2

Friction Modeling

2.1 Introduction

The desire to understand friction is old. The first concepts date as far back as to the work of Leonardo da Vinci. In this chapter we briefly review some of the various elements of the frictional contact and give a somewhat intuitive picture of the origin of friction. A more thorough description of the mechanisms behind friction can be found in Bowden and Tabor (1973). Various experimental observations of friction behavior are summarized. Most of the chapter is devoted to the issue of friction modeling and models for various purposes are reviewed. The chapter also serves as an introduction to the concepts and the terminology used when discussing friction.

2.2 The Friction Interface

Friction occurs between all surfaces in contact. In engineering applications it is common with lubricated metal to metal contacts such as in ball bearings. The discussion here will mainly relate to such contacts. To get an understanding of the mechanisms behind friction it is necessary to observe the microscopical contact between two surfaces. A simplified picture is shown in Figure 2.1. The surfaces are naturally rough and often covered by a layer of oxide or some other material. If the contact is lubricated there will also be oil or grease present in the interface. The actual contact takes place at a number of asperities and not continuously over the surface. Deformations of the contact points occur due to the load. There may also be adhesion and welding processes. As a tangential force is applied shearing of the contacts take place. This results both in elastic

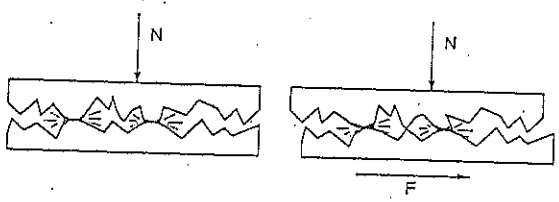

Figure 2.1 The microscopical contact between two surfaces. The contact takes place at a number of asperities which deform under normal and tangential loads.

and plastic deformations. As motion occurs between the surfaces more of the lubricant is brought into the interface.

The mechanisms behind friction can roughly be divided into four regions of behavior, see Armstrong-Hélouvry (1991). These depend on the relative velocity between the surfaces. Figure 2.2 shows a relation between friction force and velocity divided into four so called lubrication regimes. The first regime covers the case when no gross motion occurs, i.e., when the surfaces stick. A good mental picture of the behavior is shown in Figure 2.3. The contact can be viewed as formed by a number of springs. As a force is applied the springs are extended which results in the friction force. If the extension becomes too large the springs snap and sliding occurs. Thus, when sticking, the friction force is due to elastic and plastic deformations of the asperity contacts.

For the second regime motion occurs between the surfaces but hardly no lubrication is present in the interface. The friction force in this boundary lubrication regime is due to the shearing resistance of the asperity

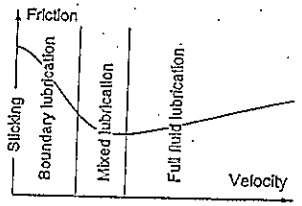

Figure 2.2 A typical relation between velocity and friction force. The relation can be divided into four so called lubrication regimes. The different regions are due to the velocity dependent causes of friction.

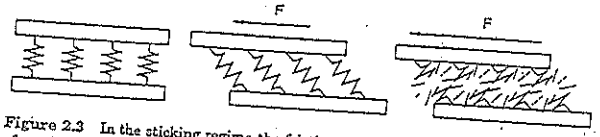

Figure 2.3 In the sticking regime the friction contact can be viewed as consisting of small springs which give rise to the friction force as they are extended. If the displacement becomes too large the springs snap and gross sliding can occur.

contacts. Normally the surfaces are covered with oxides or other compounds. The shearing resistance of these are much lower than for the metal.

As the sliding speed increases more and more lubricant is brought in, which increases the separation of the surfaces. The transitional region before full fluid lubrication is called the mixed lubrication regime. The friction force is partially due to the lubricant and its viscosity and partially due to asperity contacts.

Finally, as the two surfaces are completely separated by the lubricant, full fluid lubrication occurs. The origin of the friction force for this lubrication regime is, of course, found in the hydrodynamics of the lubricant.

Naturally, dynamics are involved in all the mechanisms like the formation of a lubrication film and the shearing of junctions. The complete behavior of the friction force is therefore very complex. Next we describe some of the rich behavior that friction may exhibit.

2.3 Experimental Observations

The behavior of friction has been extensively examined during the 20th century. Many experiments have, contrary to the conditions for engineering applications, been performed with clean surfaces and for stationary conditions, e.g., constant velocity. Lately the interest in friction dynamics has increased. Some experimental observations of friction are reviewed below. The collection is by no means complete but serves to illustrate the many facets of friction behavior.

Steady Velocity Friction

The friction force as a function of velocity for constant velocity motion typically looks as the curve in Figure 2.2. The relation is called the Stribeck curve after Stribeck (1902). In particular the shape of the force in the

boundary lubrication regime and mixed lubrication regime is called the Stribeck effect. No universal function can be given that describes friction as a function of velocity. Instead the relation is application dependent and varies with material properties, temperature, wear etc.

Many friction phenomena do not appear for constant velocity experiments. A number of observations of the dynamic behavior of friction are listed next.

Static Friction and Break-Away Force

The friction when sticking is called static friction and the force necessary to initiate motion, i.e., to overcome the static friction, is called the break-away force. Many experimental investigations were performed in the 50s to study the nature of these two concepts.

Rabinowicz addressed the transition between sticking and sliding in Rabinowicz (1951). He argued that the transition cannot be described satisfactorily as a function of velocity. Instead he investigated friction as a function of displacement. A simple experiment was devised to determine the relationship. A block was placed on an inclined plane and a ball was rolled on the plane and impacted the block. The inclination was such that sliding motion of the block was sustained, but if the block was at rest, it stayed at rest. The distance that the block moved due to the energy transferred by the ball was determined and then used to describe the friction transition in a diagram as in Figure 2.4. The break-away force is given by the peak of the curve. As seen in the figure the maximum friction force occurs at a small displacement from the starting point.

