

Basics of Sliding Friction between Solid Bodies

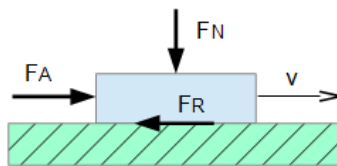
Introduction

Dry friction between two bodies under relative movement is crucial in some technical fields. The most common application areas are certainly the various braking systems of vehicles, as well as the clutches in vehicles with combustion engines. Also sliding bearings for rotation and linear movement, which work without external or internal liquid lubrication are widespread. While the coefficient of friction at braking systems and clutches should be mainly constant and rather high at all operating points, the development target of sliding bearings is normally a low coefficient of friction. Further requirements for all of this application areas are: low wear and especially for braking systems: a very high mechanical and thermal safety.

So far, there is no theoretical basis to forecast the coefficient of friction of two solid bodies moving relative to each other without experiments. The state of the art has so far developed from empirical values and tests. It is the target of this publication to offer this theoretical basis so that further research can build on it.

I. The Primary Principle of Dry Friction

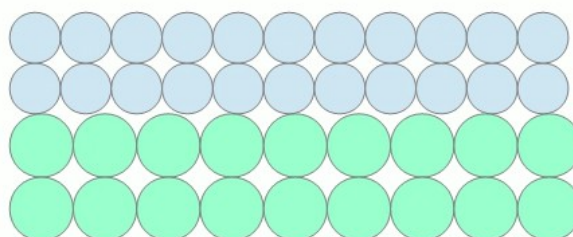
The friction between two solids can be described in the state with or without relative movement to each other (sliding friction or static friction) by the following model.



Here the flat surfaces of two bodies are loaded by a vertical pressure force F_N , which often corresponds to the inertia force of the upper body under gravity. According to the law of Coulomb, there is a constant ratio by definition between the friction force F_R and the load F_N for every material combination for sliding friction and as well as for static friction, named the coefficient of friction μ , whereby the magnitude of the friction force is identical with the moving force F_A .

Therefore it is: $\mu = F_R / F_N$

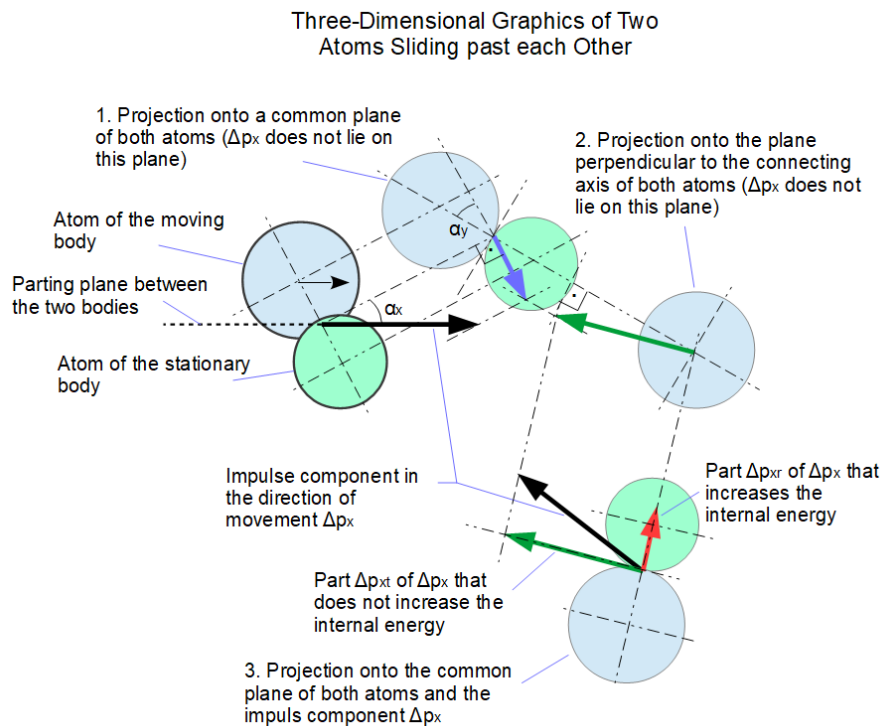
In the theories of heat transfer, the energy transfer through heat in the contact area of two solid bodies is explained by the temperature-dependent intensity of the atomic oscillations, which stimulate each other through pulse transfer. In a thermally balanced state (both bodies have the same temperature) the flows of energy between the bodies are identical. This is shown in principle with the following animation [animated only at dry-friction.com]:



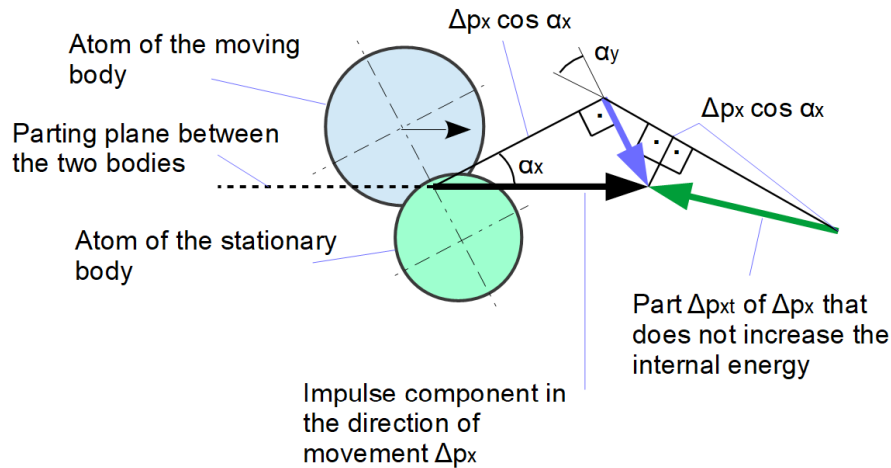
By combination of this theory with the law of Coulomb, it becomes clear that the opposite atoms of both bodies in the contact area are in relative movement to each other at temperatures above absolute zero, even if there is no relative movement at the macroscopic level between the bodies (static condition). In the static condition, the force F_N is absorbed by a finite number of atom contacts without any further energy transfer. A single atomic contact transmits (at a time) the force ΔF_N . For statistical reasons, the vector sum of all these forces has no component along the contact plane.

Because the total energy balance in the static and thermal balanced condition is zero, the impulse exchange between the single opposite atoms does not need to be considered quantitatively and therefore in the condition of relative movement of the two bodies to each other, only the hereby caused vectorial component is relevant in the impulse exchange. In this condition, every atomic oscillation of the moving body receives an additional component in the direction of movement. This has an effect to the opposite atoms in the contact area, which have to be passed and - as already mentioned - transfer an individual part ΔF_N of the normal force F_N . According to the following figure these oscillation components caused by the movement can be resolved in a (red) part directed to the center of the opposite atom and so increases the internal energy of the bodies (temperature increases) and a (green) part directed tangential along the atomic "spherical surface" without resistance. A proportion of the individual force ΔF_N must be provided in the same ratio in order to supply this energy.

However, the representation of the atoms as balls does not allow the energy transfer to be viewed as a non-central impact of two balls, because it is a contact between two force fields and not between material balls with friction to each other. So no moment is transmitted over the distance between the contact point and the atomic center.



The amount of the (green) vector Δp_{xt} can be calculated from the right-angled triangles formed by the vectors.



The amount of the (red) vector Δp_{xr} can then be determined using the Pythagorean theorem.

Then the following relationship between Δp_{xr} and Δp_{xt} is obtained:

$$\frac{\Delta p_{xr}}{\Delta p_{xt}} = \frac{\sin \alpha_x \cdot \cos \alpha_y}{\sqrt{1 - \sin^2 \alpha_x \cdot \cos^2 \alpha_y}}$$

So far only those constellations have been considered in which atoms of the moving body have to pass atoms of the resting body. Statistically, however, there are just as many pairings in which an opposite atom is contacted “downhill” generating thereby no impact component to increase the inner energy.

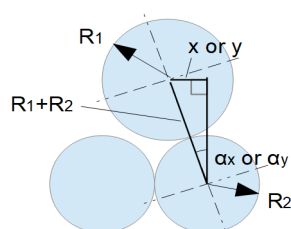
A primary coefficient of friction μ_p can be defined, which represents the ratio between the sum of all red and all green individual impacts. The total coefficient of friction μ_g is then:

$$\mu_g = \mu_p \times \mu_s$$

where μ_s is the secondary coefficient of friction resulting from secondary effects influencing the moving force. These effects will be discussed later.

