On the microphysical foundations of rate-and-state friction

Thibaut Putelat\textsuperscript{a,c}, Jonathan H. P. Dawes\textsuperscript{b}, John R. Willis\textsuperscript{c}

\textsuperscript{a}Institute of Theoretical Geophysics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA

\textsuperscript{b}Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY

\textsuperscript{c}Department of Applied Mathematics \& Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA

Abstract

The rate-and-state formulation of dry friction is well established as a phenomenological yet quantitative description of dry friction dynamics, in particular the onset of stick-slip instabilities. First, we discuss the physical origins of two theories for the derivation of friction laws used in rate-and-state models. Second, we propose a general form for the state evolution law, in the form of a first-order differential equation describing relaxation to a steady state. We show that the differential equation can be estimated directly from experimental measurements of the onset of stick-slip behaviour, and we illustrate this using a specific dataset.

Key words: rate-and-state friction, asperity creep, Eyring transition-state theory, memory length.

PACS: \textsuperscript{1}

Contents

1 Introduction .......................................................... 2

2 The friction coefficient ............................................ 3

2.1 Constitutive theory of asperity deformation ................. 4

2.2 Thermodynamic theory for slip events ...................... 6

3 The state evolution law ............................................. 8

4 Discussion ............................................................. 11

---

\textsuperscript{1} Corresponding author

Email addresses: T.Putelat@damtp.cam.ac.uk (Thibaut Putelat), J.H.P.Dawes@bath.ac.uk (Jonathan H. P. Dawes), J.R.Willis@damtp.cam.ac.uk (John R. Willis)

Preprint submitted to Elsevier

March 16, 2010
1. Introduction

The physical description of friction at interfaces has a long history and has been the subject of particularly careful theoretical and experimental work over the last few decades. Much of the current state of research is surveyed by Persson (2000), Müser et al. (2003) and Baumberger and Caroli (2006). A major step forward that has underpinned many subsequent theoretical developments was the proposal by Bowden and Tabor (1939, 1954) that friction forces are determined by the plastic deformation of micro-asperities: this sets the real area of contact between moving interfaces, and is in direct proportion to the normal load. This proposal gives a physical explanation of the first Amontons law which states that there is a linear relationship between the tangential friction force and the applied normal load. Bowden & Tabor’s proposal also explains the second Amontons law which states that the friction coefficient is independent of the apparent macroscopic area.

Later work by Rabinowicz (1957) recognized that intrinsic length and timescales of asperities and their dynamics must play important roles in the evolution of frictional stick-slip processes. This overturned the classical idea of Coulomb friction in which the friction force is only governed by the instantaneous sliding conditions. With this view Rabinowicz proposed that the friction coefficient depends on the average interfacial sliding velocity during slip over a critical distance. An important outcome of this concept was the correspondence between the critical distance and the minimum slip necessary for stick-slip to appear, which first gave a possible physical explanation why system stiffening and increasing velocity eliminate oscillations.

However, it was not until the early 1980s when the rate-and-state framework for friction was proposed that one could understand quantitatively why stick-slip oscillations only occur when the system stiffness is less than a critical value which, in turn, depends on the macroscopic velocity difference across the whole interface (Dieterich, 1979, Ruina, 1980, 1983, Rice and Ruina, 1983, Gu et al., 1984). In the rate-and-state framework, the stick-slip instability can be described as a time-periodic oscillation which originates from a Hopf bifurcation from the uniformly sliding state.

The rate-and-state framework for friction proposes that the shear stress \( \tau \) at a frictional interface depends on the interfacial slip rate \( v \) and an internal variable \( \phi \) which describes the current state of the interface (i.e. its resistance to sliding), in addition to the usual dependence on the normal stress \( \sigma \). The state variable \( \phi \) therefore models the complicated deformation dynamics of the interfacial asperities responsible for the memory effects of friction.

Formally, a rate-and-state friction law is defined by a pair of equations of the form (e.g. Rice et al., 2001)

\[
\begin{align*}
\tau &= F(v, \phi; \sigma), \\
\dot{\phi} &= -G(v, \phi; \sigma).
\end{align*}
\]

(1)

The first of these equations gives the dependence of the shear stress \( \tau \) on the variables described above. An equivalent formulation would be to write the friction coefficient \( \mu \equiv \tau/\sigma = F(v, \phi; \sigma)/\sigma \). The second equation above describes the time evolution of the interfacial state variable.

Widely used examples of such friction laws are those known as the Dieterich–Ruina laws, or equivalently the ageing/slip laws, for which the friction coefficient is given by

\[
\mu(v, \phi) = \mu_0 + a \ln(v/V_s) + b \ln(V_s \phi/L),
\]

(2)

and the state evolution law is respectively taken to be either the aging law

\[
\dot{\phi} = 1 - v\phi/L,
\]

(3)

or the slip law

\[
\dot{\phi} = -(v\phi/L) \ln(v\phi/L),
\]

(4)

where \( V_s \) is a typical sliding velocity and \( L \) is a characteristic length scale (the ‘memory length’) over which the interfacial strength varies in response to changes in the slip velocity (see Marone, 1998, for a review).

