeds, and some overlapping to easy the exchange), or by means of an automatic transmission system (with a set of gears, or with a continuous range)..
- Present ICEs run on fossil fuels strongly contaminant (and unrenewable). In spite of the electric vehicle's boom since year 2000, it is not expected that ICE can be superseded in the next decades (e.g. compare the 4 kWh/kg of propulsive work for gasoline in a cheap fuel tank, with the expensive Li-ion batteries storing just 0.16 kWh/kg), and the most reasonable way out is the development of less-contaminant biofuels and synthetic fuels (e.g. hydrogen produced with renewable energies may be burned in IC engines while fuel cells take up the fleet). The main technologies to minimise contamination are:
    - Improving engine efficiency, achieved by renovation of the fleet, what reduces fuel consumption for the same task. Proper maintenance of both, engine and vehicle (tyre pressure, clean lube oil, clean air filter...) also contribute to fuel efficiency.
    - Using less-contaminant fuels, starting by adding biofuels to fossil fuels; without engine modification, most SI-ICE may run on gasoline with 10..20 % bioethanol, and CI-ICE may run on diesel with 5..10 % biodiesel. Or changing to gaseous fuels (LPG, CNG, biogas) on city fleets (to cope with lower autonomy), although the future here seems to be the hybrid vehicle.
    - Improving combustion efficiency, achieved by exhaust gas recirculation (EGR), what reduces NO_x emissions (in gasoline and diesel engines), by homogeneous charge compression ignition (HCCI) with lean mixtures, etc.
    - Using exhaust catalysers and particulate filters.

A relative minority (but growing) of land vehicles use electric traction, i.e. electric motors powered either by electrochemical batteries (short range), or by electricity from the infrastructure (e.g. electric trains), or by photovoltaic cells (very low power).

By far the most used land vehicle is the car (automobile), with a fuel tank, an IC engine, and four wheels (with traction in two of them or the four). In automotive design, the automobile layout describes where on the vehicle the engine and drive wheels are found. Most popular modern cars have a transversal front engine with front-wheel drive, for compactness and fuel efficiency. Notice that only the driving wheels can accelerate the vehicle (and only proportional to their normal load (which decreases while accelerating in front-wheel drive cars), whereas all wheels have brakes that contribute to deceleration. By 2014, some self-controlled goal-driven driverless cars are being tested on public roads.

Almost all wheeled vehicles have brakes of some sort, even supermarket trolleys. The most used braking system in land vehicles is the disc brake, where brake pads produce friction with the brake rotors to slow or stop the vehicle. Additional friction is produced between the slowed wheels and the surface of the road. Friction converts the car's kinetic energy into heat. When traveling downhill some vehicles can use their engines to brake, either by increasing internal dissipation, or by regenerative braking in hybrid cars. The braking force must be external to the vehicle, i.e. it is not provided by the brakes mounted on the wheels but from solid friction at the road contact (and aerodynamic resistance, only significant at high speed); essentially, what disc brakes or drum brakes do is forcing the wheel to rotate slower and cause a slippage in its rolling over the road or rail.

PROPULSION NEEDS FOR CARS

Propulsion is always associated with energy expenditure , either invested in increasing the vehicle energy (kinetic or potential), or dissipated to the environment to compensate resistance in in steady motion. In the case of braking (and descending), however, some mechanical energy could be recovered from the vehicle motion and stored within for future use, although this may require additional equipment that cancels the saving (e.g. it can be easy in electrical drives with batteries, but too much complicated in a combustion engine).

The minimum power needed to advance at a steady speed v_0 is $\dot{W}_{\text{min}} = Fv_0$, where F is the traction force (thrust) to compensate for friction forces at the ground contact and in the air. As the traction force F is applied on the ground by the wheels (of radius R), a torque $Q=FR$ must be applied to the wheels, and the propulsion needs may be stated in power units (for $\dot{W}$), or in force units (for F), or in torque units (for Q), but only energy is conservative (force and torque may be varied at will with a lever, but energy remains the same).

The power required to propel a car (a wheeled vehicle in general) of mass m may be expressed as:

$$\dot{W}_{\text{total}} = \dot{W}_{\text{accel}} + \dot{W}_{\text{climb}} + \dot{W}_{\text{roll}} + \dot{W}_{\text{fluid-drag}} \tag{1}$$

with:

$$\dot{W}_{\text{accel}} = mav, \quad \dot{W}_{\text{climb}} = mgv\sin\theta, \quad \dot{W}_{\text{roll}} = \mu mgv\cos\theta, \quad \dot{W}_{\text{fluid-drag}} = c_D A_{\text{f}} \frac{1}{2}\rho v^3 \qquad (2)$$

where $\dot{W}_{\text{accel}}$ is the increases the car's kinetic energy (positive or negative), $a \equiv \dot{v}$ is the linear acceleration, v the actual speed of the car (notice that the reference frame is fixed to the ground), $\dot{W}_{\text{climb}}$ is the increases the car's potential energy (positive or negative), θ the inclination of the path (grade, or slope), $\dot{W}_{\text{roll}}$ is the dissipation on the tyres by the rolling (with a possible sliding contribution) against the pavement (positive, heating the tyre and the road), μ a friction coefficient (in pure rolling motion it is typically 10^{-2} for rubber tyres and 10^{-3} for rail systems, increasing in both cases to some 10^{-1} when sliding for acceleration or braking), and $\dot{W}_{\text{fluid-drag}}$ the power transmitted to the air (positive, later dissipated by turbulence), where c_D is a fluid-drag coefficient defined together with a reference area, A_{f}; for a single car, A_{f}, is chosen as the frontal area (the area projected by the body in the direction of motion), and then c_D is of order 1 for blunt bodies, but may drop to <0.1 for streamlined bodies, being about c_D=0.35 for modern cars; for very long vehicles (trains), the total envelop area in contact with air must be used, and a different drag coefficient used. Both, in aircraft and marine propulsion, the wetted surface area is used instead of just the frontal area. The density of the fluid medium is ρ.

It is usual to rearrange the terms in (1) in the form:

$$\dot{W}_{\text{total}} = \dot{W}_{\text{cruise}} + \dot{W}_{\text{excess}} \qquad (3)$$

where $\dot{W}_{\text{cruise}}$ is the power spent in travelling at constant speed in a horizontal path (i.e. with $a=\theta=0$), and $\dot{W}_{\text{excess}}$ is the excess power available to accelerate and/or climb. In view of legal speed limits in road cars (e.g. 120..130 km/h in motorways), or maximum safe speed in railways (e.g. 350 km/h), maximum engine power is dictated by acceleration performance (or admissible slope at maximum speed), usually stated in the time to reach 100 km/h from rest. The world record on a public road climb slope is 35 % (19º).

Notice that energy in a vehicle is not only needed for propulsion but to power all auxiliary systems, the main consumer being air conditioning (not only on cars, but in airplanes and passenger ships).

Typical values for a small family car (5 seats) are: $\dot{W}_{\text{max}}$=70 kW (50..90 kW), m_{max}=1500 kg (1100 kg tare, from which 150 kg the engine block), able to accelerate from 0 to 100 km/h in 15 s, with frontal area A_{F}=0.9 m^2, c_D=0.35, μ_{roll} =0.015, fuel-tank capacity 45 L, 4 cylinders in-line engine with total displacement of 1.6 L (1.4..1.8 L), direct fuel injection, turbocharged, with a fuel consumption of 4 L / 100 km on road (20 % more in mix traffic), CO_2 emissions 100..150 g/km, and a cost of around 15 000 €. Air conditioning compressor may have $\dot{W}_{\text{max}}$ =5 kW (typically using R134a as working fluid). Specific fuel energy is h_{LHV}=43..44 MJ/kg, with ρ=750 kg/m^3 for gasoline, ρ=830 kg/m^3 for diesel. Larger family cars and executive cars have engines of 100..200 kW with 6 or 8 cylinders in V and 2..4 L displacement. Buses have diesel engines of 200..400 kW with 6 or 8 in-line cylinders and 8..15 L displacement, and the full vehicle may cost around 200 000 €.

