*Int. J. Solids Structures* Vol. 20, No. 7, pp. 637-647, 1984
Printed in the U.S.A.

0020-7683/84 $3.00 + .00


# A THEORY OF FRICTION

A. CURNIER
Research Engineer, Laboratoire de Mécanique Appliquée, Ecole Polytechnique Fédérale de Lausanne, Switzerland

(*Received* 6 *April* 1983; *in revised form* 26 *September* 1983)

**Abstract**—A rather general theory of friction, inspired from the classical theory of plasticity is proposed. It includes the contact impenetrability condition as a by-product.

## INTRODUCTION

In spite of the resemblance which exists between plastic and frictional phenomena, the degrees of development reached by the theories of plasticity and friction are rather disparate. Indeed, although plasticity has long received a rigorous and general framework in the name of the classical theory of plasticity (e.g. [1–3]) encompassing a wide variety of material behaviour, the theory of friction has engaged very episodic attention[4–6] remaining almost exclusively limited to Coulomb's law of perfect friction which only covers a restricted range of tribological situations. Moreover, the study of plasticity at finite strain is already well advanced (e.g. [7–9]) whereas the kinematics of friction are still restricted either to small amounts of slip or else to steady state sliding along planar or cylindrical surfaces. Finally this lag of friction theory is even more flagrant when considering the progresses accomplished during the last decade with regards to the formulation of contact problems by means of variational inequalities[10–12], calling for more sophisticated laws of friction.†

This paper is a modest attempt to reduce the gap accumulated between the two disciplines by proposing a rather general theory of friction (yet limited to small amounts of slip) inspired from the theory of plasticity (at small strains) along a line originally explored by [4] and further pursued by [5].

By construction this theory is compatible with the principles of continuum mechanics and susceptible to include not only the influence of the normal load on the friction force but also other factors such as wear, adhesion and heat, although the latter will not be explicitly treated in this presentation. However it is restricted either to moderate amounts of slip or to nominally flat contact surfaces.

At the beginning of the paper, a variational formulation of contact problems between deformable bodies provides a concise introduction to both the kinematic and the static variables to be related by the constitutive law of friction under way. Then a theory of friction which includes the contact impenetrability condition as a by-product, is proposed and discussed at length with systematic reference to the theory of plasticity and numerous comments about its physical interpretation. Within the context defined by this theory, the construction of a specific law of friction reduces to the choice of a slip criterion and a wear law. A few possibilities are mentioned in the course of the discussion to illustrate the potential of the theory.

The (open) questions of existence and uniqueness of a solution, (e.g. [10]), to contact problems resorting to this theory of friction are out of the scope of this paper.

## 1. BRIEF PRESENTATION OF CONTACT MECHANICS

A variational formulation of the contact between two deformable bodies constitutes a good basis for the development of a constitutive law of friction including an impenetrability condition.

†If of any relevance here, it is fair to recognize that friction involves the interaction of two materials (not to mention the presence of oxydation films) whereas plasticity is limited to the behavior of one material. However this duplication of the constituents is largely compensated by the ease of observation and therefore of interpretation of a surface process as opposed to an interior one.

To this end consider two bodies, one called the striker and the other the target, bound to contact one another within a surface $A$ characterized by the unit outward normal $N$ to the target (Fig. 1). In principle the roles of the striker and target are interchangeable; however, in practice it is preferable to take the striker to be the most convex of the two bodies (and therefore the target as the flatest and stiffest one) in the presumed area of contact $A$, whenever applicable and possible.

Basically, to the standard weak statements of equilibrium† of the two deformable bodies, needs to be added a contact term of the form[13]:

$$\int_A F \cdot \dot{D}\, \mathrm{d}A = \int_A (F_N \cdot \dot{D}_N + F_T \cdot \dot{D}_T)\, \mathrm{d}A \quad (\geqslant 0). \tag{1}$$

In the above the vector $D$ represents the *distance* of contact (or gap) separating each point $S$ on the striker from a point $T$ on the target, i.e. $D = S - T$, which are (were or will be) in contact with one another (depending on the instant). A superposed dot denotes a rate or an increment (or even a variation) depending on the temporal (or functional) context. It proves convenient to resolve the relative velocity $\dot{D}$ into a normal and a tangential components $\dot{D}_N = (\dot{D} \cdot N)N$ and $\dot{D}_T = \dot{D} - \dot{D}_N$. In general this decomposition is not applicable to the contact distance $D$ itself due to the variations in the definition of the unit normal $N$ with the curvature and deformation of the contact surface. There exist two exceptions when it remains valid, namely for small amounts of slip and nearly flat and rigid contact surfaces.