Rabinowicz also investigated the static friction force as a function of dwell-time, i.e., the time spent sticking, see Rabinowicz (1958). This was done using so called stick-slip motion experiments. It was concluded that

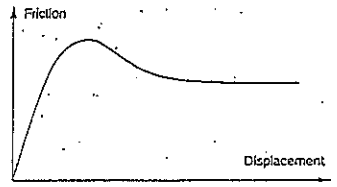

Figure 2.4 The relation between friction and displacement as found by Rabinowicz (1951). The experimental results suggested that friction should be described as a function of displacement and not velocity.

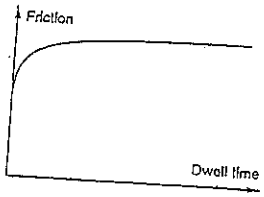

Figure 2.5 Relation between static friction and dwell-time as found by Rabinowicz (1958). The experiment suggested a temporal dependence on the level of static friction.

the static friction increased with the dwell-time. Relations of the form in Figure 2.5 were found.

In Johannes *et al.* (1973) it was pointed out that for stick-slip experiments the dwell-time and the rate of increase of the external force are related and hence the effects of these factors cannot be separated. They therefore redesigned the experiment so that the time sticking and the rate of increase of the applied force could be varied independently. The results showed, contrary to the results of Rabinowicz, that the break-away force did depend on the rate of increase of the external force but not on the dwell-time; see also Richardson and Nolle (1976). A characteristic behavior is seen in Figure 2.6.

Another investigation of the behavior in the sticking regime was done by Courtney-Pratt and Eisner (1957). They studied the spring-like behavior before gross sliding occurs. Their results were presented in diagrams showing force as a function of displacement. A typical relation is seen in Figure 2.7. This should be compared with Figure 2.4.

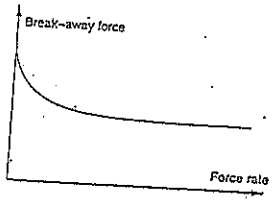

Figure 2.6 Characteristic relation between rate of force application and break-away force as found by Johannes *et al.* (1973). The experiment suggested that the break-away force decreases with increased rate of force application.

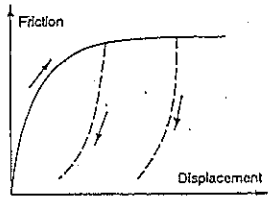

Figure 2.7 Pre-Sliding displacement as found by Courtney-Pratt and Eisner (1957). The result agrees with Figure 2.4 for small displacements. Releasing the applied force results in a permanent displacement as indicated by the dashed lines.

The experiments examining static friction have shown that even in the sticking regime microscopic motion occurs. This is often called pre-sliding motion. The break-away force which is necessary to cause gross sliding varies with the experimental conditions.

Frictional Lag

Dynamics are not only important in the sticking regime; they also affect the behavior in the other lubrication regimes.

The paper Hess and Soom (1990) has drawn a lot of attention. Hess and Soom chose to perform experiments with a periodic time-varying velocity superimposed on a bias velocity so that the motion becomes unidirectional. The velocity was varied in such a way that all lubrication regimes were covered. Typically the friction–velocity relation appeared as in Figure 2.8. Hysteresis was observed as the velocity varied. The size of the loop increased with normal load, viscosity and frequency of the

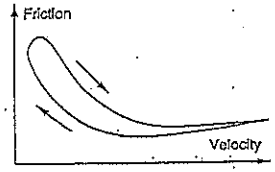

Figure 2.8 The friction–velocity relation observed in Hess and Soom (1990). The friction force is lower for decreasing velocities than for increasing velocities. The hysteresis loop becomes wider as the velocity variations become faster.

velocity variation. Hess and Soom explained the hysteresis using a pure time delay in a static relation between velocity and friction force.

The experimental observations have shown that it is necessary to consider the dynamics of friction in order to get a detailed understanding of the friction behavior.

2.4 Friction Models

A model of friction is necessary for many purposes. In some cases it is desirable to have a model which provides insight into the physical mechanisms of the friction interface. In others it suffices with a model that can predict the global, qualitative behavior of a system with friction. If the model is to be used for friction compensation there may be limitations on the computational complexity of the model. There are thus many purposes of friction modeling:

- mathematical analysis
- off-line simulations
  - global, qualitative behavior
  - high-fidelity behavior
- physical insight
- friction compensation

It is important to bear the purpose in mind when discussing friction modeling.

This section summarizes previous work on friction modeling and gives a variety of examples of friction models. The intention is not to give a complete review of all existing friction models but to give a glimpse of the wide spectrum that exists. The models have been divided into three categories: static models, dynamic models, and special purpose models. The first category includes models that to a varying extent gives a qualitative understanding of friction. The models in the second category include dynamics in order to more accurately describe the friction phenomena. The final category includes some models that give an understanding of the physical mechanisms behind friction.

Discussion

Figure 2.9 shows a block on a surface. It also introduces some of the notations that are used for the friction contact throughout the thesis.

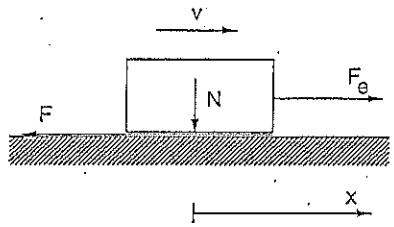

Figure 2.9 Notation for the variables affecting the friction force at the contact between two surfaces. The friction force is denoted F and any external tangential force F_e. Velocity and position are always denoted v and x, respectively. Motion is only considered in the horizontal direction.