Exercise 1. Estimate the time it takes for a 70 kW engine to accelerate a 1500 kg car from 0 to 100 km/h.

Sol.: Considering just the linear acceleration of the car, i.e. neglecting all other power consumption terms (engine and wheels angular acceleration, and all internal and external frictions), the power needed is $\dot{W}_{\text{accel}}=mav$, which at the maximum point, i.e. when v=28 m/s (100 km/h), with an average acceleration $a=v/t$, yields $\dot{W}_{\text{accel}}=mv^2/t$, i.e. $t=mv^2/\dot{W}_{\text{accel}}=1500\cdot28^2/(70\cdot10^3)=17$ s. The average acceleration is $a=v/t=28/17=1.7$ m/s^2, and the corresponding traction force $F=ma=1500\cdot1.7=2500$ N.

Notice that, in a front-wheel drive car, only the front wheels accelerate the car (the rear wheels contribute to the extra loads here neglected, like the internal bearings). If half the weight is supported on the front axle, and the traction is modelled with Coulomb's law, $F=\mu W$ (see Friction, below), then, with axle load $W=mg/2=1500\cdot9.8/2=7350$ N and the above traction force F=2500 N, $F=\mu W$ yields a friction force coefficient of $\mu=F/W=2500/7350=0.34$. Friction coefficients are later presented (Table 1), with a very small value for pure rolling, μ=0.01, and a high value for pure sliding, μ=0.7, so that it can be concluded that the driving tyres are partially sliding during acceleration.

Engine to wheel mechanical transmission

A variable-speed transmission is used to change both speed and torque (its product, transmitted power, is maintained except for the running losses).

Internal combustion engines need a gearbox (really a multispeed transmission, be it with gears, changed by the driver or automatic, or a continuous variable transmission) to match their narrow operating-speed range (e.g. 1000..4000 rpm in a diesel car, i.e. a ×4) to the large driving-speed range of the vehicle (e.g. 1..50 m/s or more; a ×50); e.g. a car engine in first gear (almost no reduction), can easily accelerate the car from 0 to 15 m/s, but would spoil the engine if further.

Manual transmission is most used in ICE-vehicle applications. It uses a driver-operated clutch engaged and disengaged by a foot pedal (automobile) or hand lever (motorcycle), for regulating torque transfer from the engine flywheel to the transmission; and a gear stick operated by foot (motorcycle) or by hand (automobile).

Electric motors of variable speed as used in electric vehicles (brushless DC and all AC motor) do not need a gearbox because even if their torque decreases with speed (e.g. from 400 Nm at low speed to 100 Nm at high speed in a car), there is still plenty of torque at cruising speeds. No start-engine, no clutch, no reverse gear are needed (the motor starts, stops, and turns opposite electrically). Electric drive perfectly matches propulsion needs, where large torque is necessary for acceleration (most needed at small speeds), whereas large power is needed for high-speed cruise. The electric motor gives maximum torque at low speed (at stall if DC, or about 1500 rpm if AC), and power increases with speed. An typical best efficiencies are around 30 % for ICE and 90 % for electric drive.

SOLID CONTACT PATCH

A glance at the tyres of a parked car will show that the region in contact with the road is flattened; the same happens on rail wheels, but it is far more difficult to detect: the contact area in a loaded car tyre is about 10^{-1} m ($10..20\ \mathrm{cm}^2$), and in a loaded rail wheel about 10^{-2} m ($1..2\ \mathrm{cm}^2$). It is through this contact patch that the normal load-force is transmitted to ground for support, and that tangential force is exerted for propulsion.

Exercise 2. Estimate the size of the contact patch in a 1500 kg car with tyres inflated to 250 kPa.

Sol.: Assuming perfectly flexible rubber, i.e. negligible rubber elasticity and thus uniform pressure distribution on the contact patch, with pressure on ground equal to air pressure inside, the total contact area is $A=mg/p_g=1500\cdot 9.8/(250\cdot 10^3)=0.059\ \mathrm{m}^2$ (i.e. about 150 cm^2 at each wheel). Notice that the 250 kPa is understood as gauge pressure.

In practice, contact area is smaller because of tyre rigidity (some 90 % of the vehicle weight is supported by the trapped air and the rest by the rubber stress); for the same tyre load and tyre pressure, wider tyres have larger contact areas because the rubber contributes less.

Contrary to the tyre-wheel-on-road contact, where the road is assumed to remain flat and the tyre deforms, in the steel-wheel-on-rail contact it is the rail that deforms the most; the average pressure on the contact patch on railways may be 400 MPa, against the typical 250 kPa for cars and mountain bikes (racing bicycles may have 600 kPa); for comparison, an adult human standing on planar shoes may have 20 kPa (rising to an average of 60 kPa on bare foot, where the middle foot is arched; with a peak of almost 400 kPa at the forefoot centre (but walking on spike-heel shoes may give 10 MPa, though higher heels shift peak-pressure from the heel to the forefoot). Modern high-speed railways are surface hardened on the rail top to decrease elastic deformation and wear.

Aircraft tyres are specially designed to withstand extremely heavy loads for short durations (landing impact), and are filled with nitrogen or another inert gas to prevent combustion promotion in case of accident. Tyre pressure is about 1.5 MPa in airliners (with a burst pressure about 5.5 MPa), and even higher in business jets.

The first theory to compute the size of the contact patch was developed by Hertz in 1881 using linear elasticity (valid only for hard materials). Result for two parallel cylinders of radii R_1 and R_2 pressed together with a normal force per unit length F_N/L is:

$$a=\sqrt{\frac{4R}{\pi E}\frac{F_N}{L}},\quad p_0=\sqrt{\frac{E}{\pi R}\frac{F_N}{L}},\quad \frac{1}{R}\equiv\frac{1}{R_1}+\frac{1}{R_2},\quad \frac{1}{E}\equiv\frac{1-\nu_1^2}{E_1}+\frac{1-\nu_2^2}{E_2} \tag{4}$$

where a is the half-width of the contact strip, p_0 the maximum pressure (central to the patch), R the composite radius, and E the composite Young's modulus with Poisson's ratio (ν) effects, as defined on (4) . The pressure distribution falls elliptically to zero at the borders of the strip, with a mean value $p_m=(2/3)p_0$.

Typical values are R=0.45 m (wheel radius), F_N=80 kN (axle load 16 t), L=0.02 m (estimation of the real finite length of the contact strip), E=100 GPa (E_{rail}=200 GPa, E_{wheel}=210 GPa, ν=0.3 in both cases), and substitution in (4) yield a=5 mm and p_0=60 MPa (the ultimate normal stress may be more than ten times higher: $\sigma_{ut,rail}$=700 MPa, $\sigma_{ut,wheel}$=900 MPa). This two-dimensional model can be enhanced to include the typical rounding of the rail head; in this case, the contact of a cylindrical wheel (still without considering any rim conicity), with a cylindrical rail of approximately the same radius of curvature, but of perpendicular axes, is an ellipse. Real contact patch shape depends on geometrical details, being elliptic, oval, a lobular, about 10 mm in the rolling direction (in agreement with the strip with of the above 2D-example) by some 20 mm across; the approach beyond un-deformed contact is about 0.1 mm (elastic yield penetration).

The contact patch between a wheel and the ground (road or rail) is the way to vehicle traction and braking. The contact patch is different when the vehicle is in motion (it is displaced forward to the vertical) from when it is static (it is symmetric). The size and shape of the contact patch, as well as the pressure distribution within the patch, are important not only to vehicle propulsion, but to vehicle control (safety) and wheel/road wear (maintenance).