In order to determine the distribution of all impacts between two bodies in relation to the contact area it is useful to replace the angular references by the distances x (direction of movement) and y (perpendicular to the direction of movement) between the opposing pairs of atoms projected onto the contact plane.

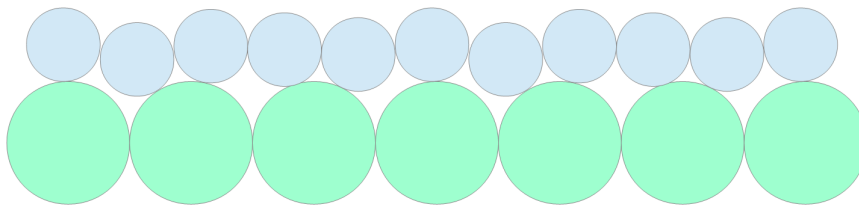
Two-Dimensional Illustration of the Dependence between the Contact Angle α_x or α_y and the Distance x or y



However, this model assumes that the atoms of the “upper” body are free to immerse themselves in the “valleys” between the opposing atoms, while the atoms of the lower body do not change their position. Of course, this does not correspond to reality.

Looking at a larger section two-dimensionally, it becomes clear that the atoms must certainly move out of their surrounding structure in order to come into contact with the opposite atom. To find out which of the two bodies is more flexible, another physical property has to be taken into account, namely the modulus of elasticity. The smaller the modulus of elasticity, the more flexible is the structure. If both bodies have different E-modules, then the atomic structures with a lower E-module adapt more to the surface of the opposite body than vice versa. This in turn has an effect on the largest occurring angle α_x or α_y and thus also on the distance x or y .

Two-Dimensional Illustration of the Contact Zone between Bodies with Different Atomic Sizes and very Different Moduli of Elasticity



The model for determining the relationship between the contact angle α_x or α_y and the distance x or y comes close to the real conditions when the upper atom belongs to a material with a significantly lower E-module than the lower one, which is actually the case in many technical applications and will be discussed later. The development of a model that applies to any relationship between the E-modules would certainly be desirable, but should not be pursued further at this point.

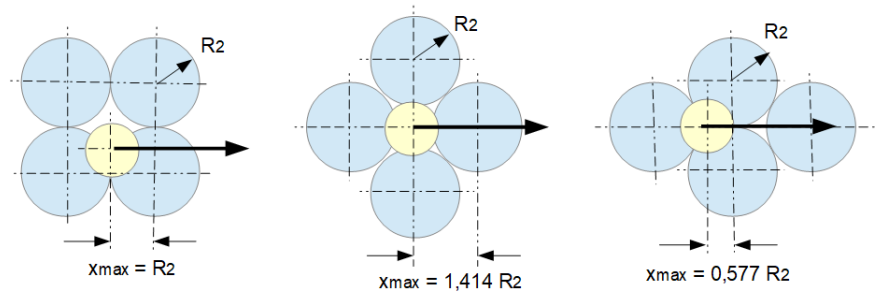
The following applies under the mentioned E-module ratios:

$$\alpha_x = \arcsin\left(\frac{x}{R_1 + R_2}\right) \quad \text{and} \quad \alpha_y = \arcsin\left(\frac{y}{R_1 + R_2}\right)$$

So the equation for $\Delta p_{xr} / \Delta p_{xt}$ will change to:

$$\frac{\Delta p_{xr}}{\Delta p_{xt}} = -\sqrt{\frac{x_r^2 y_r^2 - x_r^2}{x_r^2 y_r^2 - x_r^2 + 1}}$$

This equation, from which the primary coefficient of friction μ_p could be calculated, should not be further developed here in a generally applicable manner, since the application to a real case is not trivial, because the maximum values of x and y that have to be known for calculating the integral of this equation are different depending to the orientation of the opposing atoms, as the following extreme examples show in the top view at the contact plane.



In the general case, the “upper” atom moves along any oblique line to the “lower” atomic structure.