Motion is considered only in the horizontal direction and friction occurs at the contact interface. The variables shown in the figure are: the friction force arising at the interface F, the relative velocity between the surfaces v, the relative position x, an external applied force F_e, and the normal force N. Furthermore, the contact area between the surfaces is denoted A. The early work on friction modeling tried to determine how these variables were related, e.g., Amontons (1699) and Coulomb (1785). It was, for example, investigated if the contact area affected the friction force. As a result friction became known as force which opposes motion and which magnitude is given by a coefficient of friction times the normal force. The relation thus looks like

$$F = \mu N \tag{2.1}$$

It is still very common to discuss friction in terms of the friction coefficient μ. Typical variations in the friction force which are handled using μ are variations caused by material properties and contact geometries. The description (2.1) is too simplistic for many purposes. The dependence on normal load is more complicated than described by (2.1), see SKF (1970). Another complication is that the normal force may change. Slow changes can be accounted for by scaling or adapting parameters of the friction model. Rapid variations may, however, effect the friction force significantly. This effect is not considered in the work in this thesis. We have chosen to model the entire friction force and not just the friction coefficient. However, if so desired, the models can also be thought of as modeling μ instead of F.

For certain applications, such as gears, friction may depend on rotational position and also on the direction of motion. This can be handled for example, by letting the coefficients of the models be position dependent. Such variations have not been taken into account in the model discussion.

Static Friction Models

The friction models from past centuries regard friction as a static function of velocity. Therefore, they are called static friction models and are described next. The category includes the classical descriptions of friction but also refinements that have been made in order to adapt these models to simulation demands.

Classical Models: The classical models of friction consist of different components, which each take care of a certain part of the friction force. The purpose of these models was to give a rough understanding of the behavior of friction, which agreed with simple experiments that could be performed.

The work on friction of da Vinci was rediscovered by Amonton in the late 17th century and then developed by Coulomb in the 18th century. The main idea is that friction opposes motion and that the friction force is independent of velocity and contact area. It can therefore be described as

$$F = F_C \operatorname{sgn}(v) \tag{2.2}$$

This description of friction is termed *Coulomb friction*, see Figure 2.10 a). Other names include kinetic friction, dynamic friction, sliding friction and relay-type friction. The Coulomb friction model does not specify the friction force for zero velocity. Depending on how the sign function is defined it may either be zero or it can take on any value in the interval between $-F_C$ and F_C. The Coulomb friction model has, because of its simplicity, often been used for friction compensation, see for example Friedland and Park (1991) and Børil (1993). Some confusion in the terminology exists and sometimes Coulomb friction denotes also velocity-dependent sliding friction as in Brandenburg and Schäfer (1991).

In the 19th century the theory of hydrodynamics was developed leading to expressions for the friction force caused by the viscosity of lubricants, see Reynolds (1886). The term *viscous friction* is used for this force component, which is normally described as

$$F = F_v v \tag{2.3}$$

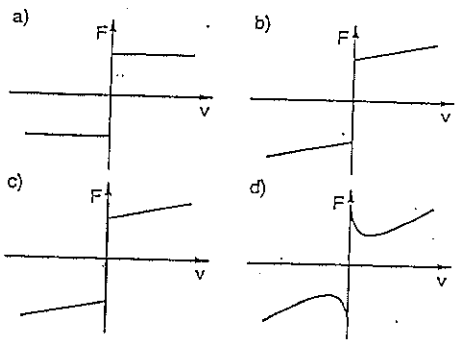

Figure 2.10 Examples of static friction models. The friction force is given by a static function except possibly for zero velocity. Figure a) shows Coulomb friction and Figure b) Coulomb plus viscous friction. Stiction plus Coulomb and viscous friction is shown in Figure c) and Figure d) shows how the friction force may decrease continuously from the static friction level.

Viscous friction together with Coulomb friction is shown in Figure 2.10 b). It is not always correct to let the viscous friction be linear in velocity. A more general relation is

$$F = F_v |v|^{\delta_v} \operatorname{sgn}(v) \tag{2.4}$$

where δ_v depends on the geometry of the application, see SKF (1970) and Andersson (1993).

Stiction is short for static friction as opposed to dynamic friction. It describes the friction force at rest. Morin (1833) introduced the idea of a friction force at rest that is higher than the Coulomb friction level. Static friction counteracts external forces below a certain level and thus keeps an object from moving. However, if the external force is too large then the friction force cannot prevent motion. It is hence clear that friction at rest cannot be described as a function of only velocity. Instead it has to be modeled using the external force in the following manner.

$$F = \begin{cases} F_e & \text{if } v = 0 \text{ and } |F_e| < F_S \\ F_S \operatorname{sgn}(F_e) & \text{if } v = 0 \text{ and } |F_e| \geq F_S \end{cases} \tag{2.5}$$

The friction force for zero velocity is a function of the external force and not the velocity. The traditional way of depicting friction in block diagrams with velocity as the input and force as the output is therefore incorrect. If doing so, stiction must be expressed as a multi-valued function that can take on any value between the two extremes $-F_S$ and F_S. Specifying stiction in this way leads to non-uniqueness of the solutions to the equations of motion for the system, see Bliman and Sorine (1995).

The classical friction components can be combined in different ways, see Figure 2.10 c), and any such combination will be referred to as a classical model.

The classical models include components that are either linear in velocity or constant. If accurate measurements of friction for steady velocity motion is performed these may reveal different dependencies. Stribeck (1902) observed that for low velocities the friction force is normally decreasing continuously with increasing velocities, not in a discontinuous manner. The phenomenon is termed Stribeck effect and the extra low-velocity friction force above the constant Coulomb level is called Stribeck friction. A more general description of friction than the classical models is, therefore,

$$F = \begin{cases} F(v) & \text{if } v \neq 0 \\ F_e & \text{if } v = 0 \text{ and } |F_e| < F_S \\ F_S \operatorname{sgn}(F_e) & \text{otherwise} \end{cases} \tag{2.6}$$

where $F(v)$ is an arbitrary function, which may look as in Figure 2.10 d). It can be given either as a look-up table or as a parameterized curve that fits experimental data. A number of parameterizations of $F(v)$ have been proposed, see Armstrong-Hélouvry (1991). The most common is of the form

$$F(v) = F_C + (F_S - F_C)e^{-|v/v_S|^{\delta_S}} + F_v v \tag{2.7}$$

where v_S is called the Stribeck velocity. Such models have been used for a long time. Tustin (1947) used the parameterization with $\delta_S = 1$ and Bo and Pavelescu (1982) suggested δ_S in the range of 0.5 to 1. Armstrong-Hélouvry (1991) used a Gaussian parameterization with $\delta_S = 2$ and also proposed to use the sum of two exponentials to match experimental data better. Other parameterizations can of course also be used such as

$$F(v) = F_C + (F_S - F_C)\frac{1}{1 + (v/v_S)^{\delta_S}} + F_v v$$

and a parameterization used by Canudas de Wit (1993)

$$F(v) = F_S - F_d|v|^{1/2}$$

This parameterization has the advantage that it is linear in the parameters but it is only valid in a limited velocity range because of the term $-F_d|v|^{1/2}$, which should account for the Stribeck effect. Outside this interval the friction force may have the wrong sign.