FRICTION

Friction is resistance to relative motion, either measured as a friction force (always opposing velocity), or as a friction energy (work always consumed). Solid friction (also named dry friction) takes place when a solid acts upon a solid system, i.e. by direct contact (either when sliding or in rolling), whereas fluid friction is the internal dissipation in deformable systems (liquid or gas, and real non-rigid solids), including lubricated solid friction (i.e. with a fluid layer in between). Different disciplines study these two different fields:

- Tribology (Gr. *τριβω*, to rub) studies surface friction, wear, and boundary lubrication (i.e. fluid-aided lubrication when the solid surfaces are so close together that there is appreciable contact between opposing asperities).
- Rheology (Gr. *ρηειν*, to flow) studies internal dissipation (viscous, and viscoelastic). Mixed boundary and bulk friction may be considered (Stribeck curve).

Solid friction between two pressing solids is due to the elastic and inelastic deformations of asperities, i.e. surface roughness; the most polished surface to our eyes (a mirror) has asperities thousand times higher than atomic sizes, which are crashed by the normal force between the two solids (we do not consider here adhesion, i.e. friction without compression). Friction is proportional to asperity size.

In spite of the often undesirable connotations of friction in Engineering, friction is a blessing in many circumstances; without friction we could not walk, we could not start or stop our vehicles, and our trousers may come down. On the other hand, friction produces wear, generates heat (by dissipation of friction work), and noise (friction-induced vibrations in solids, which propagates along the solid and is radiated to the environmental fluid). Wear is damage to one or both surfaces in contact with relative motion due to:

- Plastic deformation, i.e. permanent loss of dimensions by asperity compacting, i.e. at loads lower than the yield stress of the bulk material.
- Erosion, i.e. materials removal by asperities being broken and wiped away, causing accumulative loss of dimensions that mark the end life of the solids in contact (wheel and ground (road or rail).
- Fatigue, i.e. the materials weakening when subjected to repeated loading and unloading, by the formation and propagation of microscopic cracks that may eventually reach a critical size and material failure (when micro-cracks join and propagate suddenly).

When considering the contact patch between two solids, and tangential forces being applied (besides the normal pressing against each other), three cases may be distinguished:

- Sticking. Tangential forces may be not strong enough to force relative motion (either sliding or rolling) between the two contacting faces, and the bodies remain solidary. Shear stresses at the interface are bounded to a 'static friction' force, above which relative motion appears (perhaps with some material becoming loose or molten, if the stress was too high).
- Sliding. Tangential forces may cause the two contacting faces to have relative motion at all the contact points. In the simplest case, a small object slides over a larger one that appears almost flat to the other. Sliding usually requires a relative large contact patch to prevent the smaller body to nose down over the other (or a guided motion like in cams). If not properly lubricated with a fluid layer, sliding dissipates a lot of energy that heats the two contacting solids (sometimes, it seems that only the small object being rubbed heats up, but this is by energy accumulation on the same piece of solid).
- Rolling. Tangential forces may cause the two contacting faces to have relative motion except at some point, the instantaneous centre of rotation. In the simplest case, a small object rolls over a larger one that appears almost flat to the other. Rolling usually requires a relative small contact patch to favour the smaller body to nose down over the other, so that the rolling body must be round (or with many small tooth heads, like gears).

Sliding

Sliding (where all contact points move) is not only important to some transport means (sleds, skis, ice skates), but is key to accelerated rolling motion, and to many other mechanisms (pistons, bearings, gears, cams...). The slip velocity is the relative velocity of two points from different bodies in contact, either in pure slide motion, or in mixed sliding-rolling motion.

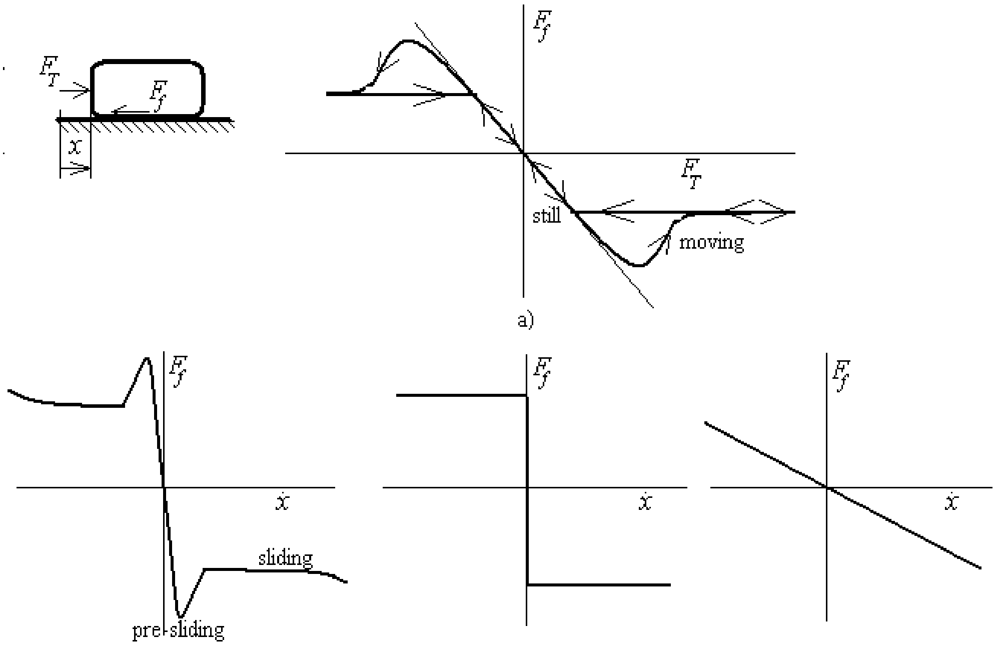
The key concept on solid friction sliding is the friction force, F_f (the opposite of the tangential force imposed F_T) defined as the component of the reaction of the reference solid against the imposed motion of the other solid in contact, located at the centre of the contact surface, and having the same direction of the motion and opposite sense. The origin of this force lies in elastic and plastic deformations of the asperities in contact (if the contact surfaces were atom-size smooth, they would form a single solid (as can be done by rub-welding). The friction force depends on materials properties, compressive force, motion kinematics and interface topology (roughness). Air is always air trapped in between the two contacting solids, except under high vacuum, where friction is much larger; friction between moist surfaces is smaller, and friction between oiled surfaces is much smaller.

In 1699 Amontons proposed a model where the sliding force of a weighting mass is independent of the contact area, $F_f \neq f(A)$, i.e., a parallelepiped block sliding over a surface shows the same friction force whatever the face in contact.

In 1785 Coulomb proposed that the friction force was independent on the sliding speed too, $F_{\mathrm{f}} \neq f(\dot{x})$, and established the most-used model for friction: $F_f=\mu F_N$, where F_f is the tangential force to drag the body, F_N is the normal force that keeps both solids in contact, usually just the contribution of weight, $F_N=W=mg$ on a horizontal surface ($F_N=W\sin\theta$ on a θ-sloping plane), and μ is a friction coefficient depending on the materials and state of their surfaces, which can easily be determined experimentally (but too complicated to model analytically). Typical values are $\mu=0.1..1$ (Table 1), but there are low-friction materials like teflon and diamond-like films. Even this most-simple model has a profound non-linear character because it really means that $F_f = -\mathrm{sign}(\dot{x})\mu F_{\mathrm{N}}$ for $\dot{x} \neq 0$ (sliding regime) and $F_f=-\min(F_T,\mu F_N)$ for $\dot{x}=0$ (pre-sliding regime), where F_T is the applied tangential force. Moreover, it has been found that different values for μ may give better results: a larger μ_{sta} for the static or pre-sliding regime, and a smaller μ_{kin} for the kinematic (or dynamic) regime. Solid friction shows hysteresis, i.e. a response depending on previous state and having some time lag (not recovering instantly).

Notice that the friction force is exerted at the contact patch, so that, if the towing force is not level to ground but at a vertical distance for it (but still horizontal, a force moment appears, which at a steady state is compensated by forward displacement of the normal resisting force (i.e. the solid tends to pitch down). If the dragging force has a normal upward component, the tangential component (friction force) decreases proportionally to the decrease in apparent weight supported at the contact patch.

