1
Introduction

The study of friction, wear and lubrication between two surfaces in relative motion is called *tribology*. This term is derived from the Greek verb 'tribos', which means 'to rub'. On one hand tribology aims at a scientific foundation of these phenomena. On the other hand it aims at a better design, manufacture and maintenance of devices which are affected by these 'annoyances'. Tribology has a very important economical outcome. According to one of the first reports on this issue, tribological problems accounted for 6% of the Gross Domestic Product in industrialized countries in the 1960s [160]. This percentage may have increased by now. Tribological problems are found in pinions, pulleys, rollers and continuous tracks, in pin joints and electric connectors, and may cause more failure than fracture, fatigue and plastic deformation. On the other hand, friction is highly desirable, or even essential, in power transmission systems like belt drives, automobile brakes and clutches. Friction can also reduce road slipperiness and increase rail adhesion. Before starting our rather theoretical description of tribology, it is important to recall the milestones that have marked the progress in this subject from the dawn of civilization.

1.1 Historical notes

More than 40 000 years ago a complex process such as the generation of frictional heat from the lighting of fire was already well known. Nowadays the same process is studied by a branch of tribology, which is known as 'tribochemistry' and is focusing, more generally, on friction-induced chemical reactions. The early use of surface lubricants to reduce friction is unambiguously proven by a famous painting from ancient Egypt, in which a 'prototribologist' supports the work of a few dozen slaves by pouring oil in front of the heavy sled that they are pulling (Fig. 1.1). More than four thousand years later Leonardo da Vinci (1452–1519) started a systematic investigation of tribology, as documented by his drawings (Fig. 1.2). Leonardo's

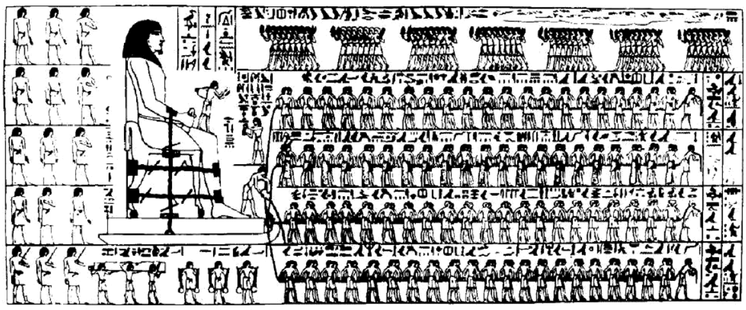

Figure 1.1 Transportation of an Egyptian colossus from the gravestone of Tehuti-Hetep (ca. 1880 B.C.). Note the officer at the feet of the statue lubricating the ground in front of the sled.

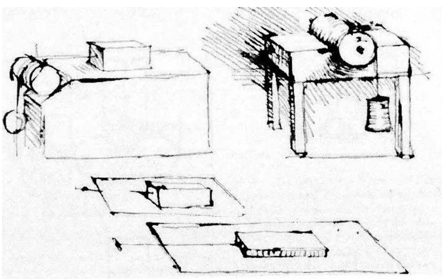

Figure 1.2 Original sketches of the friction experiments performed by Leonardo.

intuition and perseverance resulted in the formulation of the first friction law, which states the proportionality between friction and normal force. Nevertheless, this key observation quickly sank into oblivion until the French physicist Guillaume Amontons rediscovered it in 1699. The Swiss Leonhard Euler, possibly the greatest mathematician of the eighteenth century, was the first person who clearly distinguished between static and kinetic friction. Euler also made an attempt to relate friction to microscopic processes by speculating that friction is ultimately caused by the interlocking of rigid irregularities. A few years later, Charles de Coulomb, best known for his work on electricity and magnetism, observed that the kinetic friction is almost independent of the sliding velocity, whereas the static friction may vary depending on the time of stationary contact of the surfaces.

A turning point in the history of tribology was the theory of frictionless contact of non-conformal elastic solids. This theory was developed by the German physicist Heinrich Hertz in 1882 when he was only 23 years old, and forms the basis of modern contact mechanics. The Hertz theory was extended to include the contribution of adhesive forces by Kenneth Johnson and coworkers almost 90 years later. The difference between apparent and real contact areas was pointed out only in 1942 by Frank Bowden and David Tabor. They also proposed that the friction between two clean metal surfaces originates from the formation and rupture of cold weld junctions and concluded that, if the deformation of the junctions were entirely plastic, the coefficient of friction should be around 0.33, as indeed is measured in many metal pairings. The relation between friction and roughness was further investigated, among others, by John Archard (1957), Greenwood and Williamson (1966) and Bo Persson (2002), who proved, under more and more realistic conditions, that the real contact area is approximately proportional to the normal force.