Friction depends on many factors. No particular parameterization can yet be theoretically motivated. Which parameterization to choose depends on the specific application and can be determined from simple experiments.

The main disadvantage when using a model such as (2.6), either in simulations or for control purposes, is the problem of detecting or determining when the velocity is zero. A remedy for this is found in the model described next.

The Karnopp Model: Karnopp proposed the model in Figure 2.11 for simulation purposes, see Karnopp (1985). The model in the figure is drawn for the system

$$m\frac{dv}{dt} = F_e - F \tag{2.8}$$

The Karnopp model was developed to overcome the problems with zero velocity detection and to avoid switching between different state equations

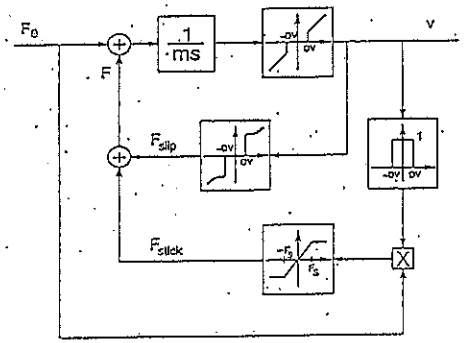

Figure 2.11 Block diagram for the Karnopp model.

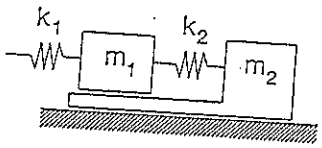

Figure 2.12 Example of a system with two friction interfaces which complicates the use of the Karnopp model.

for sticking and sliding. The model defines a zero velocity interval, $|v| < DV$. For velocities within this interval the internal state of the system (the velocity) may change and be non-zero but the output of the block is maintained at zero by a dead-zone. Depending on if $|v| < DV$ or not, the friction force is either a saturated version of the external force or an arbitrary static function of velocity. The interval $\pm DV$ can be quite coarse and still promote so called stick-slip behavior.

The drawback with the model is that it forms an integrated part with the rest of the system. The external force is an input to the model and this force is not always explicitly given. This is, for example, the case for the system in Figure 2.12. The model equations therefore have to be tailor-made for each configuration. Variations of the Karnopp model are widely used since they allow efficient simulations. The zero velocity interval does, however, not agree with real friction.

The friction models presented so far have considered friction only for steady velocities. No attention is paid to the behavior of friction as the velocity is varied. Experimental investigations have pointed out a number of phenomena that cannot be described by static models. Some of these were described in Section 2.3. Attention has lately been drawn to the dynamics of friction as the demands on control precision have increased and new hardware has enabled implementation of more advanced controllers. Following are a number of models which to a varying extent capture the dynamics of friction. They have all been developed from the 1960s and onwards with increased interest in the 1990s.

Dynamic Friction Models

A dynamic model can be obtained by slightly modifying the static models in the previous section. This is done for Armstrong's seven parameter model. Because of the relation to the static models it is described first of the dynamic friction models.

Armstrong's Seven Parameter Model: To account for some of the observed dynamic friction phenomena a classical model can be modified as in Armstrong-Hélouvry *et al.* (1994). This model introduces temporal dependencies for stiction and Stribeck effect. The modifications, however, do not handle pre-sliding displacement. This is instead done by describing the sticking behavior by a separate equation. Some mechanism must then govern the switching between the model for sticking and the model for sliding. The friction is described by

$$F(x) = \sigma_0 x \tag{2.9}$$

when sticking and by

$$F(v,t) = \left(F_C + F_S(\gamma, t_d)\frac{1}{1+(v(t-\tau_l)/v_S)^2}\right)\operatorname{sgn}(v) + F_v v \tag{2.10}$$

when sliding, where

$$F_S(\gamma, t_d) = F_{S,a} + (F_{S,\infty} - F_{S,a})\frac{t_d}{t_d+\gamma} \tag{2.11}$$

$F_{S,a}$ is the Stribeck friction at the end of the previous sliding period and t_d the dwell time, i.e., the time since becoming stuck. The sliding friction (2.10) is equivalent to a static model where the momentary value of the velocity in the Stribeck friction has been replaced by a delayed version and where it has a time dependent coefficient. As the name states, the model requires seven parameters. These are: the pre-sliding stiffness σ_0; the Coulomb friction F_C; the steady-state stiction force $F_{S,\infty}$; the viscous friction F_v; the Stribeck velocity v_S; the frictional lag τ_l; and finally the parameter determining the temporal behavior of the static friction γ. Since the model consists of two separate models, one for sticking and one for sliding, a logical statement—probably requiring an eighth parameter—determines the switching. Furthermore, the model states have to be initialized appropriately every time a switch occurs.

Another approach, which leads to dynamic friction models, is to use the simplified pictures of the physical contact in Figures 2.1 and 2.3 as the starting point. The result is models that naturally include dynamics and can be called true dynamic friction models.