The vector $F$ is the *force* of contact per unit area (also called stress or traction vector) acting on the target at the point $S = T$ whenever contact occurs, i.e. when $D_N = \dot{D}_N = 0$. $F_N = (F \cdot N)N$ and $F_T = F - F_N$ are the normal and tangential components of the contact force.

It is found convenient to attach the notion of point of contact with a material point $S$ on the striker and to regard the target surface as a sliding rink oriented by its outward normal $N$. With this convention it becomes possible to talk about the sliding trajectory of a contact point, defined as the curve described by one striker point on the target surface, without ambiguity. The actual surface of contact $C$ defined by the set of striker points in contact with the target at a given time, must be distinguished from the potential (or candidate) surface of contact $A$ which includes the actual one at all times. This biased definition of contact offers the advantages to lift the ambiguity of ubiquitous points inherent to all surfaces of material discontinuity and to avoid the need for a transversality

†For examples the principle of virtual work.

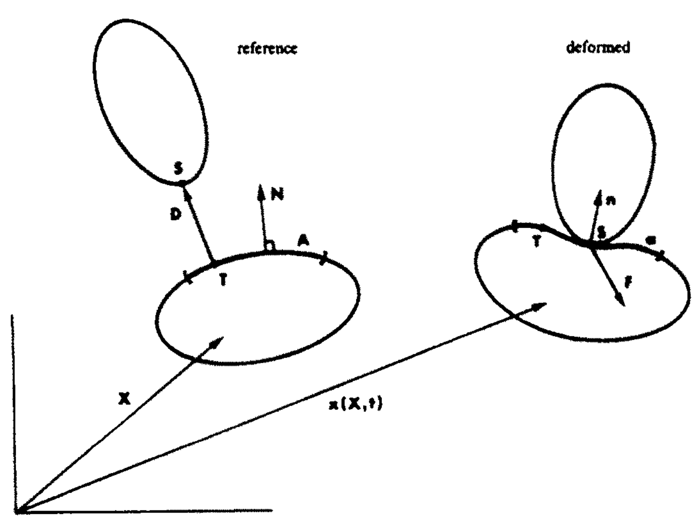


Fig. 1. Contact of two deformable solids with friction.

condition to account for the extension and recession of the contact surface with time.†

The contact integral (1) may be indifferently defined over the *reference* contact surface $A$ with unit outward normal $N$ (in material Lagrangian description), $F$ representing the first Piola–Kirchoff stress vector, or over the *deformed* contact surface $a$ with normal $n$ (in spatial Eulerian description), $F$ becoming the Cauchy stress vector. Because the amounts of slip and deformation are assumed to remain small in this presentation, the two descriptions are undistinguishable.

The contribution of the contact constraint term (1) is non-negative.

Indeed it is null whenever a positive normal gap exists ($D_N > 0$, $\dot{D}_N \neq 0$, no contact) since then $F = F_N = F_T = 0$. When the normal gap closes ($D_N = 0$, $\dot{D}_N = 0$, contact), the normal contribution $F_N \cdot \dot{D}_N$ remains zero, which accounts well for the condition of impenetrability, but the tangential term $F_T \cdot \dot{D}_T$, which represents the energy dissipated by friction, must only be non-negative according to the entropy inequality. Of course penetration ($D_N < 0$, $\dot{D}_N \neq 0$) is implicitly excluded whereas tension ($F_N > 0$, $\dot{D}_N = 0$) due to adhesion may be retained.

The dissipation due to friction vanishes in the two limiting cases of perfect stick ($\dot{D}_T = 0$) and perfect slip ($F_T = 0$). In between those extremes, a *law of friction* relating the force of friction to the amount of slip (and other internal variables to be introduced later) of the form $F_T = \hat{F}_T[D_T, \dot{D}_T, \ldots]$ is necessary. It is the purpose of this paper to propose such a law embedding an impenetrability condition of the form $F_N = \hat{F}_N[D_N, \dot{D}_N, \ldots]$ for completeness.

Enforcement of the principle of *action and reaction* across the surface of contact is achieved by taking a variation of the contact distance or velocity $\check{D}$, i.e. separate variations of $\check{S}$ and $\check{T}$, like in a classical displacement formulation. To the contrary kinematic *compatibility* is enforced by taking an independent variation of the contact force $\check{F}$ according to the force method. More explicitly the contact term (1) can be expanded into the sum of two terms

$$\int_A (F \cdot \check{D} + \check{F} \cdot \dot{D})\,\mathrm{d}A \tag{2}$$

where a "vee" denotes a variation and overrides the "dot" whenever superposed. The formulation (1) is adapted to the implementation of penalty methods whereas the corollary (2) is suitable for a treatment by Lagrange multipliers.‡ The penalty approach is prefered in the sequel of this paper.