A key process, when two surfaces slide past each other, is the so-called stick–slip. Stick–slip is caused by elastic instabilities and, in the context of atomic-size contacts, it was first modeled by Ludwig Prandtl in 1928, and experimentally substantiated sixty years later by atomic force microscopy (AFM). The slip can be thermally activated, which leads to characteristic variations of friction with temperature and driving velocity. Thermal activation is also important in capillary condensation and plastic flow, and may lead to various 'ageing' effects which are the focus of numerous theoretical and experimental investigations nowadays. On geological scales, stick–slip is also a key mechanism in earthquakes, as first recognized by Brace and Byerlee in 1966.

The advent of experimental techniques allowing one to measure friction down to the nanoscale and of fast computers allowing one to simulate the atomic interactions between two sliding surfaces resulted in the rise of the so-called 'nanotribology'. While the stick–slip motion of nanometer sized asperities can be readily investigated by AFM, other techniques such as the quartz crystal microbalance and the surface force apparatus have allowed researchers to measure the friction between adsorbate films and substrates or, respectively, between two atomically flat surfaces with intercalated lubricant films. On the other hand, molecular dynamics simulations are throwing light, more and more accurately, on the atomic origins of friction.

Without lubricants, almost no machine made of metal would work, and the Industrial Revolution would not have occurred. The theory of hydrodynamic lubrication was pioneered by Euler, Bernoulli, Poiseuille, Navier and Stokes between 1730 and 1845. It was the last mentioned who discovered that the frictional drag on a spherical particle slowly moving in a fluid is proportional to the velocity of the sphere. A series of key experiments was conducted by Gustave Hirn, who observed that the friction in a bearing is proportional to the sliding velocity and to the viscosity of the lubricant oil. An interpretation of his results, based on hydrodynamic lubrication and not on the more established concept of interlocking asperities, was first given by Nikolai Petrov in 1883, whereas the theory of fluid mechanics was fully established by Osborne Reynolds. Even if the Reynolds theory is still widely used in the design of modern lubricated machinery, this theory breaks down when the separation between the two sliding surfaces becomes comparable to their roughness. Systematic investigations of this problem were performed by Richard Stribeck, who introduced a curve which still holds his name (1902). The concept of boundary lubrication was introduced in 1922 by the biologist William Hardy while studying friction on solid surfaces covered by fatty molecules with hydrocarbon chains of different lengths.

As mentioned above, plastic flow plays an important role in contact ageing. The theory of plasticity, from the Greek verb ‘plassein’, meaning ‘to shape’, is also important in determining the stability of soils. Criteria for the yielding of these materials were proposed in early works by Coulomb and the Scottish engineer William Rankine, whereas the first scientific studies of plasticity in metals started only in 1864, when the French engineer Henri Tresca (whose name is also associated with the construction of the Eiffel Tower) published his famous criterion for yielding. This criterion was improved by Richard von Mises in 1913 and fundamental investigations of plasticity flourished in Germany in the early twentieth century under the leadership of Prandtl, who introduced the concept of plastic flow. The theory of plasticity is supported nowadays by powerful computer simulations, which are essential to control technological processes such as the rolling of strips or the extrusion of rods and tubes.

Our overview would not be complete without mentioning wear processes. Wear was well known to our ancestors, who exploited it to create artistic sculptures and useful tools by rubbing dense stones against softer ones in different ways. In spite of its importance, the variety and complexity of wear phenomena make the development of general physics laws interpreting wear processes quite challenging. Related to wear (and to friction) is the study of fracture mechanics, which was initiated by the British engineer Alan Griffith during World War I. The Griffith’s criterion, which is based on simple energetic considerations, can elegantly explain the failure of brittle materials. Fracture dynamics is not fully understood and is nowadays a subject of beautiful theoretical and experimental investigations.