The Dahl Model: In Dahl (1968) two friction models are introduced. The first model uses as its starting point the relation between strain and stress for shearing processes, as seen in Figure 2.13. The friction interface is modeled as a junction at which shearing takes place. The resulting

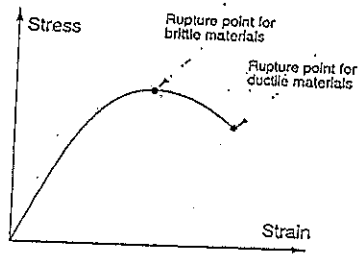

Figure 2.13 Classical stress–strain relation for brittle and ductile materials.

friction force depends on the strain caused by the external force. This corresponds to pre-sliding displacement. If the strain is large enough the junction breaks and the friction force remains constant at the level at which rupture took place. When the external force is removed the result is a permanent deformation. The model includes static friction if the stress–strain characteristic is as typical for ductile materials. It considers only friction due to contact between solids but the friction model for brittle materials is believed to work well both for rolling friction and for sliding friction between lubricated surfaces.

The second model introduced in Dahl (1968) is based on the assumption that the friction force can be described by

$$\frac{dF}{dt} = \frac{dF}{dx}\frac{dx}{dt}$$

This implies that the friction force is only position dependent. This so called rate independence is an important property of the model for analysis especially when studying it as a hysteresis operator, see Bliman (1992). Results from simulations were compared with experimental results for ball bearings. A good agreement between the model and true friction was observed. The model exhibits hysteresis between velocity and friction force. The hysteresis depends on the rate of change of the velocity. Dahl also noted that the model is a generalization of ordinary Coulomb friction.

The second model is further studied in Dahl (1975) and Dahl (1976) where it is used to describe frictional damping of a wire pendulum. The damping of the pendulum is due to the bending of the wire and thus not from relative motion between surfaces. The model is, however, applied

even to this internal friction. The position dependency of the friction force is further explored and Dahl proposes the relation

$$\frac{dF}{dx} = \sigma_0 \left| 1 - \frac{F}{F_C}\operatorname{sgn}\left(\frac{dx}{dt}\right)\right|^i \operatorname{sgn}\left(1 - \frac{F}{F_C}\operatorname{sgn}\left(\frac{dx}{dt}\right)\right) \qquad (2.12)$$

where σ_0 is the stiffness and the exponent i a model parameter. The second factor is present to stabilize the differential equation for simulation purposes. The model (2.12) is the general Dahl model. Using the friction model and the linearized pendulum equations the damping ratio and energy dissipation of the pendulum are determined. It is concluded that for large amplitudes of oscillation the model resembles Coulomb friction, but as the amplitude decays the hysteresis and dynamics become more and more important for the damping.

The Dahl model (2.12) leads to a friction displacement relation that bears much resemblance with stress–strain relations proposed in classical solid mechanics, see Ramberg and Osgood (1943) and Sargin (1971).

In Dahl (1977) the model is used when experimentally studying friction in ball bearings. A force–deflection test is done on a ball bearing and parameters of the model are fitted to agree with the experiments. Three types of fittings are done. In the first, the exponent i in (2.12) is estimated together with $x_c = F_C/\sigma_0$ while F_C is determined as the friction force asymptote for unidirectional motion. In the other two fittings the exponent i is fixed to 1 and 2, respectively. For the first type of fitting i is estimated to 1.5 quite consistently. It is observed that the rest stiffness σ_0 is important for the fit. The cases with fixed exponents do not agree with the measured data over an as large friction force range.

When referred to in the literature the Dahl model is often simplified, using $i = 1$, to

$$\frac{dF}{dt} = \sigma_0 \left(1 - \frac{F}{F_C}\operatorname{sgn}(dx/dt)\right)\frac{dx}{dt} \qquad (2.13)$$

The Dahl model has been used for adaptive friction compensation (see Walrath (1984) and Ehrich Leonard and Krishnaprasad (1992)) with improved performance as the result.

When summarizing Dahl's papers two sentences from Dahl (1968) are worth citing, firstly:

> "The origin of friction is in quasi static bonds that are continuously formed and subsequently broken."

and, secondly:

> "The resulting functions behave as a brush whose bristles must bend as the brush moves in one direction and then flop or bend in the opposite direction if the motion is reversed."

This view of the origins of friction has been the starting point for several models and is, as we will see, also the inspiration for a new model in Chapter 3.

The Bristle Model: In Haessig and Friedland (1991) a friction model is introduced, which also is based on the microscopical contact points between two surfaces. Due to irregularities in the surfaces the number of contact points and their location are random. Each point of contact is thought of as a bond between flexible bristles. As the surfaces move relative to each other the strain in the bond increases and the bristles act as springs giving rise to a friction force. The force is given by

$$F = \sum_{i=1}^{N} \sigma_0 (x_i - b_i) \tag{2.14}$$

where N is the number of bristles, σ_0 the stiffness of the bristles, x_i the relative position of the bristles, and b_i the location where the bond was formed. As $|x_i - b_i|$ becomes equal to δ_s the bond snaps and a new one is formed at a random location relative to the previous location. The new location is determined by

$$b_i^{k+1} = b_i^k + \text{uniform}(\Delta)\, \text{sgn}(x_i - b_i^k)$$

where Δ determines the uniform distribution. A suitable number of bristles is 20–25 according to the authors but even a single bristle will work, although giving rise to larger variations in the friction force. The stiffness of the bristles, σ_0, can be made velocity dependent. The model is inefficient in simulations due to the many actions that has to be taken. The deflection of each bristle has to be checked to see if it has snapped, and if so, a new location has to be determined. The resulting behavior when sticking may be oscillatory since no damping of the bristles is present in the model.