It is important to bear in mind that the law (1) describes instantaneous frictional motion in which the interfacial state continuously evolves, whereas in steady state at a slip rate \( v = V_s \) the interfacial state can
be written as a function $\phi_{ss}(V)$ given implicitly by solving $G(V, \phi_{ss}(V)) = 0$ and hence for steady sliding we have $\tau = F_{ss}(V) \equiv F(V, \phi_{ss}(V))$ when $v = V$ constant.

In the simplest case of a single interface, neglecting wave propagation, confined between slabs of thickness $h/2$ with uniform material properties, linear elasticity provides an additional equation for the time evolution of the interfacial shear stress which is proportional to the velocity of the interface relative to the driving velocity $V$: this closes the system of equations. In dimensional form this reads

$$\dot{\tau} = k[V - v(\tau, \phi)],$$  

(5)

where $v(\tau, \phi)$ is obtained by inverting the friction law $\tau = F(v, \phi; \sigma)$, and $k = G/h$ is a measure of the stiffness of the material. For the system (1) - (5), the steady sliding solution $v = V$ becomes linearly unstable via a Hopf bifurcation at critical stiffness $k_c$ which is proportional to the negative of the slope of the steady-state friction coefficient. As a result, stick-slip dynamics can occur only when friction is velocity weakening, i.e. the steady-state friction coefficient decreases with driving velocity $V$. Recently (Putelat et al., 2010) we made quantitative investigations of the relation between the exact form of the nonlinearities involved in the friction law (1) and the supercritical or subcritical nature of the Hopf bifurcation from which stick-slip dynamics originate. We suggested that this kind of weakly nonlinear analysis enables discrimination between candidate friction laws when discussing which models fit a given set of experimental observations, taking as an example those of Heslot et al. (1994). We also believe that the rate-and-state formalism is a good candidate for explaining the relaxation dynamics of friction in response to abrupt jumps in the driving velocity $V$.

In summary it is both physically plausible and mathematically tempting to regard the rate-and-state formalism as a universal phenomenological description of friction when the sliding velocities involved are relatively low so that one may disregard flash heating. This universality motivates the probing of its microphysical foundations and the assumptions that underpin it, implicitly or explicitly, in any derivation of a friction law. In this paper we discuss these foundations, bringing together two different approaches to the derivation of the friction law (1) and making the new observation that experimental work can be used to validate the state evolution law (1).

The contents of the paper are as follows. In section 2 we first summarise Bowden and Tabor’s assumptions and then present two microphysical theories of friction which are commonly invoked in the derivation of a rate-and-state friction law. We refer to them as the ‘constitutive theory’ (section 2.1) and the ‘thermodynamic theory’ (section 2.2). Both combine (in different ways) the ideas of an Arrhenius-like thermally activated rate process and Bowden and Tabor’s theory of friction. In section 3 we discuss the state evolution law: although an essential part of the rate-and-state model, this has been neglected in comparison to the detailed microscopic arguments available to justify the friction law. Under mild assumptions on the form of the state evolution law (only that it is described by a first-order process, implying that relaxation to the steady state is monotonic) we show that it should be possible experimentally to determine the form of the steady state dependence $\phi_{ss}(V)$ of the state variable $\phi$ on the driving velocity $V$. Such an experimental result would be a major contribution to the foundations of the rate-and-state formalism, and we hope that experimentalists will take up this challenge. We conclude in section 4.

2. The friction coefficient

In this section we begin with brief remarks on rate processes and on Bowden and Tabor’s foundational ideas for the study of friction. We then discuss each of the microphysical theories in turn.

The transition-state\(^1\) theory of Eyring rate processes (Eyring, 1935) accounts for the rearrangements of matter by defining the frequency $\nu$ (in Hz) with which an event occurs when it has to overcome a potential energy barrier of height $E_a$. If body deformation is present, the height of the energy barrier is reduced by

\(^1\)or ‘absolute reaction-rate’
an amount $\Omega \Sigma$ corresponding to the work done by the applied stress $\Sigma$ through the activation volume $\Omega$. This leads to the expression
\[ \nu = \nu_0 \exp \left( \frac{E_a - \Omega \Sigma}{k_B T} \right), \]
where $\nu_0$ is a reference frequency, $k_B$ is the Boltzmann constant and $T$ is the absolute temperature. A brief exposition of the transition-state theory of chemical reaction-rate is given by Glasstone (1948), and its application to plastic deformation of crystals due to the motion of dislocations can be found in Poirier (1985).