Contrary to solid friction, fluid friction, at least in the simplest viscous flow, shows a linear resistive force, $F_{\mathrm{f}} = -\mu\dot{x}$, so convenient for analysis, that even solid friction is sometimes modelled that way (it might be good for lubricated joints). Sometimes, for oscillating sliding solids, the law $F_{\mathrm{f}} = -\mu\dot{x}/\omega_0$ is used, where ω_0 is the applied oscillation frequency and $x=A\sin(\omega_0 t)$; it is easy to deduce that this model yields an energy dissipation $E_{mdf}=\pi\mu A^2 n$, n being the number of cycles executed, that is independent on the applied frequency, as the basic Coulomb law. High-Reynolds number flows show a quadratic dependence on speed, $F_{\mathrm{f}} = -\mathrm{sign}(\dot{x})\mu\dot{x}^2$. Figure 1 gives a summary view of what has been explained.



a)

b) Generic behaviour without hysteresis c) Coulomb model d) Linear viscous model

Fig. 1. Friction force F_f opposing an applied tangential force F_T on a sliding object, and some models for the dependence of friction force on relative speed.

Rolling

We refer here to a wheel rolling over a surface, as used in solid traction on ground (the wheels are mounted on axles connected in a frame, undercarriage, on which a load can be supported and transported). Rolling elements are also used in machinery, like in rolling bearings, and in materials processing (rolling is a forming process in which some stock is passed through a pair of rolls to reduce the thickness and make it uniform). As for sliding, we only consider rolling in real cases; in theory, one may consider both sliding and rolling motions without friction, staying in steady state without any force applied, but these ideal-limit models are of no use in propulsion. Some railways were initially built with gears because it was thought that metal-to-metal contact would yield no traction (by lack of friction, but in fact, a 100 t locomotive may allow almost 500 kN of rolling traction force; 50 kN under the worst conditions).

Hence, rolling requires a friction force, otherwise, both, a towed wheel and a spinning wheel, would slip. The force of rolling resistance, F_r (or F_f since it is a friction force) may be modelled with Coulomb law as for slipping, $F_r=\mu F_N=\mu W$, where W is the weight supported by a wheel (the normal pressing force, F_N, in general), and μ (or μ_r) is the dimensionless rolling resistance coefficient or coefficient of rolling friction, with typical values in Table 1. Similarly to sliding, some tangential force must be overcome to initiate rolling motion from rest, but relative uncertainties in rolling are much larger than in sliding, and no data on this static force is given in Table 1 (rolling coefficient refers to the kinematic force to keep the rolling). Moreover, the rolling coefficient shows a clear dependence with the contact area, being almost proportional to it (that is why it is so important to run with appropriate tyre pressure, and the difference between bike-tyre and car-tyre coefficients in Table 1. Notice that rolling resistance may refer to the energy loss in rolling.

A major advantage of rolling versus sliding is that wheels of large radius can easily get across small sudden steps on the road (without butting). Another major advantage of tyres is that they provide great grip on lateral steering, since in this case is the slide friction that acts; the camber force is the force generated perpendicular to the direction of travel of a rolling tyre due to its camber angle and finite contact patch.

Table 1. Typical rolling and sliding friction coefficients, $\mu \equiv F_T/F_N$.

Contact	Rolling resistance coefficient[a)]	Sliding resistance coefficient
Steel wheel on steel rail	0.001[b)]	0.6 (static) 0.5 (kinetic, dry) 0.1 (lubricated[c)])
Byke tyre (thin):		
-on wood track	0.001	0.5 (kinetic)
-on concrete	0.002	0.7 (kinetic)
-on asphalt	0.004	0.8 (kinetic)
Car tyre:		
-on ice	0.1[d)]	0.2 (kinetic)
-on concrete	0.008	0.8 (static), 0.6 (kinetic), 0.3 (wet)
-on asphalt	0.01	1 (static), 0.7 (kinetic), 0.4 (wet)
-on sand	0.3	2 (static), 1.5 (kinetic), 1 (wet)
Rubber on rubber		2
Teflon on teflon, or on steel		0.04

[a)] Only for pure rolling, and accounting just the contact-patch force, not the axle and other transmission forces. Rolling acceleration and deceleration require a combination of rolling and sliding.

[b)] See note a). In railways, for maximum acceleration in dry conditions, a value of μ=0.2 is used for traditional trains, and μ=0.4 for modern trains with traction control. For maximum deceleration, a lower value is used, μ=0.2, to guarantee skid control. Under wet conditions (rain, snow, ice, or dead leaves) it is difficult to get over μ=0.1.

[c)] Lubrication can range from humid rail (μ=0.4 for relative humidity of air RH>60 %), to wet rail by rain (μ=0.2), to oily rail by dead leafs in autumn (μ=0.1). To mitigate the slippery rail problem, a paste of sand and metallic particles is delivered before the contact patch.

[d)] It seems that rolling resistance is much higher on ice than on asphalt because of the local ice crushing, but depends a lot on the state of the surface.

Acceleration and deceleration of a rolling wheel always requires some sliding. For rolling with sliding, the distance covered by the axle (translation), $v_0 t$, is not equal to the path length of a rim-point on the rotating wheel, $\omega R t$, and the relative difference is named wheel slip or creepage, ξ)

$$\xi \equiv \frac{\omega R - v_0}{\frac{1}{2}\left(\omega R + v_0\right)} \tag{5}$$

Two broad cases may appear, besides the particular states of ξ=0 (pure rolling), ξ=1 (spinning without rolling), and ξ=–1 (pure sliding):

- Under-rolling, $\omega R < v_0$ (i.e. slip $\xi < 0$), i.e. when the axle translates faster than in the pure-rolling case. This happens when a rolling vehicle brakes, or when a non-rolling wheel touches a moving surface (e.g. on landing). Braking is a main safety issue.

- Over-rolling, $\omega R>v_0$ (i.e. slip $\xi>0$), i.e. when the axle translates slower than in the pure-rolling case. This happens when accelerating a standing vehicle, or when a spinning wheel touches a still surface. Acceleration is key to fast travel.

The contact patch on a standing wheel is symmetric relative to the advance direction, but as soon as the wheel moves, the shape of the contact patch modifies and takes a forward displacement, with a clear pressure distribution yielding a vertical resultant ahead of the wheel axle (where the load acts), creating a force moment that balance the torque applied to the wheel by the propulsion system.

In rolling, sliding friction determines the maximum force (torque) that can be transmitted through a wheel, either to start rolling and accelerate, or to brake and stop (and to help bending the trajectory). In fact, there is always some sliding involved in rolling, since wheel particles that enter the contact area (from the right, in Fig. 2a) adhere to opposing particles of the rail but the creepage increasingly tries to shear the two surfaces. The traction curve in rolling (Fig. 2b), is a plot of the friction coefficient (tangential force divided by normal force, F_T/F_N) versus the slip ξ (5); in pure rolling (no slipping) the tangential force is $F_T=\mu F_N$, with $\mu=0.01$ for cars and $\mu=0.001$ for trains (Table 1), although in a standing still wheel it may take any value between $\pm\mu F_N$ before rolling starts. But, on acceleration or deceleration, some slipping occurs, and friction F_T quickly and greatly grows, reaching almost the pure-sliding value, $\mu=0.1..0.6$, according to surface conditions (Table 1), for small creepings (the maximum may be at $\xi=0.01..0.05$); increasing the slip does not increase the friction but reduces it due to overheating at the contact patch (Fig. 2b).