## 2. RATE INDEPENDENT THEORY OF FRICTION WITH IMPENETRABILITY

Relying on basic experimental observation[14], the theory of friction proposed here is independent of the slip rate, uses a standard slip rule for all internal variables and presents high resistance to penetration.

By analogy with plasticity, the theory rests upon four basic principles.

### 2.1 *Decomposition of the contact distance into adherence and slip* (*cf. decomposition of the strain into elastic and plastic parts*)

The theory is based upon the decomposition of the distance of contact at a point $D$ into the sum of two parts (Fig. 3a): one reversible, rather unusual, called *adherence* and denoted $D^A$ and the other irreversible, more familiar, called *slip* and denoted $D^S$, which

†A more symmetric treatment of contact may be easily obtained by alternating the roles of the striker and target over complementary subsets of the potential contact surfaces. Schematically the contact term (1) could be decomposed into

$$\int_{A^S} F \cdot \dot{D}\,\mathrm{d}A + \int_{A^T} F \cdot \dot{D}\,\mathrm{d}A \quad \text{with} \quad A^S \cup A^T = A \quad \text{and} \quad A^S \cap A^T = \emptyset\,.$$

‡Recent formulations resorting to "penalty augmented Lagrangians" provide a unified treatment of these alternative approaches.

can be resolved into normal and tangential components as before

$$
\begin{aligned}
D &= D^A + D^S && \text{(a)} \\
&= (D_N{}^A + D_N{}^S) + (D_T{}^A + D_T{}^S) && \text{(b)} \\
&= D_N + D_T. && \text{(c)}
\end{aligned}
\tag{3}
$$

This double decomposition is implicitly based upon two assumptions regarding the contact distance definition and the contact normal variation respectively (Fig. 2). First, for the decomposition (3a) into adherence and slip to remain meaningful (i.e. to keep track of the sliding distance covered), it is essential to conserve the same origin $T$ (defined as the initial point of contact of the striker point $S$ on the target surface $A$) for all subsequent measurements of the distance of contact attached to the striker point $S$, since by definition $D = S - T$. Next, for the resolution (3c) into normal and tangential components to remain meaningful (i.e. to obtain accurate measures of the normal gap and tangential slip), it is imperative that the direction of the outward normal $N$ to the target surface remains nearly constant throughout the sliding process. This will be the case if either the amounts of slip are small or else if the target surface remains nominally flat. These two assumptions are two strong limitations of the present theory. A radical remedy to relax these assumptions is to use the contact velocity $\dot{D}$ instead of the contact distance $D$ in the kinematic decomposition (3). Considering the introductory nature of this paper this substitution will not be attempted here to avoid some unnecessary complications.

The decompositions (3) of the contact distance may be compared to the strain decompositions into elastic and plastic parts on one hand and into bulk and deviatoric components on the other hand used in plasticity:

$$E = E^E + E^P = (\bar{E}^E + \bar{E}^P) + (E'^E + E'^P) = \bar{E} + E'$$

with an obvious notation. Note that these decompositions are subject to caution at finite strains[7] and it is sometimes advocated to work with strain rates instead[8, 9].

A physical interpretation of (3) is the following. The examination with a microscope of a material surface, however polished it may appear to the eye or feel to the touch, shows a rough profile. Consequently the real area of contact between two nominally smooth surfaces occurs in fact at the top of asperities in regard[14].

The relative displacement $D^A$ called adherence can be attributed to the elastic deformations of the asperities of the two bodies in contact. To the contrary, the slip term $D^S$ may be imputed partly to the plastic deformations of these asperities and mainly to the rupture of the junctions occuring at their tips. Although the micro-shifts due to adherence are negligible in comparison to the macro-slips due to sliding, their existence rests upon serious experimental evidence[4]. Moreover their introduction will prove very convenient for the formulation.

The terms "adherence" and "slip" must be taken in an enlarged acceptation of these words, since they include not only the tangential relative displacements as usual but also the normal ones. This double connotation prepares the simultaneous treatment of both the normal impenetrability condition and the tangential slip condition which characterizes this theory.

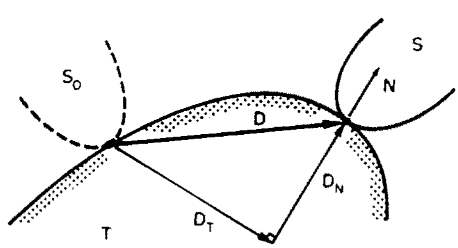


Fig. 2. Limitations of the contact distance decomposition.