We let the nominal contact area of the interface be $A$ and let the macroscopic normal force and friction force be denoted by $W$ and $F$, respectively. These define the macroscopic normal and shear stresses $\sigma$ and $\tau$ through $\sigma = W/A$ and $\tau = F/A$. We now introduce the twin assumptions that Bowden and Tabor make in their theory of friction: that both the macroscopic normal force $W$ and the macroscopic friction force $F$ are proportional to the real contact area $A_r$, i.e.
\begin{align}
W &= A_r \sigma_c, \\
F &= A_r \tau_c. \tag{7} \tag{8}
\end{align}
In the sequel, we refer to expressions (7) and (8) as hypotheses BT1 and BT2. These hypotheses relate the macroscopic interfacial normal and shear stresses $\sigma$ and $\tau$, to the stresses at the contact scale $\sigma_c$ and $\tau_c$. Consequently the coefficient of friction can be defined by
\[ \mu \equiv \frac{F}{W} = \frac{\tau}{\sigma} = \frac{\tau_c}{\sigma_c}, \tag{9} \]
in agreement with the second Amontons law of friction, according to which the friction force is independent of the apparent area. We refer to (9) as the Bowden and Tabor postulate of friction. In this framework, it is important to recall that the local stresses $\sigma_c$ and $\tau_c$ are supposed to correspond to plastic yield stresses because the hypotheses BT1–BT2 follow from considering the plastic deformation of asperity junctions.

Most importantly, we stress that there is experimental evidence showing that “at the microscopic level the friction force $F$ is proportional to the net applied load $W$ and not to the real area of contact” (Berman et al., 1998). Berman et al. even conclude that “Amontons’ law (9) is also obeyed directly at the molecular level and does not emerge indirectly because of some fortuitous correlation between the net applied load and the local contact area or shear strength, as is commonly supposed”.

2.1. Constitutive theory of asperity deformation

The constitutive theory first assumes the hypothesis BT2 and decomposes the friction force as a product, letting the asperity shear strength $\tau_c$ depend only on the rate $v$, and letting the true contact area $A_r$ depend only on the state $\phi$, i.e.
\[ F = A_r(\phi) \tau_c(v). \tag{10} \]
To justify such a decomposition it is necessary to consider a constitutive law and a model for the asperity deformation, which we now discuss in detail. The essence of this argument originates from Bréchet and Estrin (1994), Dieterich and Kilgore (1994), Estrin and Bréchet (1996) and is invoked in many references, for example see sections 13.2 and 13.3 of Persson (2000) and Berthoud et al. (1999), Baumberger et al. (1999), Baumberger and Caroli (2006). It is usually assumed that interfacial asperities deform by a thermally activated creep process whose constitutive law relates a scalar measure $\dot{\epsilon}_c$ of the strain rate of the asperities to their local stress $\sigma_c$ or $\tau_c$. The state dependence of $A_r$ and the rate dependence of $\tau_c$ are then obtained by artificially dividing the creep of asperities into a one-dimensional pure shear deformation and a one-dimensional simple shear deformation, respectively.

More precisely, the state dependence of the true contact area $A_r$ follows from its time dependence and the interpretation of the interfacial state as a contact time. Assuming that the deformation is plastic and that volume is conserved, the asperity strain rate is related to the evolution of the contact area through $\dot{\epsilon}_c = \dot{A}_r/A_r$. Note that the strain rate measure is defined so that it is positive when an increase of the contact
area is observed: as a consequence, we use the convention that the stress is positive in compression. Under a constant macroscopic load \( W \), the contact normal stress \( \sigma_c \) must then be given by \( \sigma_c = W/A_r \) according to hypothesis BT1. Using a constitutive law \( \dot{\varepsilon}_c = f(\sigma_c) \), an evolution equation for the true contact area is obtained:

\[
\dot{A}_r = A_r f(W/A_r).
\]  

(11)

The approximate integration of (11) is possible if we assume that the deformation of asperities produces little variation of the true contact area and also that the constitutive law \( f(\sigma_c) \) is given by a Nabarro–Herring creep-type law (Bréchet and Estrin, 1994, Berthoud et al., 1999, Persson, 2000). The constitutive law corresponds to a rate process of the form (6) with a reference strain rate \( \dot{\varepsilon}_0 \) that we write as

\[
\dot{\varepsilon} = \dot{\varepsilon}_0 f(\sigma_c) \equiv \dot{\varepsilon}_0 \exp \left( \frac{E_{\perp} \sigma_c - \sigma_Y}{k_B T \sigma_Y} \right),
\]  

(12)

derived from (6) with \( \sigma_Y \equiv E_{\perp}/\Omega_{\perp} \). The energy and volume of activation \( E_a \) and \( \Omega \) are denoted \( E_{\perp} \) and \( \Omega_{\perp} \) to recall the pure shear nature of the asperity deformation at this stage of the reasoning. Then, in the limit \( k_B T/E_{\perp} \equiv \beta \ll 1 \) we may integrate (11) at leading order in \( \beta \) to obtain

\[
A_r(t) = A_r(0) \left[ 1 + \beta \ln(1 + t/t_0) \right] + O(\beta^2),
\]  

(13)

where \( t_0 = \beta/\dot{\varepsilon}_0 \) defines a cut-off time. In an experiment in which \( W \) remains constant, the requirement to set \( W = A_r(0) \sigma_Y \) allows both \( \sigma_Y \) to be interpreted as a plastic yield stress of asperities and the choice of the initial true contact area \( A_r(0) \) to be the contact area after plastic yielding first occurred under the normal load \( W \). Note that in this way we recover here some elements of Bowden and Tabor’s reasoning.