The Reset Integrator Model: A second model also proposed in Haessig and Friedland (1991) is the reset integrator model. This model is similar to the bristle model and can be seen as representing a single bond. Instead of snapping the strain in the bond is kept constant by a

logical statement, which shuts off the strain increase at the point of rupture. The model utilizes an extra state to determine the strain in the bond, which is modeled by

$$\frac{dz}{dt} = \begin{cases} 0 & \text{if } (v > 0 \text{ and } z \geq z_0) \text{ or } (v < 0 \text{ and } z \leq -z_0) \\ v & \text{otherwise} \end{cases}$$

The friction force is given by

$$F = (1 + \alpha(z))z\sigma_0(v) + \sigma_1 \frac{dz}{dt} \tag{2.15}$$

where $\sigma_1 dz/dt$ is a damping term that is active only when sticking. The damping coefficient can be chosen to give a desired relative damping of the resulting spring-mass-damper system. Stiction is achieved by the function $\alpha(z)$, which is given by

$$\alpha(z) = \begin{cases} a & \text{if } |z| < z_0 \\ 0 & \text{otherwise} \end{cases}$$

If $|z| < z_0$ the model describes sticking where the friction force is a function of z. As the artificial deflection reaches the maximum value z_0 the variable z remains constant and the friction force drops since $\alpha(z)$ becomes zero. The friction force when slipping is an arbitrary function of the velocity given by $\sigma_0(v)$. The reset integrator model is far more efficient to simulate than the bristle model but, still, a test is necessary to check whether $|z| > z_0$ or not.

The Models by Bliman and Sorine: Bliman and Sorine have developed a family of dynamic models in a series of papers: Bliman and Sorine (1991), (1993) and (1995). The starting point for the models is the experimental investigations by Rabinowicz in the 1950s, see Rabinowicz (1951). Results as in Figure 2.4 were obtained when studying the break-away behavior. It is seen in Figure 2.4 that the maximum friction force is reached after a small distance. After some further motion the Coulomb friction level has been reached. Bliman and Sorine stresses velocity independence, i.e., that the curve will be the same no matter how fast the distances are covered. The rate independence makes it possible to express the models with distance instead of time as the independent variable. They hence replace the time variable t by a space variable s through the transformation

$$s = \int_0^t |v(\tau)| d\tau$$

The models are now expressed as

$$\begin{aligned} \frac{dZ_s}{ds} &= AZ_s + Bv_s \\ F &= CZ_s \end{aligned} \tag{2.16}$$

where $v_s = \text{sgn}(v)$. The models are linear systems when expressed in the new space variable. A first-order model is given by

$$A = -\sigma_0/F_C, \quad B = \sigma_0/F_C \quad \text{and} \quad C = F_C. \tag{2.17}$$

Expressed with time as the independent variable this becomes

$$\frac{dF}{dt} = \sigma_0 \left(v - |v| \frac{F}{F_C} \right)$$

which coincides with the Dahl model, see (2.13). The first-order model does not yield a friction peak as in Figure 2.4. This can, however, be achieved by a second-order model with

$$\begin{aligned} &A = \begin{pmatrix} -\sigma_0/(\eta(F_C + \Delta F)) & 0 \\ 0 & -\sigma_0/\Delta F \end{pmatrix} \quad B = \begin{pmatrix} \sigma_0/\eta(F_C + \Delta F) \\ -\sigma_0/\Delta F \end{pmatrix} \\ &\text{and} \quad C = \begin{pmatrix} F_C + \Delta F & \Delta F \end{pmatrix} \end{aligned} \tag{2.18}$$

and with $0 < \eta < \Delta F/(F_C + \Delta F)$, see Bliman and Sorine (1995). The second-order model actually consists of two first-order models in parallel. The fastest model has the highest steady-state friction. The friction force from the slower model is subtracted from the faster model, which gives the resulting friction force. Both the first- and second-order models can be shown to be dissipative. The velocity or rate independence also makes them attractive, since hysteresis theory can be applied. Bliman and Sorine also show that, as σ_0 goes to infinity, the models behave as a classical Coulomb friction model (first-order model) or as a classical model with Coulomb friction and stiction (second-order model). It should be noted that the Stribeck effect of the second-order model claimed by the authors is not the same as observed in Stribeck (1902). The emulated effect by the second-order model is only present at a certain distance after motion starts. This means that it will not appear when slowing down, as the true Stribeck effect would. The friction peak is instead the equivalent of stiction for a dynamic model.

The models described so far have all been applicable to the general contact shown in Figure 2.9. They have also, without exceptions, been considered for control purposes. Advanced friction models have, of course, also been derived outside the control area. These models often have different purposes and aim to explain, for example, phenomena for idealized contacts, dynamics of lubricants, and even earthquakes. Some of these modeling efforts are now shortly described in order to highlight the various aspects of friction modeling.

Special Purpose Models

The special purpose models we describe include a category of models that describes contact forces using continuum mechanics. Another is based on the hydrodynamics of lubricated contacts. There are also special purpose models for road–tire friction and rock mechanics.

Continuum Mechanics: The relation between stress and strain for various materials is treated in classical solid mechanics. If friction is seen as a one-dimensional shearing process such relations could form the basis for friction models. This was the case for the Dahl model (2.12). Another approach to friction modeling is to describe the friction interface using continuum mechanics. This is done in Oden and Martins (1985). The method combines a simple friction relation, which holds locally, with a complicated model of the contact. In particular, the model includes motion in the normal direction of the interface.

A typical friction interface is shown in Figure 2.14. Empirically the local normal stress can be described as a power function of the compression of the surface, i.e.,

$$\sigma_N = c_N a^\gamma$$

where $a = t_0 - t$. The friction force per area unit is then given by the

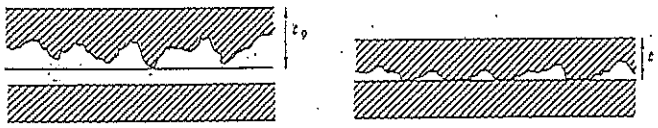

Figure 2.14 Microscopical picture of the friction interface. The asperities deform as contact is made. The compression is given by $t - t_0$. For simplicity the lower surface is shown as being rigid.