To complete this kinematic description of a frictional contact, it remains to introduce the *cumulated* slip $D^c$, a scalar defined as the integral of the magnitude of the directional slip $D_T^S$ over the sliding process

$$D^c = \int_0^t \mathrm{d}D^c \quad \text{where } \mathrm{d}D^c = \sqrt{\mathrm{d}D_T^S \cdot \mathrm{d}D_T^S}. \tag{4}$$

It is the direct analog of the equivalent plastic strain $E^c$ and its role will become clear in the sequel.

The directional slip $D^S$ and the cumulated slip $D^c$ are the two internal variables (one vector and one scalar) used to take into account the permanent memory characteristic of the phenomenon of friction.

### 2.2 *Laws of adherence, tear and wear (cf. elastic, kinematic- and isotropic-softening laws)*

To the kinematic variables of adherence $D^A$, directional slip $D^S$ and cumulated slip $D^c$ are associated, by energetic duality, three dynamic quantities which are called the force of *friction F*, the force of "*tear*" $F^S$ and the force of "*wear*" $F^c$. The forces of friction and tear are vectors whereas the (generalized) force of wear is a mere scalar. Strictly speaking, the term "force of friction" should be reserved for the tangential component $F_T$ of the contact force $F$ (a similar remark applying to the force of tear). However it is found expedient to refer to both the resultant and the tangential components of these vector forces by a single generic name to avoid a shower of new terms. Also, since these forces act per unit area, it is recalled that more rigorous but less attractive terms would be surface tractions or stress vectors.

The forces of tear and wear are introduced to characterize two different forms of a single and same phenomenon: the running-in (i.e. the grinding-in) of contact surfaces in relative sliding motion. The first form occurs when the sliding motion of the two bodies is monotone and oriented along some preferential direction, resulting into an *anisotropic tear* of the contact surfaces which requires a vector entity for its modelization: the tear force. The second form occurs when the two bodies rub against one another in alternate arbitrary directions, resulting into an *isotropic wear* of the surfaces, sufficiently well described by a scalar quantity: the wear force. Both processes produce a reduction of the force of friction. In a monodirectional experiment, the forces of tear and wear may be interpreted as the drops in the force of friction due to directional and cumulated slips respectively (Fig. 3b).

The forces of friction, tear and wear are the analogs of the stress, kinematic stress and isotropic (or equivalent) stress, respectively, in plasticity.

In order to define these forces, three constitutive laws are needed. For the sake of

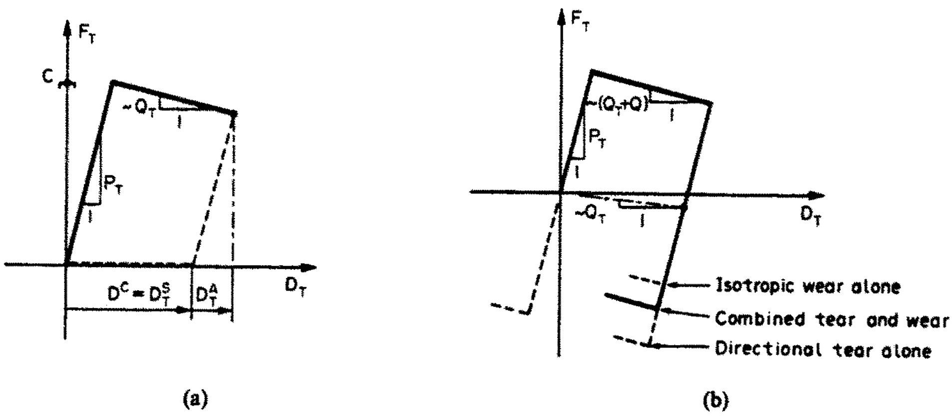


Fig. 3. Kinematic decomposition and friction laws, (a) distance = adherence + slip, (b) combined tear and wear laws.

simplicity three linear laws are proposed here but any reversible relationships would be perfectly legitimate:

$$F = \mathbf{P}D^A \quad \text{(a)}$$
$$F^S = \mathbf{Q}D^S \quad \text{(b)} \qquad (5)$$
$$F^c = QD^c \quad \text{(c)}$$

where **P** is a "penalty matrix" representing the elasticity of the asperities, **Q** is a "rugosity matrix" which characterizes the directional tear of asperity tips and junctions and $Q$ a "rugosity modulus" playing a similar role for the cumulated wear.†

These three constitutive laws of adherence, tear and wear are the analogs of the elastic, kinematic and isotropic softening laws of plasticity

$$S = \mathbf{E}E^E, \quad S^p = \mathbf{H}E^p, \quad S^c = HE^c$$

where $S$ is the stress, $S^p$ the plastic softening stress, $E^c$ and $S^c$ the equivalent plastic strain and stress, **E**, **H** and $H$ the various associated moduli.