A creep law of Nabarro–Herring type (12) can also be used to derive the rate dependence of the local interfacial shear strength \( \tau_c(v) \), by applying it to the simple shear of asperities considering that its energy and volume of activation are \( E_a \) and \( \Omega_a \). After inverting the creep law, we obtain

\[
\tau_c = \tau_Y \left[ 1 + \alpha \ln(\dot{\gamma}/\dot{\gamma}_0) \right],
\]  

where we define the yield stress \( \tau_Y = E_a/\Omega_a \) and the energy scale \( \alpha = k_B T/E_a \), for the reference shear strain rate \( \dot{\gamma}_0 \). Without further information we must assume that the energy barrier \( E_a \) and the volume of activation \( \Omega_a \) in simple shear are different to those in pure shear. Dimensional analysis suggests that the shear rate \( \dot{\gamma} \) for interfacial simple shear scales like the interfacial slip rate \( v \) divided by a characteristic depth \( \delta \)

\[
\dot{\gamma} \sim v/\delta.
\]

At this level, a more sophisticated model of the simple shear of asperities would be necessary to identify \( \delta \) and \( \dot{\gamma}_0 \) precisely. As a result, and after defining a reference slip rate \( V_s \sim \delta \dot{\gamma}_0 \), the interfacial shear strength is given by

\[
\tau_c(v) = \tau_Y \left[ 1 + \alpha \ln(v/V_s) \right].
\]  

(14)

Thus, at leading order in \( \beta \) and \( \alpha \), the dependence of the friction force on \( t \) and \( v \) is given by

\[
F = A_r(0) \tau_Y \left[ 1 + \beta \ln(1 + t/t_0) \right] \left[ 1 + \alpha \ln(v/V_s) \right],
\]

and the friction coefficient \( \mu \) is given by

\[
\mu = \mu_s \left[ 1 + \beta \ln(1 + t/t_0) \right] \left[ 1 + \alpha \ln(v/V_s) \right],
\]

where we recall that \( A_r(0) = W/\tau_Y \) and we define \( \mu_s = \tau_Y/\sigma_Y \). The Dieterich–Ruina law (2) follows from neglecting terms of \( O(\alpha \beta) \). It is particularly interesting to note that this derivation gives specific physical interpretations of the material coefficients \( \mu_s \), \( \beta \), \( \alpha \), \( t_0 \) and \( V_s \) and also shows that the rate and state dependence of the friction force can be thought of as a higher order correction to the Bowden and Tabor theory in the limit of low thermal activation (and hence low rates of transitions) since \( \beta \) and \( \alpha \) are small.
Finally we remark that if backward jumps are taken into account in the expression for the local asperity shear strength, then the Nabarro-Herring creep law relating shear rate to shear strength takes the form 
\[ \dot{\gamma} = 2\gamma_0 e^{-E_/k_B T} \sinh (\Omega_*/(k_B T)) . \]
With the notation above, this leads to the slip rate dependence
\[ \tau_e(v) = \alpha \tau_Y \sinh^{-1} \left[ \left( v/(2V_e) \right) \exp(1/\alpha) \right], \]
recalling that \( \alpha \tau_Y = k_B T/\Omega_*. \) We note that this expression regularises (14) in the limit \( v \to 0 \) and can be moreover approximated by (14) when the argument of the \( \sinh^{-1} \) function is large, that is, over a wide range of the dimensionless slip rate \( v/V_e \) as long as \( k_B T \ll E_*, \) i.e. \( \alpha \ll 1. \) Writing \( a = \mu_* \alpha \) and \( b = \mu_* \beta, \) the constitutive theory finally gives the friction coefficient
\[ \mu = \frac{\alpha}{a} \left[ 1 + \left( \frac{b}{\mu_*} \right) \ln \left( 1 + \frac{t}{t_0} \right) \right] \sinh^{-1} \left[ \frac{v}{2V_e} \exp \left( \frac{\mu_*}{a} \right) \right]. \]
(15)

2.2. Thermodynamic theory for slip events

The thermodynamic theory for slip events originates from Schallamach (1953, 1963) who observed that the temperature dependence of the sliding velocity at constant pulling force of rubber obeys an Arrhenius equation
\[ v \propto \exp(-E_/(k_B T)), \]
which suggests that frictional sliding arises from a rate process determined by an activation mechanism.