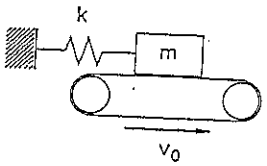

Figure 2.15 Set-up of the example in Oden and Martins (1985).

classical relation

$$\sigma_F = \mu \sigma_N \operatorname{sgn}(v)$$

The total friction force is achieved by integrating over the contact area. The continuum problem can be formulated using variational methods and be solved by, for example, finite element methods. The variational formulation requires a regularization of the sign function. The approach should be compared with the combination of an elaborate friction model with a simple model of the interface. The difference is illustrated with a simple example taken from Oden and Martins (1985).

The set-up of the example is shown in Figure 2.15. Assuming motion with one degree of freedom the system can be described by

$$\frac{d^2x}{dt^2} = F - kx$$

where F is the friction force and x the horizontal position. To account for various friction phenomena an elaborate friction model may be required, which introduces additional states. The continuum mechanics approach is instead to model the deformation between the block and the belt. With the block assumed rigid we get the situation shown in Figure 2.16. Three degrees of freedom are necessary for the two dimensional motion. With the coordinates introduced in Figure 2.16 the system is described by

$$M\frac{d^2\xi}{dt^2} + K\xi + P(\xi) + J\left(\xi, \frac{d\xi}{dt}\right) = F_e$$

where

$$\xi = \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} \quad M = \begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad K = \begin{pmatrix} k & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad F_e = \begin{pmatrix} 0 \\ mg \\ 0 \end{pmatrix}$$

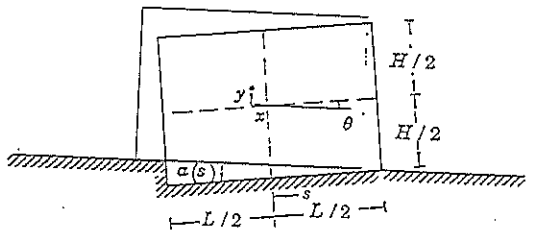

Figure 2.16 Deformation model for the contact between the rigid block and the belt.

and

$$P = \begin{pmatrix} 0 \\ P_y \\ P_\theta \end{pmatrix} = Bc_N \int_{-L/2}^{L/2} (y - s\theta)^\gamma(s) \begin{pmatrix} 0 \\ 1 \\ -s \end{pmatrix} ds$$

$$J = \begin{pmatrix} J_x \\ 0 \\ J_\theta \end{pmatrix} = B\mu c_N \operatorname{sgn}\left(\frac{dx}{dt} + \frac{H}{2}\frac{d\theta}{dt} - v_0\right) \begin{pmatrix} 1 \\ 0 \\ H/2 \end{pmatrix} \int_{-L/2}^{L/2} (y - s\theta)^\gamma(s) ds$$

where B is the breadth of the block. The term $P(\xi)$ includes the force and torque caused by the deformation of the belt. The friction force is given by the term $J(\xi, d\xi/dt)$. The system description is of order six.

Lubricated Contacts: For most engineering applications lubrication is present in the friction interface. Friction models have therefore been derived using hydrodynamics. A simple example is viscous friction as given by (2.3). Other models also exist. Hess and Soom (1990) used the following model to describe the friction coefficient for steady sliding:

$$\mu = \frac{\mu_b}{1 + c_1(\eta v/\sqrt{WE})^2} + c_2\frac{\eta v L}{W} \qquad (2.19)$$

where μ_b is the coefficient of friction in the boundary lubrication regime, η the viscosity coefficient, W the normal load, E Young's modulus, and L the contact length. The second term is equivalent to linear viscous friction and the first term accounts for friction due to boundary lubrication. The model thus covers all lubrication regimes except sticking. Dynamics

are introduced by substituting the velocity in (2.19) with $v(t-\Delta t)$. Comparisons with experimental data show that a constant time lag gives good agreement.

In Harnoy and Friedland (1993) a model based on the hydrodynamics of a lubricated journal bearing is introduced. The model covers both the boundary, mixed and hydrodynamic lubrication regimes. It stresses the dynamics of the friction force. The eccentricity ε of the bearing is an important variable in determining the friction force which, with some simplification, is given by

$$F = K_1(\varepsilon - \varepsilon_{tr})^2\Delta + \frac{K_2}{\sqrt{1-\varepsilon^2}}v. \tag{2.20}$$

The first term is due to the shearing of asperity contacts and the second term is due to the viscosity of the lubricant. The function Δ is an indicator function that is one for $\varepsilon > \varepsilon_{tr}$ and zero otherwise. This implies that for small eccentricities there is no friction due to asperity contacts. The eccentricity is given by a fourth-order differential equation, which determines the pressure distribution in the lubricant. The model requires five parameters. Simulations show a behavior very similar to the observations in Hess and Soom (1990). An extension including sleeve compliance is given in Harnoy and Friedland (1994). The model then becomes more complicated and requires determination of initial values when switching between slipping and sticking. The paper shows that a low sleeve compliance may be advantageous in precise motion control.

Rock Mechanics: The desire to predict earthquakes has inspired research that aims to model the friction between the crustal plates of Earth. This is described in Dieterich (1972). The stability of tectonic sliding has been analyzed using a special friction model in Ruina (1983). The model has also been used in connection with control, see Dupont (1994).

The "Magic Formula Tire Model": When simulating the behavior of road vehicles it is important to know the contact force between the road surface and the tires of the vehicle, see Bakker *et al.* (1987) and Pacejka (1991). This is also of considerable interest when designing antilock braking systems, compare with Example 1.1 in Chapter 1. Since the wheels rotate, we get a slightly unfamiliar way of describing friction. A simplified picture of a tire on a surface is shown in Figure 2.17. The rubber of the tire is elastic and therefore the tread deforms as a force is transmitted over the contact. The deformation starts as the tread comes into contact with the ground and ceases as it loses contact. This results in

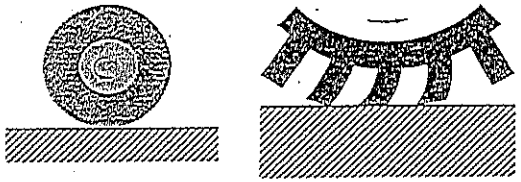

Figure 2.17 A simplified picture of a road–tire contact. The tread deflects when a force is transferred by the contact between tire and road surface.