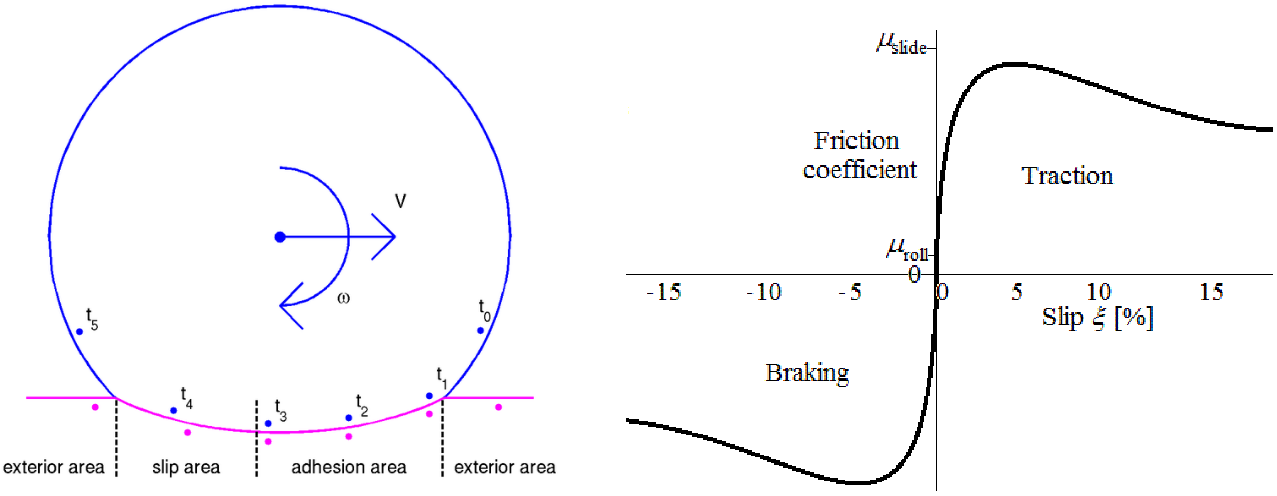


Fig. 2. a) Relative position of two particles, coincident at the entry contact point, during rolling. b) Traction curve.

The standard test to measure rolling resistance and the traction curve of new tyres is to use a large (1.5 m in diameter) steel drum as ground; first the drum is set to spin to a pre-set peripheral speed (90 km/h), and the torque (or the power) demanded measured; second, the tyre, mounted on a movable axle parallel to the drum, is approached until light-contact makes it spin with the same peripheral speed (the rolling resistance, measured in any of the two axis, should be negligible); third, the normal force between axles is being step-by-step increased, and the rolling resistance measured (a simple measure is by the force needed to maintain the wheel axle without translation around the drum).

It is important to avoid accelerations and decelerations with large creeps in wheel propulsion (either on railways or roadways), not only because friction decreases (Fig. 2b), but due to the increase in wear by the overheating. Special care must be taken to avoid the two extreme cases of slip:

- Blockage, or skidding (freezing of wheel rolling), i.e. pure sliding motion ($\xi=-1$). Must be avoided because all the wear concentrates on the same area of the wheel, and the abrasion tends to make it planar (flatted or squared wheel). The anti-lock braking system (ABS) in cars, and the wheel slide protection (WSP) systems in railways, are safety devices to this purpose; notice that the skidding in roadways would be without lateral control, what aggravates the problem.
- Spinning, or excessive wheel rolling, i.e. pure rotation without axle translation ($\xi=1$). Must be avoided because all the wear concentrates on the same area of the rail (or road), with the consequent grinding effect.

Tyres

A tyre (or tire) is a hollow rubber cushion at the periphery of a wheel, providing:

- Traction, i.e. tangential force for acceleration, deceleration (braking), and steering.
- Support, i.e. normal force to balance the vehicle weight, and dynamic loads.
- Cushion, i.e. shock absorption for a smooth ride over surface irregularities, usually provided by a compressed-air chamber (pneumatic tyre).

Tyres are used in practically all land vehicles (and aerospace landing gears), except in railways (some trams have both, steel wheels, and tyre wheels, on separate tracks, to decrease noise level in urban areas), and when track chains are used in heavy industrial and military vehicles (to apply lower pressure and avoid sinking on softer terrain).

A tyre is composed of several parts (Fig. 3); from outside in:

- Tread, what comes in contact with ground, with grooves to channel away water in wet roads; must be at least 1.5 mm deep.
- Beads, the wire-reinforced parts in contact with the rim on the central wheel.
- Sidewalls, bridging the tread and the bead, made of rubber reinforced with fabric or steel cords to increase tensile strength to transmit shear from axle to the contact patch.
- Ply, layers of flexible but relatively inextensible cords embedded in the rubber to minimise its stretching.
- Compressed air, either directly in contact with the ply and wheel in modern cars, or hold inside a separate tubular chamber (in cycles and heavy vehicles); some 90 % of the vehicle weight is supported by the trapped air and the rest by the rubber. A valve stem, mounted directly to the wheel rim in the case of tubeless tyres, or as an integral part of the inner tube, allows tyre inflation. Tyre-pressure monitoring systems are being incorporated to the driver's instrument panel, to signal low pressure conditions, to minimise rolling resistance (thus increasing overall fuel efficiency), and to avoid accidents due to tread disintegration (under-inflated tyres are the first cause of tyre failure, followed by sharp braking and curb hitting). Very small tyres are not

pneumatic but made of solid rubber, like those in carts, office furniture, lawnmowers, wheelbarrow...).

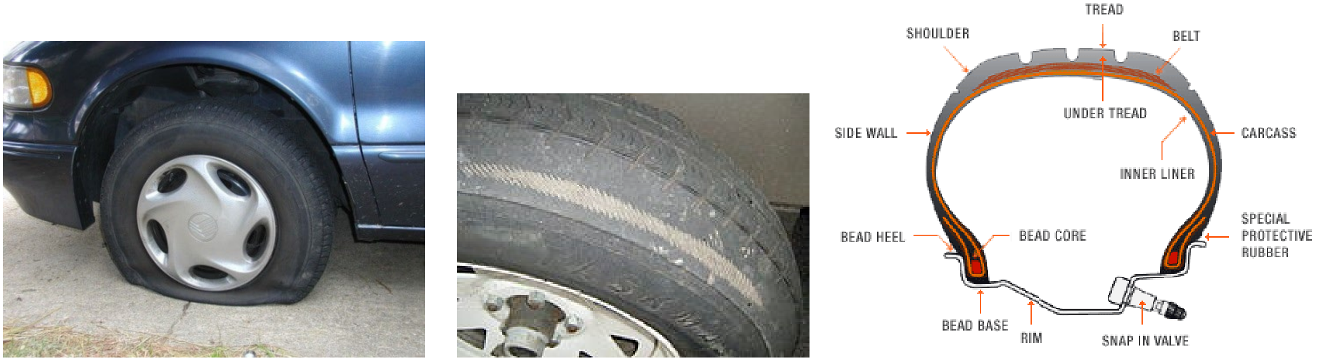


Fig. 3. a) Flat tyre (Wiki). b) Sidewall wear down to the ply (Wiki). c) Tyre structure.

The primary cause of tyre rolling resistance is rubber hysteresis (the energy of deformation is greater than the energy of recovery, heating the rubber); some 70 % of the power dissipation occurs at the contact patch, and the rest at the sidewalls. Wider tyres increase dissipation area and thus remain cooler (on hard braking, the rubber melts and lubricates de contact). Large-diameter wheels have less rolling resistance than smaller ones because there is less rubber deformation frequency (the spinning rate reduces). Notice that tyre rotation usually refers to the moving tyres to different car positions (e.g. front to rear), not to wheel rotation rate.

Exercise 3. Estimate the force needed to push a still car of 1500 kg over a horizontal asphalt road.

Sol.: The force to push a stalled car (without gears engaged, of course) depends on how well lubricated the wheel's bearings and other transmission elements are, because this is usually the highest friction resistance, and not the pure rolling of the tyre. An upper bound might be the case of pure sliding (non-rotating wheels); with data from Table 1 for a car-tyre sliding on asphalt, $F_f=\mu W=\mu mg=1\cdot 1500\cdot 9.8=14.7$ kN, well beyond the capability of a dozen people (that would be the case of trying to push a braked car). A lower bound would be to just account for the rolling friction force at the wheel-ground contact; with data from Table 1 for a car tyre rolling on asphalt, $F_r=\mu W=\mu mg=0.01\cdot 1500\cdot 9.8=147$ N to keep it rolling, well under the capability of an adult person. In practice, the pushing force for a still car to start rolling may be more than 500 N but usually below 1000 N (it increases with hard pushing, so start slowly!), i.e. suitable for two adults. Once the rolling started, the force required may drop to 200..300 N (enough for one pusher for a short while, since at normal pace of 1 m/s, pushing with 300 N delivers $Fv=300\cdot 1=300$ W, a hard personal work).