In general the matrices **P** and **Q**, best partitioned into the normal-tangential directions for an easier interpretation, will be full. In practice however, they may be significantly simplified according to the following arguments.

First, since the micro-shifts $D^A$ are generally negligible in comparison to the macro-slips $D^S$ as already mentioned, the asperity stiffness may be taken arbitrarily large (thus the name of penalty) without any serious consequence on the resolution of the theory. Indeed as these penalties tend to infinity, the conditions of perfect impenetrability and adherence, characteristic of ideally polished surfaces in contact, may be asymptotically approached.

Moreover the matrices **P** and **Q** may be assumed diagonal, which amounts to neglect stiffness and tear coupling in the diverse directions. For instance, in the absence of impacts, it is reasonable to neglect hammering effects on fiiction. In the tangential plane it makes sense to suppose that running in the asperities in one direction may create a grip in the negative direction‡ but not in one (and only one) transverse direction. This is a classical assumption in plasticity.

If in addition the rugosity is isotropic, then the tangential coefficients must be equal. Finally if normal adhesion is either neglected or supposed inaffected by hammering then the normal wear component vanishes.

To summarize the adherence, tear and wear laws may be reduced to

$$\begin{pmatrix} F_N \\ F_T \end{pmatrix} = \begin{bmatrix} P_N & 0 \\ 0 & P_T \end{bmatrix} \begin{pmatrix} D_N^A \\ D_T^A \end{pmatrix} \quad \text{(a)}$$
$$\begin{pmatrix} F_N^S \\ F_T^S \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & Q_T \end{bmatrix} \begin{pmatrix} D_N^S \\ D_T^S \end{pmatrix} \quad \text{(b)} \qquad (6)$$
$$F^c = QD^c. \quad \text{(c)}$$

Similar laws, but expressed in terms of rates, may be found in [5]. It is recalled that the adherence, tear and wear laws model the interaction of a pair of materials in contact. Consequently the coefficient, $P_N$, $P_T$, $Q_T$, $Q$ must represent the average surface behaviour of these materials.

Finally it is pointed out that the adherence, tear and wear laws relate global quantities by opposition to incremental ones. This approach proves much simpler and more reliable in practice, whenever applicable as it is for the rate independent theory restricted to small slips at aim.

†The dimension of **P**, **Q** and $Q$ is $ML^{-2}T^{-2}$.
‡Counterpart of the Bauschinger effect observed in plasticity (Fig. 3b).

2.3 *Slip criterion* (*cf. yield criterion*)

To activate the kinematic decomposition of the distance of contact into adherence and slip, a *slip criterion* (also called friction criterion) is interrogated to decide which one of the two modes occurs. If adherence reaches (respectively drops below) a certain threshold called slip or adherence limit, then the relative motion contributes exclusively to slip (adherence) respectively.

It is emphasized that since both the normal and tangential components of the contact distance enter the definition of adherence, the slip criterion includes the contact criterion as well, i.e. the impenetrability condition.

In general the adherence-slip limit depends on all the state variables which for the purely mechanical theory under consideration gives

$$Y(D^A, D^S, D^c) < 0 \quad \text{contact and adherence} \qquad (7)$$
$$= 0 \quad \text{gap or slip.}$$

Upon substitution of the reversible constitutive laws (5) into the kinematic criterion, a dynamic *friction criterion* is obtained which is equivalent to the slip criterion but more appropriate to state eventual associated slip rules

$$Y(F, F^S, F^c) \leqslant 0. \qquad (8)$$

The friction-slip criterion is the homologue of the plastic yield criterion used in plasticity indifferently written in strain or stress space as

$$Y(E^E, E^P, E^c) \leqslant 0 \quad \text{or} \quad Y(S, S^P, S^c) \leqslant 0$$

with elastic deformation below and plastic beyond.

The normal impenetrability condition is comparable to the incompressibility of certain materials.

Without aiming at an exhaustive study of possible friction criterions, it is instructive to consider two typical cases (Fig. 4).

*Perfect friction*. The law of perfect friction states that the force of friction is proportional to the load and is independent of the apparent area of contact and the other state variables. Combined to the impenetrability condition the criterion of perfect friction