The calculation for a real case is even more difficult by the fact that materials normally consist of molecules with different elements, i.e. the materials consist of atoms with different atomic radii. It must also be taken into account whether the various elements of a material all come into contact with the elements of the opposite body with the same probability. This is certainly also dependent on the arrangement of the elements within the molecules - some elements are e.g. always inside a molecule and can therefore never come into direct contact with the opposite body. In the case of metallic surfaces, the corrosion layer must also be taken into account, which is always present, if only sometimes with a minimal layer thickness. This important topic will be discussed later. Furthermore, some materials consist of different molecules, which would make the solution of the above equation even more complex.

As mentioned at the beginning, the described primary principle of dry friction only relates to the state of sliding friction. Static friction is subject to completely different laws, since no energy transfer takes place. Therefore the widespread assumption is wrong that the static friction between two solid bodies is always higher than the sliding friction. The static friction will not be discussed further in this work.

II. Secondary Frictional Influences

In addition to the primary cause of friction between solids that has been dealt with so far, there are also a number of secondary influencing variables that can be summarized in the already mentioned secondary coefficient of friction μ_s and are considered in the following:

As soon as wear occurs on the contact surfaces during movement, work has to be expended to detach the wear particles from the surfaces. Welding processes between materials that are ripped out of one of the two bodies as a result of movement also act in the same way. In both cases, this increases the moving force F_A .

Two materials, which have contact surfaces with a high-quality flatness in connection with a very low roughness, have a "force of attraction" to each other that is not negligible and adds to the compressive force F_N and thus also increases the moving force F_A . This effect is often ascribed to adhesion, but is at least partially based on the following fact: Surfaces that are actually in contact with each other no longer have any ambient air between them. This means that the ambient air pressure does not act on these surfaces, but continues to act on all other surfaces of the bodies involved. This increases the load between the contact surfaces under normal ambient conditions by 0.1 MPa. A pressure at this level is not negligible, since e.g. in real applications with unlubricated linear plain bearings, the surface pressures are in the majority of cases well below 1 Mpa¹. That would be a

component in explaining why the measured coefficient of friction in many of these bearings increases with increasing wear. When new, the actual contact areas due to manufacturing inaccuracies (e.g. with injection molded plastic sliding elements) are still very small compared to the theoretical available contact area. The air pressure effect then hardly plays a role. With increasing running distance between the contact partners and unavoidable wear, the actual contact area increases and the air pressure effect becomes more noticeable. In an application with a designed surface pressure of 0.5 MPa for example, the actual surface pressure due to the air pressure then increases by 20% to 0.6 MPa. This effect can be minimized if one of the two surfaces shows negligible wear in connection with a certain roughness or waviness in the macro range, which ensures that a large proportion of the ambient air can remain between the bodies in contact. In turn, high humidity can lead to an increase in these forces if water displaces a part of the air in the remaining spaces.

It is also imaginable that the air pressure effect causes resonances and thus noises, since the effect stops almost suddenly as soon as the material contact is interrupted (of course the air also needs a certain time to flow in). A renewed contact could cause a recurring loss of contact through the resonance behavior of the materials.

Other secondary effects to be mentioned are:

- Material transfer (e.g. in the case of plastics), whereby worn particles of one body stick to the opposite body or can even form a complete coating over time and thus change the tribological properties of the system. Environmental contaminations (e.g. dust) that are deposited on the contact surfaces also have a similar effect.

- plastic deformation of the surface structure of a material or both materials in the micro range needing additional energy.

- elastic deformation of the surface structure of a material or both materials in the micro range needing additional energy if the material has a high degree of damping (such as elastomers). This effect is also speed dependent and is possibly the explanation for the strong speed dependence of many tribological pairings with plastics.

- Speeds that are so slow that creeping processes within a body (e.g. with plastics) play a role (quasi-fluid friction) (e.g. floating bearings for bridges with the sliding pairing PTFE against metal).

- Partial phase transformations of a material during contact with the opposing body from solid to liquid (e.g. with ice). This effect causes an improvement in the overall coefficient of friction μ_g . In the case of a phase transformation due to mutual atomic stimulations, the impact component that is directed to the center of the opposite atom is reversibly converted. It can only be assumed here that the force components of ΔF_N , which are needed for this reversible conversion, finally will be neutralized by forces which occur when the phase transformation reverses.