The thermodynamic theory combines the same elements of the formalism discussed in section 2.1, but makes use of the transition-state theory for events at a much larger scale: the scale of slip patches instead of single asperities. It begins by considering a local slip event to occur when the energy barrier \( E_* \) is overcome in combination with the simple shear of asperities, collectively, at constant stress \( \tau_e \) over an activation volume \( \Omega_* \). Then, a statistical theory for slip events implies that the rate at which the energy barrier is surmounted in either the forward (+) or backward (−) direction is given by
\[ \nu^\pm = \nu_0 \exp \left( -\frac{E_* + \Omega_* \tau_e}{k_B T} \right), \]
where \( \nu_0 \) is a reference frequency of slip attempts. In similar fashion to the creep flow laws of the constitutive theory, the volume of activation is defined by \( \Omega_* = E_*/\tau_Y \) which means that the simple shear of asperities occurs once the local shear stress reaches the yield stress \( \tau_Y. \)

The (local average) shear stress \( \tau_e \) at a slip patch does an amount of work \( \Omega_* \tau_e \) which promotes forward-slip transitions and restrains backward motion across the energy barrier \( E_* \). The net frequency of slip processes is then given by
\[ \nu \equiv \nu^+ - \nu^- = 2 \nu_0 \exp \left( -\frac{E_*}{k_B T} \right) \sinh \left( \frac{\Omega_* \tau_e}{k_B T} \right), \]
which implies that the macroscopic slip rate \( v \) resulting from successful slip events is
\[ v = 2 \lambda \nu_0 \exp \left( -\frac{E_*}{k_B T} \right) \sinh \left( \frac{\Omega_* \tau_e}{k_B T} \right), \]
(17)
in which the characteristic length \( \lambda \) is interpreted in fact as an average separation between the energy barriers (Heslot et al., 1994). From this we may define a reference slip rate by \( V_e = \lambda \nu_0 \). When inserted in (17), we can use the Bowden and Tabor postulate expressed by equation (9), seen either as a first principle or as a consequence of the hypotheses BT1 and BT2 (equations (7) and (8)), to relate the interfacial slip rate \( v \) to the macroscopic stresses in the form
\[ v = 2 V_e \exp \left( -\frac{E_*}{k_B T} \right) \sinh(\tau/(a\sigma)), \]

\[ \text{These two hypotheses are not necessary in this point of view.} \]
where we define the coefficient $a = k_B T / (\Omega_\sigma \sigma_c)$. Inverting this relation gives the rate dependence of the friction coefficient

$$\mu = \tau / \sigma = a \sinh^{-1} \left( \frac{v}{2V} \exp \left( \frac{E_*}{k_B T} \right) \right).$$

(18)

In contrast to the constitutive theory of section 2.1 therefore, in the slip-event theory the friction coefficient is derived directly from the rate equation and the BT postulate (9). When relations such as (18) are used, the coefficient $a$ is generally considered to be time-independent which in turn implies that $\sigma_c$ is constant. The form (15) of the constitutive theory is recovered from (18) if hypothesis BT1 is invoked, implying that $\sigma_c = W / A_r(t)$, giving immediately

$$a = \frac{k_B T}{\Omega_\sigma \sigma_c} \left( \frac{k_B T}{\Omega_\sigma W} A_r(t) \right).$$

Thus, in comparison with the constitutive theory of section 2.1, the thermodynamic theory neglects at this stage the effect of the logarithmic increase of the true contact area as involved in (13): equivalently it assumes $\sigma_c = \sigma r$.

We now turn to mechanisms for the inclusion of state dependence $\phi$ in the friction coefficient (18). The natural way to include state dependence is to consider variations over time in the height of the thermal energy barrier $E_*$, reflecting the maturity of the contact population (Heslot et al., 1994, Rice et al., 2001). Since for large values of the argument of the $\sinh^{-1}$ function we have the approximation $\sinh^{-1}(x) \approx \ln(2x)$, the friction coefficient (18) can be simplified to obtain

$$\mu = a \ln (v / V_\sigma) + \frac{E_*}{\Omega_\sigma \sigma_c}. \quad (19)$$

This expression compared to the Dieterich–Ruina friction coefficient (2) suggests the form

$$E_*(\phi) = \Omega_\sigma \sigma_c \left[ \frac{\mu_* + b \Psi(\phi)}{\Omega_\sigma \sigma_c} \right]$$

with $\Psi(\phi) = \ln(\phi / \phi_*).$ As the transition-state theory requires $E_* = \Omega_\sigma \tau r$, we conclude that the state dependence of $E_*$ may come from the variation of the yield stress in the form

$$\tau r = \left[ \frac{\mu_* + b \ln(\phi / \phi_\sigma)}{\sigma c} \right]$$

(20)

either by ageing or hardening of the interfacial asperities in relation to their pure shear deformation promoted by the normal pressure. In the case of ageing, the state variable $\phi$ would be interpreted as time. In the case of hardening, we expect the state variable to correspond to a strain measure of the asperity creep. As the Dieterich–Ruina law (2) implies that $E_*(\phi)$ is an increasing function of $\phi$, this evolution of $E_*$ from hardening is intuitively more appealing phenomenologically because the observation (Bowden and Tabor, 1954, Dieterich and Kilgore, 1994) of the increase of the true area of contact $A_r$ with the time of (quasi-)stationary contact can be incorporated in the modelling, with the contact area $A_r$ considered as the strain measure.