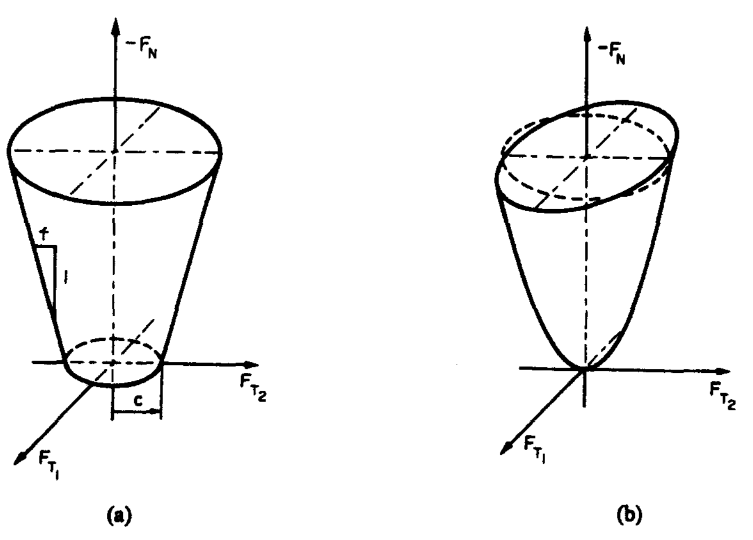


Fig. 4. Examples of slip criterions, (a) Coulomb's cone, (b) anisotropic paraboloid.

takes the form

$$Y(F) = \begin{cases} F_N \leqslant 0 & \text{contact} \\ |F_T| + fF_N - C \leqslant 0 & \text{slip} \end{cases} \quad (9)$$

where $|F_T| = \sqrt{F_{T_1}^2 + F_{T_2}^2}$ denotes the euclidian norm of $F_T$, $f$ the *coefficient of friction* and $C$ a constant characterizing adhesion[6]. In geometric terms the criterion assumes the shape of a truncated cone in the normal-tangential axes attached to the contact point. It is the analogue of the Drucker–Prager criterion in plasticity and constitutes the determining ingredient of Coulomb's law of friction.

For $f = 0$ the cone degenerates into a cylinder which corresponds to a force of friction independent of the load (e.g. for mica). It is the analogue of the deviatoric energy criterion of Von Mises.

The replacement of the euclidean norm $|F_T|$ by an elliptic norm

$$|F_T| = \sqrt{\frac{F_{T_1}^2}{a^2} + \frac{F_{T_2}^2}{b^2}}$$

(where $a$ and $b$ are the principal axes of the ellipse) transforms the isotropic criterion into an *anisotropic* one, accounting for preferential roughnesses in the tangential directions 1 and 2[5].

*Friction with wear.* Experiments show that real contacts often deviate from the law of perfect friction specially for very light and heavy loads. In the context of the present theory such deviations may be accounted for by acting on the *shape* of the slip criterion and including running in or *tear* and *wear* mechanisms similar to plastic softening. For instance

$$Y = \begin{cases} F_N \leqslant 0 & \text{contact} \\ (F_T - F_T^S)^2 + (fF_N - F^c) \leqslant 0 & \text{slip} \end{cases} \quad (10)$$

This criterion is a truncated paraboloid which axis may translate with the force of tear $F_T^S$ and which radius may decrease with the force of wear $F^c$ starting for instance from a virgin value $F^c = C$.

To improve the flexibility, the quadratic power may be replaced by an adjustable exponent $n$ and in the absence of tearing due to $F_T^S$, a fairly general, yet simple, friction criterion results:

$$Y = |F_T|^n + fF_N - F^c \leqslant 0. \quad (11)$$

2.4 *Slip rules (cf. flow rules)*

The slip direction (including take off) is governed by slip rules deriving from a convex potential $Z(F, F^S, F^c)$

$$\begin{aligned} \mathrm{d}D^S &= \lambda \frac{\mathrm{d}Z}{\mathrm{d}F} && \text{(a)} \\ -\mathrm{d}D^S &= \lambda \frac{\mathrm{d}Z}{\mathrm{d}F^S} && \text{(b)} \\ -\mathrm{d}D^c &= \lambda \frac{\mathrm{d}Z}{\mathrm{d}F^c} && \text{(c)} \end{aligned} \quad (12)$$

where $\lambda$ is a, yet arbitrary, constant expressing the collinearity of the slip increment with the outward normal to the potential $Z$. The fact that the three increments derive from the same potential characterizes a model of *standard generalized friction*. In particular the laws (12a,b) imply the dependence of the potential $Z$ on the relative force $F - F^S$ issued from the tear force.

The replacement of slip velocities $\dot{D}^S \ldots$ by slip increments $dD^S = \dot{D}^S dt \ldots$ is typical of a theory of friction *independent of the slip rate*. (Time entering the equations as an homogeneous variable may be eliminated.)

The slip rules play the role of equations of motion for the additional internal variables introduced in this theory which are the slips with wear forces as conjugates.