- Partial phase transformations of a material in the form of a change in the crystal structure during contact with the counter body (an effect that may play a role at PTFE, which changes its crystal structure at 19 °C and 35 °C and amorphous components at 127 °C pass over in a supercooled liquid ²). This effect also improves the overall coefficient of friction μ_g and can be explained like the phase transformation from solid to liquid.

- Static charge, if both bodies are not electrically conductive. There are no reliable theories about this, but an influence should not be excluded as long as the irrelevance is not proven.

III. Simplified Calculation Example

Despite the complexity of the real atomic arrangements on the contact surfaces of two bodies moving against each other described in Chapter I, it would be interesting to consider a simplified case and to determine the approximate value of the primary coefficient of friction.

Therefore, the simplest case of two bodies, each consisting of only one element, shall be considered. Then the last determined equation

$$\frac{\Delta p_{xr}}{\Delta p_{xt}} = - \sqrt{\frac{x_r^2 y_r^2 - x_r^2}{x_r^2 y_r^2 - x_r^2 + 1}}$$

can be solved, if mean value ranges for x_r and y_r are assumed that lie somewhere between the largest and the smallest possible range for x and y .

Since there is no reason why x and y will assume different ranges, they can also be equated. The following then applies:

$$\frac{\Delta p_{xr}}{\Delta p_{xt}} = - \sqrt{\frac{x_r^4 - x_r^2}{x_r^4 - x_r^2 + 1}} \quad \text{with} \quad x_r = \frac{x}{R_1 + R_2}$$

The primary coefficient of friction μ_p can be defined as 50% (because of the 50% "downhill" of the atoms) of the mean ratio of $\Delta p_{xr} / \Delta p_{xt}$ when considering this ratio over the distance x from x_{\max} to $x = 0$.

It should be assumed that the maximum value of x_r occurs on average at $x_{\max} = R_2$.

Then applies:

$$x_{r\max} = \frac{R_2}{R_1 + R_2}$$

and further:

$$\mu_p = -0,5 \int_{x_r = \frac{R_2}{R_1 + R_2}}^{x_r = 0} \sqrt{\frac{x_r^4 - x_r^2}{x_r^4 - x_r^2 + 1}} dx_r$$

This equation can only be solved numerically.

The sliding pairing of polyethylene (PE) against the anodized surface of aluminum should be considered. This also fulfills the requirement of a large difference in the modulus of elasticity between the two bodies (PE: 1,500 MPa and Al₂O₃: 406,000 MPa). PE has the chemical formula C₂H₄, whereby the hydrogen (atomic radius 32 pm) is arranged on the

outer shell of the PE molecule, which therefore and from the fact that it occurs twice as often as carbon, has by far the greatest probability to come into contact with the anodized surface. The anodized surface consists of aluminum oxide (Al_2O_3) with 40% aluminum atoms (atomic radius 143 pm) and 60% oxygen atoms (atomic radius 66 pm).

If calculating the primary coefficient of friction only with the atomic radius of the oxygen on the aluminum oxide surface, the result would be $\mu_p(\text{O}) = 0.160$, with the atomic radius of aluminum only it would be $\mu_p(\text{Al}) = 0.180$. With proportional weighting and assuming a linear dependency we get $\mu_p(\text{Al}_2\text{O}_3) = 0.172$ for the anodized surface.

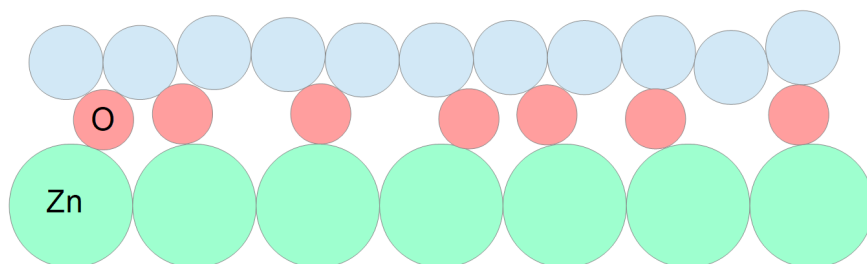
In order to determine the associated secondary coefficient of friction μ_s theoretically, the theoretical basis for the effects described is still largely missing. After successfully calculating the primary coefficient of friction, however, it would be possible to determine the sum of all secondary coefficients of friction indirectly by measuring the total coefficient of friction μ_g for a pairing under certain operating conditions (speed, surface pressure, temperature) using the equation: $\mu_s = \mu_g / \mu_p$