2

Dry friction and damped oscillators

In this chapter we introduce the two categories of friction forces experienced by a rigid object sliding on a solid surface or moving in a viscous fluid. These forces have a different nature. Sliding friction increases with the normal force and is usually independent of the velocity. Viscous friction depends on the shape of the object and is proportional to the velocity, provided that this is low enough. Furthermore, while an object in a fluid can be set into motion by an arbitrarily low force, this is not the case if the same object lies on a solid surface, since a static friction force needs to be overcome in this case. Static friction allows us to join objects together using screws. It also has a key role in the propulsion and braking of vehicles and in transmission belts. Sliding (or kinetic) friction is important in pivots and collar bearings, not to mention uncountable situations in everyday life. Viscous friction can be exploited in mechanical dampers to mitigate the effects of forced oscillations. Since the theory of these oscillations is of pivotal importance in physics and engineering, it will be recalled in this chapter, whereas a detailed description of various situation involving viscous drag is provided in the last part of the book.

2.1 Amontons' law

In order to start and to keep moving a solid block on a solid surface, different *friction forces* F_{fric} have to be overcome and opposed. The *static friction* F_{s} corresponds to the minimum tangential force required to initiate sliding. The *kinetic friction* F_{k} perfectly balances the tangential force needed to maintain the sliding at a given (average) speed. These forces are intrinsically different. The static friction does not do any work, while the kinetic friction equals the dissipative work done at the interface divided by the distance covered by the block.

According to *Amontons' law* [5], the friction force is proportional to the normal force F_{N} acting on the block:

$$F_{\mathrm{fric}} = \mu F_{\mathrm{N}}. \tag{2.1}$$

Table 2.1 *Typical coefficients of static and kinetic friction.*

Physical situation	μ_s	μ_k
Rubber on concrete	1.0	0.8
Steel on steel	0.74	0.57
Aluminum on steel	0.61	0.47
Glass on glass	0.94	0.4
Copper on steel	0.53	0.36
Wood on wood	0.25–0.5	0.2
Wood on wet snow	0.14	0.1
Metal on metal (lubricated)	0.15	0.06
Wood on dry snow	–	0.04
Teflon on teflon	0.04	0.04
Ice on ice	0.1	0.03
Synovial joints in humans	0.01	0.003

Furthermore, it is independent of the nominal area of contact. The ratio between F_{fric} and F_N is the *coefficient of friction* μ, and it is usually different for static and kinetic friction. The static friction coefficient depends on the time of stationary contact (so-called *contact history*), on the elastic and geometric properties of the contacting surfaces and on the way in which the driving forces are applied. On the other hand, the kinetic friction coefficient is much better defined once the temperature, humidity, velocity and surface properties are reproducible. The values of the friction coefficient are usually lower than one, and the static coefficient μ_s is always equal to or larger than the kinetic coefficient μ_k.

As far as we are concerned with macroscopic contacts, we will also accept the validity of *Coulomb's law* and assume that, under dry conditions, F_k is independent of the sliding velocity. This is not the case at very low or very high velocities, where thermal effects or, respectively, inertial effects become important.

A representative list of friction coefficients is given in Table 2.1. For lubricated metal surfaces typical values of μ_s are in the range of 0.1–0.3. Higher values are observed after prolonged sliding if the lubricant film is worn off. For common engineering surfaces the friction coefficient does not depend significantly on the surface roughness, unless the surfaces are extremely smooth or rough. Amontons' law is also modified in the presence of strong adhesive forces.

Suppose now that a block rests on a plane inclined by an angle α, as in Fig. 2.1. If α is slowly increased, the block will start moving when

$$\tan\alpha = \mu_s. \tag{2.2}$$

This value defines the *angle of friction* (or *angle of repose*) α_c. If $\mu_s = 0.1$ the angle of friction is about 6°. Thus, the coefficient of static friction can be simply estimated by measuring α_c.

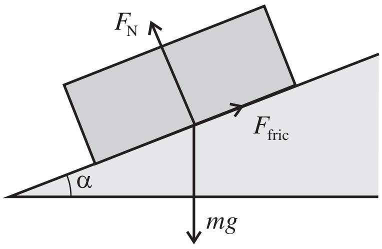

Figure 2.1 Forces on a solid block of mass m resting on an inclined plane ($g = 9.8$ m/s^2 is the acceleration due to gravity).

2.2 Applications to representative mechanical systems

Amontons' law is essential to understand the operation of machines and common mechanical parts. Here, we will discuss a few examples to illustrate its usefulness. We will not distinguish between static and kinetic friction coefficients since the values to be used are clear from the context.