We now briefly consider the details of two lines of argument to incorporate the state variable into the expression (19) for the friction coefficient. A first attempt of this kind, formally, would be to assume that the deformation within a slip patch corresponds to that for a hyperelastic material whose strain measure is based on Hencky’s natural strain (Hencky, 1931, Anand, 1979). For instance, denoting the principal stretches by $\lambda_1$ and $\lambda_2$, an isochoric pure shear would involve principal components of the Cauchy stress and a deviatoric stress measure of the forms, respectively, $\sigma_t = -p + 2G \ln \lambda_1$ and $\sigma_t - \sigma_2 = 2G \ln \lambda (\lambda_1 = \lambda_2^{-1} = \lambda)$ on which the yield stress $\tau r$ could be defined. One sees in this case that the state variable would be associated to a principal stretch which then measures relative variations of area between two different configurations.

Secondly, and to draw a parallel with the constitutive theory, another possible mechanism might be to observe that expression (20) suggests considering another deformation mechanism with the creep law

$$\dot{\epsilon} = \dot{\epsilon}_0 \exp \left( - \frac{E_* - \Omega_\sigma \tau r}{k_B T} \right),$$

7
where the yield stress $\tau_Y$ for the simple shear deformation controlling the frequency of slip events is set by the asperity creep, for which the energy barrier $E_+ \doteq \Omega_+ \sigma_c$ defines the yield stress $\tau_+ = E_+/\Omega_+$. As a result, we define $b = \mu_* k_B T / E_+ = k_B T / (\Omega_+ \sigma_c)$. Comparing the creep law above and (20), we obtain that
\[
\tau_+ = \mu_* \sigma_c
\]
and that the state variable corresponds to a strain rate of patches of asperities. Equation (21) gives an interpretation for $\mu_*$ and suggests that the physical process from which the energy barrier $E_+$ would originate results from the classical Amontons–Coulomb friction as supported by the experimental evidence of Berman et al. (1998) according to which Amontons’ law also applies at the microscopic level.

Actually, without referring to any precise microphysical mechanism, one could conclude from contemplating expression (20) that the variations of the energy barrier $E_+$ of a slip event could directly follow from some state dependent frictional phenomenon at the microscopic scale. Then, the simplest possible process leading to the form (20) can be delivered by a transition-state Eyring process for which
\[
\phi = \phi_\ast \exp \left( -\frac{\mu_* \sigma_c - \tau_Y}{b \sigma_c} \right) = \phi_\ast \exp \left( -\frac{E_+ - \Omega_+ \tau_Y}{k_B T} \right).
\]
As a result, the interfacial state $\phi$ corresponds to the average frequency of the transition-state process and the coefficients $\mu_*$ and $b$ are defined by
\[
\mu_* = \frac{E_+}{\Omega_+ \sigma_c} \quad \text{and} \quad b = \frac{k_B T}{\Omega_+ \sigma_c}.
\]

Finally, we conclude that the memory effects of friction responsible for the evolution of the activation energy $E_+$ of slip event could have different origins depending on the material properties of the frictional system under consideration. In particular we gave different examples leading to the same logarithmic state dependence of $E_+$ which suggests that the physical mechanisms causing frictional memory effects may not be unique. A possible discrimination between them would need a careful study of the state evolution law.

3. The state evolution law

In this section we turn to the derivation of the state evolution law (12) for which it appears that no microscopic arguments have been put forward. In some modelling approaches, the state law is almost redundant in the sense that $\phi$ has been considered to be exactly the time since the interfaces were first placed in contact, and hence the ‘state evolution law’ is simply $\dot{\phi} = 1$. However, in the rate and state framework one would prefer the more general supposition that the state variable $\phi$ evolves so as to relax towards an equilibrium value appropriate for the current sliding velocity $v$. Therefore it is natural to propose a general evolution law for $\phi$ in the form
\[
\dot{\phi} = \frac{1}{t_{ss}} \frac{\phi - \phi_{ss}(v)}{\phi_{ss}(v)} = -G(v, \phi; \sigma),
\]
where $t_{ss}\phi_{ss}(v) \equiv t_\phi$ is the relaxation time of $\phi$ towards its steady-state value $\phi_{ss}(v)$. Expression (22) can be thought of as describing, quite generally, linear deviations from the equilibrium value of the interfacial state $\phi_{ss}(V)$ which is achieved for steady sliding, when $v = V$. Equation (22) corresponds to a first order kinetics whose characteristic time scale is $t_{ss}$. We remark that we might also expect $t_{ss}$ and $\phi_{ss}$ to depend on the normal stress $\sigma$ but we will ignore this additional complication in order to keep the presentation simple.