IV. Special Characteristics of Metallic Sliding Partners

As already mentioned above, under normal ambient conditions, oxidation processes always occur on the surface of metals, i.e. the metals combine there with the oxygen of the air to form metal oxides. However, depending on the metal, reactions with carbon dioxide and / or the hydrogen from the air humidity also occur on the surface. The elements withdrawn from the atmosphere will mainly be present on the outermost surface of the oxide layer. Every metal oxide has a fixed ratio between the number of atoms of the metal element and the atoms of the oxide-forming elements. This results in a more or less dense surface layer of the oxide-forming elements, which has a significant influence on the primary coefficient of friction.

A first suggestion should be made here to quantitatively estimate the influence of this effect. By building the ratio of the projected areas between the oxide-forming atoms and the metal atoms, a measure of the arrangement density of the outermost oxide layer in relation to the metal structure, which normally has a very dense arrangement of atoms, can be created. In the case of zinc oxide (ZnO) this ratio is e.g. 24% or 48% for aluminum oxide (Al_2O_3), that means, that the relatively small oxygen atoms have relatively large distances with large gaps between, which would result in a large proportion of unfavorable atomic contact angles in the case of sliding friction with other substances.

Two-Dimensional Illustration of the Contact Zone between an Oxidized Zinc Surface and a Material with a Significantly Lower Modulus of Elasticity



For iron oxide $\text{Fe}(\text{OH})_2$, on the other hand, the ration in question is 242%, which means that on the surface of the iron crystal lattice, an $(\text{OH})_2$ -group cannot be arranged at every

iron atom, but certainly there will be a fairly high dense arrangement of (OH)₂ molecules, which should result in a relatively good coefficient of friction. In fact, tests have shown that zinc and aluminum surfaces tend to cause high coefficients of friction, while steel and stainless steel surfaces are more favorable.

Further studies would also have to clarify how the equilibrium between surface wear and corrosion layer formation is established quantitatively at permanent sliding movement.

Metal oxides are also specifically produced technically in order to protect against corrosion and to create decorative surfaces. Accordingly, anodizing is a widely used process for aluminum. The arrangement of the oxygen and aluminum atoms on the surface will certainly differ from the oxidation in ambient air. Experience has shown that anodized surfaces have a positive effect in terms of low friction. This indicates a surface with a very dense atomic arrangement.

V. Requirements for Technically Suitable Dry Sliding Systems

Not every material pairing is suitable for dry relative movement to one another. The less the two bodies are able to adapt to one another, the higher is the specific load on the bodies. This is especially the case with material pairings that both have a relatively high modulus of elasticity. If the load exceeds a material-specific limit value, the corresponding material is destroyed at this point. This further increases the load on other contact regions., which then results in exponentially increasing wear. There are not many material pairings that have a technically acceptable wear rate under the condition that the geometrical surface quality is suitable, and even fewer that also allow high relative speeds. Since good heat conduction from one of the two materials is always required at high speed in order to dissipate the resulting frictional heat, one of the two materials usually consists of metal or ceramic, which both have a relatively high modulus of elasticity. The other material normally has a significantly lower modulus of elasticity in order to adapt as best as possible to the surface of the partner.

Epilogue

Dynamic sliding friction between solid bodies exists in its familiar appearance only because their surfaces consist, microscopically, of an arrangement of oscillating spherical segments.

Citations:

¹ The author's experience from many years of industrial development work

² Stefan Kirchberg; *free translation from German*: Determination of the elastic properties of PTFE and PTFE compounds using dynamic mechanical analysis (DMA) - [Ermittlung der elastischen Eigenschaften von PTFE und PTFE-Compounds mittels Dynamisch-Mechanischer-Analyse (DMA)]; GRIN Verlag 2016

Author: Horst Moshhammer, Am Sändchen 8, 53545 Linz am Rhein, Germany

Status: 11-30-2020