Exercise 4. A seemingly strange case of propulsion is the forward throwing off-ground of a hoop (in a vertical plane) while imposing some backward spinning on it; after touching down, the hoop rolls toward the thrower. Analyse the motion of a hoop (a ring of mass m and radius R), initially spinning with angular velocity ω_0 in the air, when touching a floor with negligible linear speed (i.e. assuming its centre of mass initially at rest).

Sol.: This propulsion system may be viewed as a kinetic energy source, with a solid-friction propeller (the 'engine' is the solid rim material that transforms normal stress to shear stresses). For the planar motion of this rigid body with a mixed rolling-sliding motion, the linear and angular momentum equations, neglecting the rolling friction against the sliding friction (see Table 1 and Fig. 2b) are:

$$\left.\begin{aligned} v &= v_0 + at = v_0 + \frac{F_\mathrm{f}}{m}t = v_0 + \frac{\mu W}{m}t = v_0 + \mu g t \\ \omega &= \omega_0 - \alpha t = \omega_0 - \frac{F_\mathrm{f} R}{I}t = \omega_0 - \frac{\mu W R}{mR^2}t = \omega_0 - \frac{\mu g}{R}t \end{aligned}\right\} \quad v = \omega R \quad \Rightarrow \quad t_\mathrm{tr} = \frac{\omega_0 R - v_0}{2\mu g}$$

i.e. the wheel accelerates linearly across the ground due to the slipping friction force $F_\mathrm{f}=\mu W$, with a linear acceleration $a=\mu_\mathrm{slip} g$ while the same friction force creates a moment relative to the axis, $Q=F_\mathrm{f}R$, which causes an angular deceleration, $\alpha=Q/I$ (I is the moment of inertia, $I=mR^2$ for a ring). After the transient period t_tr found above (with $v_0=0$ in this exercise), the hoop attains a pure rolling motion with a linear deceleration $a=\mu_\mathrm{roll} g$ (and $\omega=v/R$) until full stop (usually preceded by falling to one side at low speeds after dynamic stability is not strong enough).

If, instead of coming into ground contact a ring spinning at ω_0 in the air with $v_0=0$, a non-spinning ring ($\omega_0=0$) with linear speed v_0 parallel to ground comes into ground contact, then the same equations above would apply, except for a reverse of sign in the accelerations, with the same result: i.e. now $t_\mathrm{tr}=v_\mathrm{o}/(2\mu g)$.

Exercise 5. Estimate the stopping distance (braking length) and the forces on the wheels of a 1500 kg car that suddenly brakes while running at 15 m/s (54 km/h) over a horizontal asphalt road. Assume that it follows a straight trajectory with the wheels blocked, with the centre of mass fixed at 0.7 m off ground and 10 % to the rear of the middle point between wheel axles, which separated 2.5 m, because of the rear load.

Sol.: It is a dynamic slipping case with $\mu=0.7$ (Table 1). We take the two front wheels together, and similarly for the two rear wheels (and label them 1 and 2, respectively instead of f and r). The relation between normal and tangential forces at each axle is $F_\mathrm{T}=\mu F_\mathrm{N}$. Before the braking, force and torque balance are $F_\mathrm{N1}+F_\mathrm{N2}=mg$, $F_\mathrm{N1}x_1=F_\mathrm{N2}x_2$, with $m=1500$ kg, $g=9.8$ m/s^2, $x_1=1.5$ m and $x_2=1$ m ($x_1+x_2=2.5$ m). Hence, the initial load share is (solving) $F_\mathrm{N1}=5900$ N and $F_\mathrm{N2}=8800$ N. Now the brakes are applied and the global deceleration force is $F_\mathrm{d}=\mu mg$, i.e. a constant acceleration $a=-\mu g$, so that the speed progresses as $v=v_0+at$, and the space as $s=v_0t+½at^2$, what yields $a=0.7\cdot9.8=6.9$ m/s^2 (a forward g-force of $0.7g$), time to stop $t=2.2$ s, and distance to stop $s=16.4$ m.

While on braking, the force and torque balance are $F_\mathrm{N1}+F_\mathrm{N2}=mg$, $F_\mathrm{N1}x_1=F_\mathrm{N2}x_2+F_\mathrm{d}z$, with $F_\mathrm{d}=\mu mg=0.7\cdot1500\cdot9.8=10.3$ kN and $z=0.7$ m (arm of the deceleration force relative to the centre of mass). Solving these two equations yield $F_\mathrm{N1}=8800$ N and $F_\mathrm{N2}=5900$ N (notice that the load share has reversed; now, the front axles is more loaded than the rear one). The tangential forces at each wheel are just $F_\mathrm{T}=\mu F_\mathrm{N}$ ($F_\mathrm{T1}=6.1$ kN and $F_\mathrm{T1}=4.2$ kN; total, $F_\mathrm{d}=10.3$ kN).

Rail systems

Using rails to guide a vehicle motion (and in machinery, in general) offers great advantage on trajectory control and interaction (e.g. rough ground can be greatly smoothed to ease rolling, electric contact to overhead lines is easier for trams than for trolleybuses). Most rail systems are based on trains, i.e. a set of in-line connected coaches or wagons. Railways were initially on wood, later on iron, and since 1857 on steel (concrete is used in some monorails). Steel wheels are always used over steel rails, but some rail systems use rubber wheels (e.g. monorails).

Similar to road vehicles, the propulsion system in trains is not only needed to overcome the rolling resistance, but mainly for acceleration, climb, aerodynamic and wind drag, and internal resistances in the transmission (including wheel bearings).

The energy source used in most modern railways is electric because of its many advantages (in spite of higher infrastructure costs): no on-board storage needed, more powerful acceleration, energy recovery in regenerative braking, cleanness... Electricity is taken from the mid-voltage AC three-phase grid, and fed as DC or AC to a single contact cable, closing the circuit through the rail track, which is wired back into the substation supplying the power. The contact cable is usually an overhead line (to allow for high voltage) supported on a catenary, but it is an overhead rail within tunnels, and a third rail on the ground in some cases. In fact, the catenary shape is for the steel cable (hanging from side posts with a brace to reach the central position), from which the contact copper cable, some 15 mm in diameter, hangs in horizontal position, held under stress tension by weights suspended at each end of its length, about 1.5 km; each length is overlapped by its neighbour to ensure a smooth passage for the pantograph, which is spring-loaded and pushes a contact shoe up against the contact wire. The contact cable is laid out with some small horizontal zigzag to minimise concentrated heating and wear in the pantograph. The pantograph shoe has an insert of a carbon paste strip which actually contacts the wire; this graphite paste provides good electrical contact, lubrication, friction-heat spreading (the contact point is moving), and easier maintenance.

According to the track (the permanent stuff), two different basic configurations exist:

- Monorail systems, where vehicles run either suspended from the rail (more stable), or straddle on the rail (usually larger than in the former case, but in any case narrower than the vehicle). Modern monorails are short haul, straddling on a large elevated steel or reinforced concrete beam (0.6..0.9 m wide), over busy populated areas (e.g. airport trains). A rubber-tired carriage contacts the beam on the top and both sides for traction and to stabilize the vehicle; electric contacts slip along the guidance beam.
- Two-rail systems (by far the most common, Fig. 4), where wheels sit on the rails without guidance except for the shape of the wheel-tread in relation to the rail-head. There are some tracks with three rails, but only two are used for wheel propulsion; the third rail is usually of different type and used for electrical connection or, in the case of three equal rails (on short tracks for two rail gauges), only two are used at a time (not a good long-term solution because of asymmetric wear and complicated switching).