a reduction in the effective circumference of the tire, which implies that the ground velocity will not equal $v = \omega r$, where ω is the rotational speed of the wheel and r the wheel radius. Instead we have an effective radius r_e. The slip rate is defined as

$$s = \frac{\omega(r - r_e)}{\omega r}$$

The contact force is modeled as a function of the slip. The slip can be measured and hence it is possible to know how much contact force that can be transferred. The relation depends, of course, on the tire and the type of road surface. A characteristic shape of the relation is shown in Figure 2.18 which is closely related to Figure 2.4. The behavior can be explained using Figure 2.17. Small values of the slip corresponds to a small deflection or displacement of the tread. Larger values implies that there is also gross sliding. This is the same behavior as noted in Figure 2.4. In Bakker *et al.* (1987) curves of the form

$$F(s) = D\sin(C\arctan(Bs - E(Bs - \arctan(Bs)))) \tag{2.21}$$

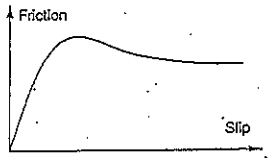

Figure 2.18 Typical relation between slip rate and contact force for a tire.

are suggested to fit experimental data. Equation (2.21) is called the "Magic Formula Tire Model".

In Bliman *et al.* (1995) the authors propose to use the second-order model by Bliman and Sorine, i.e., (2.16) and (2.18), to model the friction between tire and road. They model the contact area as a rectangle and let ξ denote the longitudinal position along the contact area. The friction is then model by means of $Z(\xi)$, which is the friction state along the contact. The model becomes

$$\frac{d}{dt}Z(\xi) = \frac{dZ}{d\xi}\frac{d\xi}{dt} = AZ|\omega(r-r_e)| + B(\omega(r-r_e))$$
$$F = C\int_0^L Z(\xi)d\xi$$

where L is the length of the contact zone. Since the wheel is rotating, it holds that $d\xi/dt = \omega r$ and the model equations can then be written as

$$\frac{dZ}{d\xi} = AZ|s| + Bs$$
$$F = C\int_0^L Z(\xi)d\xi$$

For constant slip s we can integrate along the contact area to achieve the friction force. The first-order model (2.17) gives a monotonously increasing friction–slip relation, but for the second-order model (2.18) with

$$C = \begin{pmatrix} (F_C + \Delta F)/L & \Delta F/L \end{pmatrix}$$

we get

$$F(s) = (F_C + \Delta F)\left[1 - \frac{\eta(F_C+\Delta F)}{\sigma_0 Ls}\left(1 - e^{-(\sigma_0 Ls/\eta(F_C+\Delta F))}\right)\right] - \Delta F\left[1 - \frac{F_C}{\sigma_0 Ls}\left(1 - e^{-(\sigma_0 Ls/F_C)}\right)\right] \quad (2.22)$$

for $s > 0$. This function can model the desired behavior and fit experimental data as shown in Bliman *et al.* (1995). The parameters also have a physical interpretation, which they lack in (2.21).

Simulation Efficiency and Accuracy

One of the purposes with a friction model is to use it in simulations. It is necessary that these simulations can be performed efficiently. The reviewed friction models show various difficulties when simulated.

Discontinuities are notoriously difficult to simulate, since it is necessary to find the time where the events of discontinuity occur, see Hairer *et al.* (1987). Simulation software must have a structured way of finding zero crossings, otherwise the step size of the integration routine must be decreased. The problem of detecting zero velocity is present for all the static models except for the Karnopp model. The Karnopp model avoids this problem by defining an interval as zero velocity. By doing so it is not necessary to detect exactly when the interval is entered. A coarse estimate of the first time inside the interval suffices, which implies that a larger step size can be used without event handling. The problem of zero velocity detection occurs also for some of the dynamic models such as Armstrong's seven parameter model, which further uses a time delay. This requires that the history of the velocity is stored.

Some models consists of two sub-models, one for sticking and one for sliding. It is then necessary to have a mechanism that determines when to switch between the sub-models. It is also necessary to initialize the models appropriately after a switch. This holds for Armstrong's seven parameter model and for the bristle model.

The Dahl model, the models by Bliman and Sorine, and the reset integrator model are all efficient to simulate. The differential equations, however, may be stiff depending on the choice of parameters. A stiff solver may therefore be advantageous for the simulations.

Some of the models derived from continuum mechanics and hydrodynamics include partial differential equations that require special solvers.

2.5 Summary

This chapter has given a short introduction to the mechanisms behind friction and to some of the many experimental observations of friction. The observations show that friction exhibits a very rich and dynamic behavior.

The main emphasis in the chapter lies on friction modeling. A small survey of existing friction models for various purposes is presented. The two main categories include friction models that all have been considered for control purposes. The models are categorized depending on if they include dynamics or not. The emphasis is made on a number of dynamic

	1	2	3	4	5	6	7
Arbitrary Steady-State Char.	yes	yes	yes	–	yes	yes	–
Stiction	yes	yes	yes	–	yes	yes	yes
Varying Break-Away Force	–	–	yes	–	–	–	–
Pre-Sliding Displacement	–	–	yes	yes	yes	yes	yes
Frictional Lag	–	–	yes	–	–	–	–
Simulation Efficiency	–	yes	–	yes	–	yes	yes

Table 2.1 A summary of the friction characteristics for the reviewed models: Classical models (1), Karnopp model (2), Armstrong's seven parameter model (3), Dahl model (4), Bristle model (5), reset integrator model (6), and models by Bliman & Sorine (7).

friction models that have been proposed for control purposes since the end of the 1960s. A third category describes special purpose models from a wide range of application fields. Many connections between the different models are pointed out.

Table 2.1 gives a summary of the different characteristics of the models. For the special purpose models most of the characterizations are not applicable. These models have therefore been excluded from the table.