With the assumption of the form (22) the central point of this section is to show that it is possible experimentally to deduce the form of the unknown function $\phi_{ss}(v)$ giving the dependence of $\phi$ on $v$ in
steady state. We show that $\phi_{ss}(v)$ can be derived through determination of the critical stiffness at which steady sliding becomes linearly unstable.

We begin by recalling the general expression for the critical stiffness $k_c$, below which steady sliding is unstable in the standard spring–block setup (e.g. see Gu et al., 1984, Putelat et al., 2010):

$$k_c = -G_\phi F'_{ss}(1 + MG_\phi/F_\nu) = -G_\phi F'_{ss} + M\omega_c^2,$$

where $\omega_c$, which denotes the critical frequency of small amplitude stick-slip oscillations, is given by

$$\omega_c^2 = -G_\phi^2 F''_{ss}/F_\nu.$$

From (22), we can easily compute the partial derivative

$$G_\phi = 1/(t_{ss} \phi_{ss}).$$

Now, combining (23) and (24) we see that $t_{ss} \phi_{ss}(V)$ can be obtained directly from knowledge of the steady-state friction law $F_{ss}(V)$, the critical stiffness $k_c(V)$ and the frequency $\omega_c(V)$ for different values of the driving velocity $V$, via the formula

$$t_{ss} \phi_{ss}(V) = -F''_{ss}/k_c - M\omega_c^2 = -F''_{ss}/k_c.$$

We remark that, from (23), the combination $k_c - M\omega_c^2$ does not in fact depend on the mass $M$. It corresponds to the quasi-static value $k_* = -G_\phi F'_{ss}$ of the critical stiffness when the block inertia is negligible. Therefore the quantity $\phi_{ss}(V)$ that we would compute using (25) is purely a property of the interface, as we would expect. The term $M\omega_c^2$ allows the corrections of the measurements of $k_c$ from the effect of inertia. From (25), it is furthermore natural to define the apparent memory length

$$L_a(V) \equiv V t_{ss} \phi_{ss}(V),$$

which is important to characterise the frictional sliding memory effect. From dimensional analysis, expression (26) also suggests that there exists an intrinsic memory length $L$, which is independent of $V$, and which we define as

$$L \equiv V_t t_{ss},$$

so that $L_a = (V/V_*) \phi_{ss}(V) L$. We note that the dimensionless quantity $\phi_{ss}$ is necessarily a function of $V/V_*$. A constant apparent memory length would then correspond to $\phi_{ss} \sim (V/V_*)^{-1}$. Plotting the curve $L_a(V)$ from experimental measurements of $F_{ss}(V)$ and $k_c(V)$ and using (25)–(26) would enable an experimental test of the common assumption of the use of a constant memory length ($L_a \equiv L$) for modelling frictional relaxation phenomena.

To illustrate such an experimental determination of $\phi_{ss}$, we will now consider the experimental data for Bristol paper board available in Heslot et al. (1994). The steady-state velocity-weakening regime of friction can be fitted by $\mu = \mu_0 - (b - a) \ln(V/V_*)$, with $\mu_0 = 0.369$, $b - a = 0.014$ and $V_* = 10^{-6}$ m s$^{-1}$. In particular, when inertia is negligible $k_c/M \gg \omega_c^2$ which implies that the measured $k_c \approx k_*$, the apparent memory length is

$$L_a = (b - a)Mg/k_c,$$

when the steady-state friction coefficient is approximated by the Dieterich–Ruina law for which $b - a$ represents the slope of the curve $\mu_{ss}(V)$. Figure 1 shows computations of the equilibrium interfacial state $t_{ss} \phi_{ss}$ and apparent memory length $L_a$ extracted, using (25) and (26), from the experimental measurements of $\mu_{ss}(V)$ and $k_c(V)$ made by Heslot et al. (1994). Disregarding the data points for $V > 2 \times 10^{-5}$m interpreted by these authors as corresponding to finite amplitude oscillations prior to a subcritical Hopf bifurcation, we find that $\phi_{ss}$ is a power law decreasing function of the velocity

$$\phi_{ss}(V) \approx (V_*/V)^{0.723} \quad \text{with} \quad t_{ss} \approx 10^{0.436} \approx 2.729 \text{ s}.$$
Consequently, as the apparent memory length $L_a$ increases from $O(10^{-6})$ to $O(10^{-5})$ m, its functional dependence on $V$ can be expressed as

$$L_a \approx (V/V_*)^{0.277} L \quad \text{with} \quad L \approx 10^{-5.564} \approx 2.729 \times 10^{-6} \text{ m}, \quad V_* = 10^{-6} \text{ m s}^{-1}.$$  

Thus we may deduce the empirically-determined state evolution law

$$\dot{\phi} = \left(\frac{1}{t_{\text{ss}}}\right) \left[1 - \left(\frac{V}{V_*}\right)^{\gamma}\right] \quad \text{with} \quad \gamma \approx 0.723.$$  

Being given the value of $a = 0.0349$ and $b = 0.0489$ (Heslot et al., 1994), it is interesting to note that the value of $\gamma$ that we find is close to the ratio $a/b \approx 0.714$. However, it is premature to draw more general conclusions as this method is valid for velocities which correspond to the pure logarithmic velocity weakening regime of friction where $|F_s(V)| = (b-a)Mg/V$.