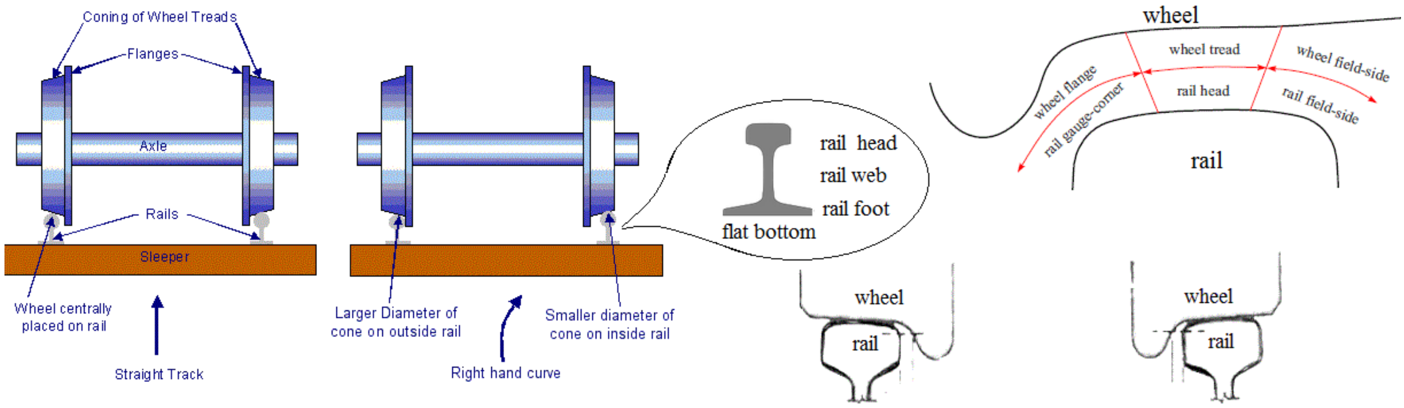


Fig. 4. Lateral stability for wheels running on rails, and details of possible wheel-rail contact zones.

Nearly all railway systems use wheels fixed to a common axle, the wheelset (the wheels on both sides rotate in unison), and two wheelsets are mounted in a bogie or truck within a rigid frame (sometimes allowing some axle steering to better accommodate bends); a coach has two bogies. Lateral stability on displacement (drift) and angle (yaw) is provided by the coning of the wheel-tread (e.g. 3° on traditional tracks; half of it in high-speed tracks), and the rounding of the rail head, in such a way that a lateral displacement creates a restoring force, as can be seen in Fig. 4 for the case of a bend to the right. Wheel flanges should stay about 8 mm from the rail inwards, without touching the rails except as a last aid to prevent derailment (producing a load squealing); the ±8 mm freedom in lateral displacement for the wheelset, the conicity of the wheel treads, and a matching tilt of the rails (sloped inwards by the same angle, e.g. 3°) gives lateral stability on a straight track (i.e. running centrally), and on curve turnings. Rail tracks for trams are not tilted but a rail guard is provided.

Rail track gauge is measured between inner lateral head faces (international standard gauge is 1435±1 mm). Most modern railways use continuous welded rail (CWR), several km long, built by flash-butt welding of long rail pieces (up to 250 m long, since joints are a source of weakness). Welding is performed by an automated track-laying machine, running a strong electrical current through the touching ends (thermite welding is used in repair), usually when ambient temperature is midway between the extremes experienced at that location. CWR reduce track vibrations and misalignment. Tracks are designed for a given maximum axle load (e.g. 20 t per axle on freight tracks, 6 t per axle on trams). Rail head is 80 mm wide (against 150 mm for the wheel), but only 15..20 mm of width is in contact at a time (the contact patch, about 10 mm wide in the direction of motion), and not always at the top of the rail but on a 40 mm wide band (and occasionally more inward).

According to the vehicle use (the rolling stuff), two different types can be distinguished depending on the type of load transported:

- Passenger trains, which may be long-haul trains (e.g. >500 km), high-speed trains (e.g. >250 km/h, 70 m/s, in some segments), trams, metros (underground)...
- Freight or cargo trains, which may be containers, bulk carriers, cisterns...

According to the number of wagons in the vehicle, the rolling stuff may be:

- Single vehicle, only used on small short-haul passenger trains and railway machinery.

- Trains, consisting of a series of linked vehicles used to transport cargo or passengers.

On trains, motive power is provided by a separate locomotive or by self-propelled multiple units, in most cases by electric energy received via overhead lines or through a third rail electric system, with diesel engines used in low-traffic or remote railways where the high investment of electrical infrastructure is unjustifiable. The use of self-propelled coaches is increasingly common for passenger trains, but rare for freight trains, which remain powered by locomotives. Traditionally, locomotives pull trains from the front, or in push-pull mode where one locomotive pulls the train from the front and another locomotive pushes it from behind. Multiple units are more energy efficient than locomotive-hauled trains because the whole weight of the train is placed on driving wheels, providing better traction. Climb slopes are kept small (<5 %) to have a meaningful speed with a given maximum power (the world record is 13.5 %, i.e. 8º, in Lisbon tram).

Train brakes are used to decelerate, to keep position on sloped tracks, and to avoid acceleration on slope down (a 0.5º track inclination, <1 % slope, may be sufficient for a non-braked train to slowly start rolling down). Braking in electric trains can be regenerative (and in non-electric systems too, using compressed air, flywheels, or even gravitational storage (the first underground train, the London Tube, was designed with stations at a higher level than intermediate tracks, with small slopes on either side, so that the trains enhanced the acceleration by sloping down the station, and enhanced the deceleration by the climb to the next station).

High-speed railways are gaining popularity for passenger transport since the successful 'bullet train' of Japan, started for the Tokyo Olympic Games of 1964. High speed is sometimes stated as >200 km/h in at least one section (as for the first Japan units), or recently as more than 250 km/h (70 m/s), and they usually require not only new vehicles, but new tracks and new aerial lines. Speed records are near 500 km/h on tests, but trains in service do not go faster than 350 km/h, requiring curve radius above 5 km (8 km for 350 km/h), and maximum slopes of 5 %. The lines may rest on traditional sleeper and ballast or on concrete slabs (14 m wide for a double track), although the latter is much more expensive to build because of subsidence of the embankment. Fences prevent access of animals to the tracks. Trains are powered via overhead cables and pantographs, usually at 25 kV alternate single-phase current. A typical set of eight coaches has a total motive power of 8..10 MW, either in one driving unit or better distributed in 2, 4, or 8 combined units (usually asynchronous three-phase motors fed through an electronic drive). Total price of the train is about 30 M€. In spite of the streamlined design of high-speed trains (at least the first 5 m in the leading coach, smooth inter-car connections, and the skirts used to smooth the structures underneath), above 200 km/h, more than 75 % of propulsion power is used to overcome air drag.

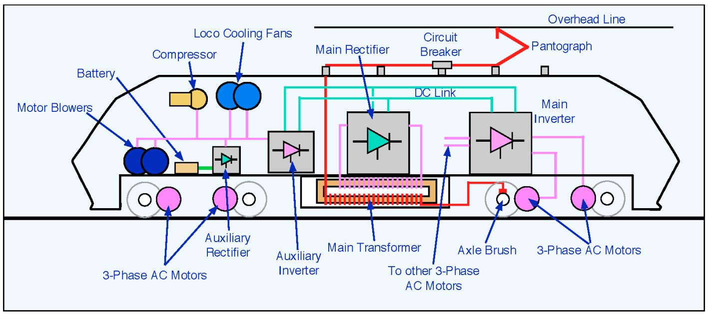


Fig. 5. a) High speed train (AVE Class 103). b) Diagram of a traditional electrical locomotive.