The slip rules are equivalent to the flow rules in plasticity which for a standard generalized material take the form [16],

$$dE^P = \lambda \frac{dZ}{dS} \quad -dE^P = \lambda \frac{dZ}{dS^P} \quad -dE^c = \lambda \frac{dZ}{dS^c}.$$

If the slip potential $Z$ is replaced by the slip criterion $Y$, the slip rule (12) becomes *associated* with the criterion. Although in plasticity the flow rules associated with the standard criterions prove realistic for relatively large classes of materials, the slip rules associated to the usual friction criterions like (9)–(11) are *not* acceptable (Fig. 5). Indeed the incompressibility of plastic deformations implied by a flow rule associated with the Von Mises criterion finds wide applications. On the contrary the normal component separating the two bodies

$$dD_N{}^S = \lambda \frac{dY}{dF_N} = \lambda f \quad (>0)$$

produced by a slip rule associated to the classical criterion of perfect friction bears no support except if the coefficient of friction vanishes, i.e. when the force of friction is independent of the load.†

Consequently a cylindrical slip potential with a hemispherical cap

$$Z = \frac{1}{2}(F_T - F_T{}^S)^2 - \frac{1}{2}F^{c2} + \frac{1}{2}\max(0, F_N|F_N|)$$

seems to be the only admissible surface which guarantees that slip begins in the tangent

†Another argument against associated slip rules may be found in [5] but it is not too convincing because as any model based on asperity slopes it seems to forget that climbing an asperity tip is systematically followed by sliding down the other side which is a globally conservative process in contradiction with the dissipative nature of friction.

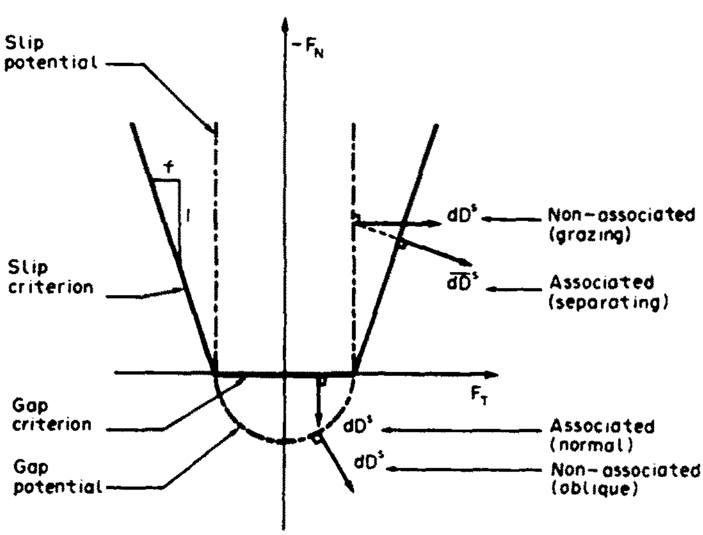


Fig. 5. Associated and non associated slip rules.

plane common to the two points in contact, as most easily verified in the absence of tear:

$$\mathrm{d}D' = \lambda F \quad \text{if } F_N > 0$$
$$\mathrm{d}D^S = \lambda F_T \quad \text{if } F_N < 0.$$

The normality rules (12) may be derived from a maximum work principle generalised to non associated slip rules but specialised to friction[4, 15].

To ascertain the signification of a slip rule, consider the simple case associated to a constant friction criterion (Fig. 6)

$$Y = \frac{1}{2} F_T^2 - C \leqslant 0 \quad (\text{for } F_N < 0).$$

Suppose that the force of friction $F_T$ just exceeds the limit of adherence by an amount $\mathrm{d}F_T$ oriented along an arbitrary direction, for instance perpendicular to $F_T$. Then the slip rule dictates

$$\mathrm{d}D_T{}^S = \lambda F_T.$$

Thus incipient slip occurs in the direction of $F_T$ (or $F_T + \mathrm{d}F_T$) but never along $\mathrm{d}F_T$ as intuition could mislead. The flow rules of plasticity follow the same logic. In fact the analogy between a slip rule associated to an elliptic friction criterion and a flow rule associated to the elliptic plastic criterion of Von Mises in plane stress is quite striking.

## CONCLUSION

The presentation of this small slip theory of standard generalised friction independent of the slip rate is completed.

The basic ingredients are recalled to consist of a kinematic decomposition (3), constitutive laws to generate the conjugate dynamics (5), a transition criterion (8) and an additional equation of evolution (12).