Screws

A *screw* is able to convert a torque into a linear force. A square-thread screw with a mean radius R and pitch[1] b can be seen as a plane inclined by an angle α such that $\tan\alpha = (b/2\pi R)$, and wrapped around a cylinder. Thus, the screw will be self-locking if $\alpha < \alpha_c$, where α_c is defined by Eq. (2.2). Due to this possibility, the applications of screws for holding objects together are uncountable.

Screws can in principle also be used in power transmission, although they are not very efficient in this case. Exploiting the analogy with the inclined plane, it can be indeed demonstrated that the *efficiency* of a screw, i.e. the ratio between the useful work done and the energy transferred to a mechanism, is

$$\eta = \frac{\tan\alpha}{\tan(\alpha + \alpha_c)}.$$

The maximum efficiency η_{max} is achieved when $\alpha = 45^\circ - \alpha_c/2$ and is equal to

$$\eta_{max} = \frac{1 - \sin\alpha_c}{1 + \sin\alpha_c}.$$

If $\mu = 0.1$, a value of $\eta_{max} \approx 0.82$ is reached when $\alpha \approx 42^\circ$. If $\alpha = \alpha_c (\approx 6^\circ)$, the efficiency drops to 0.49. These low values, compared to other transmission mechanisms such as belt drives (see below), explain why 'lead screws' are rarely used for transferring large amounts of power.

[1] The pitch of a screw is the rise corresponding to a rotation of 360°.

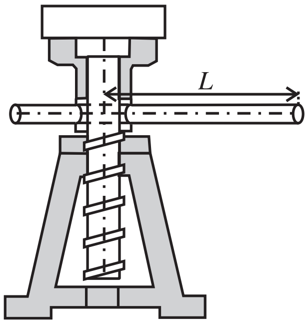

Figure 2.2 A jackscrew.

Jackscrews

If a screw is used as a lifting machine, we can define its *mechanical advantage* MA as the ratio between the load which can be lifted by the screw and the horizontal force F_{ext} applied on it. It is not difficult to see that

$$\mathrm{MA} = \frac{1}{\tan(\alpha + \alpha_{\mathrm{c}})}.$$

The mechanical advantage can be increased by one or two orders of magnitude if the force is applied to the end of a long horizontal bar connected to the screw, as in Fig. 2.2. In this case

$$\mathrm{MA} = \frac{L/R}{\tan(\alpha + \alpha_{\mathrm{c}})}, \tag{2.3}$$

where L is the length of the bar. If $L = 20R$, $\alpha = \alpha_{\mathrm{c}}$ and $\mu = 0.1$, the 'jackscrew' will be able to lift a weight almost 100 times larger than the applied force.

If a V-thread with angle of inclination β is used, the previous formulas are still valid, provided that the coefficient of friction μ in the definition (2.2) is replaced by $\mu/\cos\beta$. In this case the value of α_{c} increases, and the mechanical advantage is reduced. Nevertheless, V-threads are easier to manufacture and for this reason they are much more common than square-threads.

Pivots, collars and clutches

Pivots and *collar bearings* are commonly used to support an axial load acting on a rotating shaft. Pivots are placed at the end of the shaft, whereas collars can be located at any position (Fig. 2.3). If the pressure is uniform it is easy to see that the frictional torque acting on a flat pivot of radius R (Fig. 2.3(a)) is

$$M_{\mathrm{fric}} = \frac{2}{3}\mu F_{\mathrm{N}} R. \tag{2.4}$$

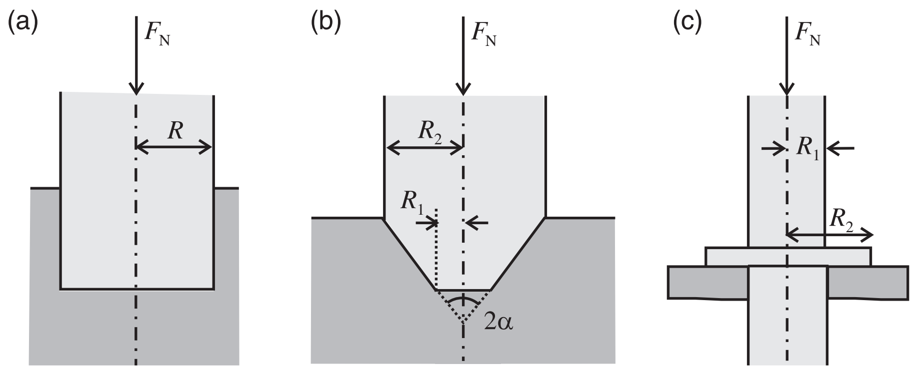

Figure 2.3 (a) A flat pivot, (b) a conical pivot and (c) a collar bearing.