Aerodynamic drag on trains includes friction drag and pressure drag, which are respectively the algebraic sums of the longitudinal components of all shear and normal forces on the train surfaces (including the pantograph). Besides longitudinal drag, other relevant aerodynamic effects on trains (particularly on high-speed trains) are: crosswind stability, noise generation, upraise of ballast due to air flow, pressure waves in tunnel entrance, sonic booms at tunnel exit, etc.

As for any other type of vehicle, power is not only needed for propulsion but for other uses on board, among which air conditioning (AC) is the most demanding (AC must keep a comfort temperature of about 22±2 ºC with ambient temperatures ranging from −20 ºC to +50 ºC, in coaches with large windows).

Propulsion needs for trains

As for tyre-road systems, railway systems need propulsion to overcome, at cruise speed, friction at the wheel's contact patch, and air-drag friction (we neglect pantograph friction); more propulsion is needed to accelerate and climb. The analysis in terms of power ($\dot{W}$) and terms of traction force (F_T) is (1):

$$\begin{aligned}\dot{W}_{\text{total}} &= mav + mgv\sin\theta + \mu mgv\cos\theta + c_D A_f \frac{1}{2}\rho v^3 \\ F_{\text{T,total}} &= ma + mg\sin\theta + \mu mg\cos\theta + c_D A_f \frac{1}{2}\rho v^2\end{aligned} \qquad (6)$$

where m is the mass of the train, a its acceleration, v the advancing speed, θ the track slope angle, μ the rolling-sliding coefficient (see Table 1, above), $c_D A_f$ the product of a drag-coefficient times its reference area, and ρ the air density. In trains, skin friction drag is more important than shape pressure drag (the streamlined nose and tail are mainly for smoothing transients in tunnel entrance and exit (tunnel mouths should have a bell shape for the same reason), so that the 'wet' area is chosen for A_f, and a drag coefficient of c_D=0.01 for streamlined trains (similar to aircraft).

The case of a high-speed train

Let have some numbers on a typical high-speed train with the following data: a 450 passengers train, of 450 tonnes, with a power of 9 MW, composed of eight coaches, each 25 m long 4 m high and 3 m width, with two bogies of two axles each (on this case, half of the wheels are powered with asynchronous triphase motors of 550 kW each), accelerating from 0 to 320 km/h in 380 s, and decelerating from 320 km/h to 0 in 3900 m.

The maximum attainable speed, neglecting friction loses in the transmission, can be deduced from (6) for a=0 (steady speed), θ=0 (level run), μ=0.001 (pure rolling, Table 1), c_D=0.01, A_f=(4×25)×(2·4+3)=2200 m^2, and ρ=1.2 kg/m^3; direct substitution yields $9\cdot10^6=0+0+4500\cdot v+13\cdot v^3$, and thus v=87 m/s (310 km/h); the stated maximum speed is 350 km/h, so that this approximate calculation is valid in view of the assumed uncertainties. Notice that at v=87 m/s, 5 % of the power goes to rolling friction and 95 % to air drag (9=0.4+8.6 MW).

The maximum climb slope at a reasonable high speed, say 200 km/h (v=56 m/s) can be estimated with (6) for a=0 (steady speed), and negligible rolling resistance; direct substitution yields $9\cdot10^6=0+247\cdot10^6\cdot\sin\theta+0+2.3\cdot10^6$, obtaining θ=0.027, i.e. just a slope of 2.7 % (1.6º).

Maximum acceleration can be estimated with (6) if an average speed is assumed (say 100 km/h) and all other terms neglected, with the result that this train can accelerate from 0 to 28 m/s at a rate of $a=\dot{W}/(mv)$=$9\cdot10^6/(0.45\cdot10^6\cdot28)$=0.7 m/s^2, in a time about $t=v/a$=28/0.7=40 s, covering a space of $s=\frac{1}{2}at^2$=570 m. To check the quoted value of "accelerating from 0 to 320 km/h in 380 s", a more detailed model, including air drag, must be used, but an upper bound may be easily checked; looking for the minimum power required to impose the linear kinetic energy (rotational energy is neglected), $E_k=\frac{1}{2}mv^2=450\cdot10^3\cdot89^2$=1800 MJ, in the stated 380 s, we need $\dot{W}_{\min}=E_k/t_{acc}$=1800/380=4.7 MW, therefore, the 8.8 MW seems enough for this task (the rest would be the power dissipated by all friction processes).

But an important point when dealing with acceleration and deceleration in wheel systems is that the contact patch has a limitation in traction and braking force. In the case of modern railways, maximum traction is obtained for μ=0.4 (see foot-note in Table 1), so that from the second equation in (6), maximum traction is bound to $F_T=\mu F_N=\mu m_T g$, where m_T is the mass supported by the traction wheels; in our case of distributed traction on half the axles, m_T=450/2=225 t, so that maximum traction is $F_T=\mu m_T g$=0.4·225·9.8=880 kN; in the case of similar trains but with two locomotives (two traction units from the eight coaches), m_T=450·2/8=112 t and F_T=440 kN. Essentially, at low speeds a control system must limit the applied power such that, with $\dot{W}=Fv$, maximum traction force is not exceeded (to avoid wheel spinning over the same rail spot), whereas at high speed traction force is limited by available power.

Braking brings two basic limitations:

- How to get the force to decelerate the vehicle. This force must be external to the vehicle, i.e. it is not provided by the brakes mounted on the wheels; it must come from solid friction at the contact patch (and aerodynamic resistance, only significant at high speed). In our case, to stop a mass of 450 t in horizontal motion, from 89 m/s to 0 within a braking distance of s=3900 m, assuming constant deceleration, the force needed is $F=ma=E_k/s=1800\cdot10^6/3900$=460 kN.
- How to dissipate the kinetic energy (and gravitational energy if slopping down). In our case, neglecting rotational kinetic energy, the dissipation rate is $\dot{W}=E_k/t_{dec}$, where the deceleration time is estimated with a mean speed of v_{mean}=89/2=44.4 m/s; i.e. $t_{dec}=s/v_{mean}$=3900/44.4=88 s, and hence $\dot{W}=E_k/t_{dec}=1800\cdot10^6/88$=21 MW.

Maximum deceleration is obtained by braking all wheels (with a control system to maximise resistance and avoid wheel freezing). For the maximum breaking force ($F_T=\mu mg$) we take μ=0.2 (see footnote in Table 1) so that now $F_T=\mu mg$=0.2·450·9.8=880 kN and $a=F_T/m=\mu g$=2 m/s^2. An estimation of the time and distance travelled for deceleration from 320 km/h (v=89 m/s) to 0 is obtained from the simple model of constant deceleration: $t=v/a$=89/2=45 s, covering a space of $s=\frac{1}{2}at^2=\frac{1}{2}\cdot2\cdot45^2$=2025 m, too short in comparison with the quoted data (3900 m to stop from 320 km/h), which suggests that the limitation is not on the force available but on the capability to dissipate at such a rate (for t_{dec}=45 s, $\dot{W}=E_k/t_{dec}=1800\cdot10^6/45=40$ MW). Dissipation is not very great at the contact patch because relative speed (slip in Fig. 2b) is not great; main dissipation is at the brakes mounted on the wheelset, which may be of different kinds:

- Friction disc brakes, where friction pads are pressed against a metal disc solidary with the wheels. They are most efficient for short-time braking because the pad material has a limited working temperature, about $T_{pad,max}$=650 K, and thermal control gets difficult for a prolonged period.
- Electromagnetic dissipators by eddy currents, where fixed electromagnets induce electric currents in a rotating metal disc that becomes hot. Thermal control is easier (it is the large disc that gets hot, instead of the smaller brake pads), and maintenance is minimal (no rubbing surfaces).
- Electromagnetic regenerative brakes, where the electric motors at the wheels are forced to work as electric generators; the generated electricity is fed back into the supply system through the pantograph, since there is too-much energy involved to be stored in batteries. These systems are always provided with a reostatic dissipator (resistor banks) mounted in a well-ventilated place, to cope with situations where the grid cannot accept the electricity produced.

Back to Propulsion