The result is a generalisation of Coulomb's law accounting for the influence on the macroscopic coefficient of friction of:

—the normal *load* (hereby resulting into a nonlinear dependence of the force of friction on the normal load);

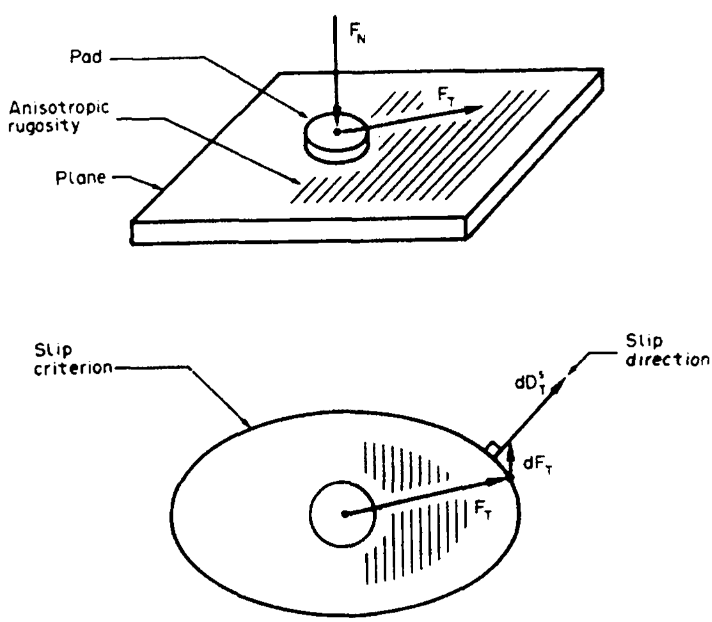


Fig. 6. Essence of the slip rule.

—the initial *rugosity* of the surfaces in contact (whether it is isotropic or anisotropic);

—the subsequent *wear* of these surfaces (whether it is monotone or alternate).

The combination of the normal impenetrability condition and the tangential law of friction into an integrated law of frictional contact is another feature of this theory which deserves a mention.

However this quasi-static infinitesimal theory is still restricted to small slips or to nominally flat surfaces of contact and the dependence of the coefficient of friction on the slip rate is limited to the schematic distinction between a static and a dynamic value of this coefficient. More practically, the theory is applicable to quazi-static contact problems involving microslips like normal approach punch problems, joints and hinges subjected to alternating loads....

Its implementation in, say, a general purpose finite element program, requires a trial and error strategy to deal with the inequality criterion and a stable algorithm to integrate the differential slip rule. A confrontation of this theory with extensive experimental data should help to assess its validity.

It is hoped that this attempt to provide the mechanics of friction with a structure well proved in plasticity will contribute to diminish the lack of rigor which afflicts some of the work in this field.

## REFERENCES


1. R. Hill, *The Mathematical Theory of Plasticity*. Oxford University Press (1950).
2. A. Mendelson, *Plasticity: Theory and Application*. Macmillan, New York (1970).
3. L. Kachanov, *Foundations of the Theory of Plasticity*. North Holland, Amsterdam (1971).
4. B. Fredriksson, Finite element solution of surface nonlinearities in structural mechanics with special emphasis to contact and fracture mechanics problems. *Comput. Structures* **6**, 281–290 (1976).
5. R. Michalowski and Z. Mróz, Associated and non-associated sliding rules in contact friction problems. *Arch. Mech.* **30**, 259–276 (1978).
6. A. S. Kravchuk, On the theory of contact problems taking account of friction on the contact surface. *J. Appl. Math. Mech. (PMM)* **44**, 83–85 (1981).
7. E. Green and P. Naghdi, A general theory of an elastic–plastic continuum. *Arch. Rational Mech. Anal.* **18**, 251–281 (1965).
8. S. Nemat-Nasser, Decomposition of strain measures and their rates in finite deformation elastoplasticity. *Int. J. Solids Structures* **15**, 155–166 (1979).
9. K. Neale, Phenomenological constitutive laws in finite plasticity. *Solid Mech. Arch.* **6**, 79–127 (1981).
10. G. Duvaut and J. L. Lions, *Les Inéquations en Mécanique et en Physique*. Dunod, Paris (1972).
11. A. D. De Pater and J. J. Kalker, *The Mechanics of the Contact between Deformable Bodies*. Delft University Press (1975).
12. T. J. Oden and E. B. Pires, Numerical analysis of certain contact problems in elasticity with non-classical friction laws. *Comput. Structures* **16**, 481–485 (1983).
13. T. J. Hughes *et al.*, A finite element method for a class of contact-impact problems. *Comput. Meth. Appl. Mech. Engng* **8**, 249–276 (1976).
14. F. P. Bowden and D. Tabor, *The Friction and Lubrication of Solids*. Oxford University Press, Vol. I (1958), Vol. II (1964).
15. A. C. Palmer, A limit theorem for materials with non-associated flow laws. *J. Mécanique* **5**, 217–222 (1964).
16. B. Halphen and N. Q. Son, Sur les matériaux standards généralisés. *J. Mécanique* **14**, 39–63 (1975).