For a truncated conical pivot:

$$M_{\text{fric}} = \frac{2\mu F_N}{3\sin\alpha}\frac{(R_2^3 - R_1^3)}{(R_2^2 - R_1^2)}, \tag{2.5}$$

where R_1 and R_2 are, respectively, the inner and outer radii of the pivot, and α is the half-angle of the cone (Fig. 2.3(b)). Equation (2.5), with $\alpha = 90°$, can be also applied to collar bearings (Fig. 2.3(c)).

However, since the wear rate at a given pressure is proportional to the sliding velocity and hence to the distance r from the axis of the shaft, a bearing will be more and more damaged at increasing values of r and the pressure distribution will consequently change with time. This means that the formulas (2.4) and (2.5) are strictly valid only for brand new elements. A good agreement with observations is found assuming that the wear rate becomes uniform. Since the wear rate is proportional to pr, where p is the pressure, it can be proven that, in this case, the frictional torque is

$$M_{\text{fric}} = \frac{1}{2}\mu F_N R$$

for a flat pivot, and

$$M_{\text{fric}} = \frac{\mu F_N}{2\sin\alpha}(R_1 + R_2) \tag{2.6}$$

for a truncated conical pivot. Equations (2.5) and (2.6) can be also applied to plate clutches (with $\alpha = 0$) and to conical clutches connecting two shafts rotating at different speed.

Belt drives

Consider two pulleys with radius R connected by a flexible elastic belt (Fig. 2.4). The initial tension in the belt is T_0. If a torque M_{ext} is applied to one of the pulleys, it

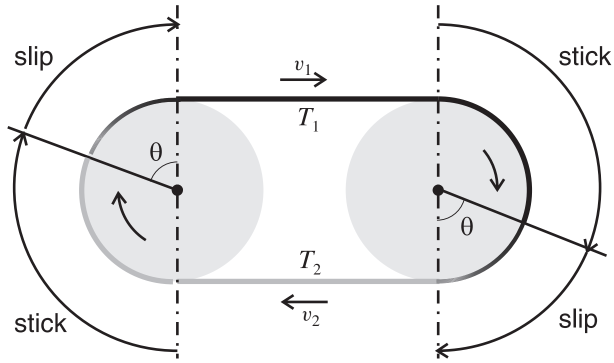

Figure 2.4 Belt transmission.

will be transmitted to the second one. This causes a difference between the tensions T_1 and T_2 on the tight side and slack side of the belt, which is given by

$$T_1 - T_2 = M_{\text{ext}}/R.$$

The tension increases over an arc $R\theta$ around the driving pulley, where the belt slips. The length of this arc is determined by the *capstan equation*

$$\mathrm{e}^{\mu\theta} = T_1/T_2,$$

which is obtained by integrating the tension variation over an infinitesimal arc $\mathrm{d}\theta$:

$$\mathrm{d}T/T = \mu\, \mathrm{d}\theta$$

(μ is the coefficient of friction between belt and pulley). The slip arc is located next to the point where the belt runs out of the driving pulley. A corresponding arc is located around the driven pulley, and this pulley runs slower than the driving pulley in proportion to the transmitted torque. If v is the mean speed acquired by the belt, the transmitted power is

$$P = (T_1 - T_2)v = 2T_0 v \frac{\mathrm{e}^{\mu\theta} - 1}{\mathrm{e}^{\mu\theta} + 1}.$$

Note that the efficiency of a belt is typically very high ($\sim$0.9). Since transmission belts do not require lubrication and are relatively cheap, they have found numerous applications ranging from automotive engines to transportation of heavy materials.

The previous formulas are also valid for a V-grooved belt, provided that μ is replaced by $\mu/\sin\beta$, where β is the half-angle of the groove profile. In this case the length of the slip arc can be reduced significantly. For this reason V-grooved belts are the most common choice for applications to power transmission.