Nevertheless, the generic velocity dependence of $\phi_{ss}(V)$ can be derived if an hypothesis is made on the analytical form of the function $F(v, \phi)$ inferred from the microphysical theories described earlier. Differentiating $F(v, \phi_{ss}(v))$ combined with equation (25) leads to a nonlinear first order differential equation for $\phi_{ss}(v)$:

$$F_\phi(v, \phi_{ss}) \phi_{ss}^\prime + F_v(v, \phi_{ss}) + t_{ss} k_s(v) \phi_{ss} = 0.$$  

In order to obtain some insight into the analytical velocity dependence of $\phi_{ss}(v)$ that the equation above implies, we consider a friction force $F = \mu(v, \phi)\sigma$ with a friction coefficient corresponding to the Dieterich–Ruina form (2) for simplicity. In this situation, introducing the scaled state variable $\psi = \phi_{ss}/\phi_s$ where $\phi_s \equiv \phi_{ss}(V_*)$ and at the reference velocity $V_*$ we also have $\mu = \mu_*$, equation (27) takes the form of the Bernoulli equation

$$b\sigma \psi' + (a\sigma/v) \psi + t_{ss} \phi_s k_s(v) \psi^2 = 0,$$

whose solution can be computed explicitly by introducing the change of variable $w = 1/\psi$ (e.g. see Polyarnin and Zaitsev, 2003) and is given by

$$\psi^{-1} = C v^{a/b} + \frac{t_{ss} \phi_s}{b\sigma} \int v (v/\bar{v})^{a/b} k_s(v) d\bar{v},$$

where $C$ is the constant of integration.

Further explicit computations are possible and of interest, in the case that the quasi-static critical stiffness $k_s(v)$ is constant. In this simplistic case we obtain $\psi^{-1} = C v^{a/b} + t_{ss} \phi_s k_s v/[b-a]\sigma$. Then, applying the condition that at the reference slip rate $v = V_*$ we have $\phi_s = \phi_{ss}(V_*)$ and $\mu = \mu_*$, we find that the constant $C$ takes the value $C = (1-t_{ss} \phi_s V_*/L)/V_*^{a/b}$. The solution of (28) for $\phi_{ss}(v)$ is then found to be

$$\phi^{-1}_{ss} = (1/\mu_*) - t_{ss} V_*/(v) V_*/V_*^{a/b} + (t_{ss} V_*/L) (v/V_*),$$

where it is natural to define the constant memory length $L = (b-a)\sigma/k_s$. Further assuming that the reference state satisfies $t_{ss} \phi_s = L/V_*$ we find that $\phi_{ss}^{-1}(v)$ is linear in $v$ and (22) simplifies to yield

$$t_{ss} \dot{\phi} = 1 - \frac{t_{ss} \phi v}{L},$$

which is exactly the Dieterich ageing evolution law (3).

However, this simple solution is not completely satisfactory because $\phi_{ss}$ is unbounded as $v \to 0$. To avoid this drawback, it is natural to require that the integrand in (29) remains finite such that

$$\int v (v/\bar{v})^{a/b} k_s(v) d\bar{v} \to b\sigma/t_{ss} \quad \text{as} \quad v \to 0.$$  

\footnote{We recall that the Dieterich–Ruina form (2) is well established experimentally over several decades in velocity and corresponds to a leading order approximation of the expression (18) derived from the thermodynamic theory for slip events. The analysis to follow is then relevant to (18)–(20).}
As a result, we construct a state variable that satisfies \( \phi_{ss}(v) \to 1^- \) as \( v \to 0 \) which implies that \( 0 \leq \phi \leq 1 \). This saturation of the equilibrium state variable towards vanishingly small velocities embeds our intuitive physical idea according to which the interface healing in stationary contact cannot be indefinite, as it is for the Dieterich law (3). The most important consequence of such a behaviour is to produce a local maximum of the steady-state friction coefficient due to logarithmic velocity strengthening of the instantaneous friction coefficient (2) (e.g. see Putelat et al., 2007, 2010, Grosch, 1963, for theoretical and experimental examples respectively). In addition, taking the opposite limit \( v \to \infty \) suggests selecting the constant of integration \( C = 0 \) and requiring a finite limit for the integral

\[
\int_0^v (v/\tilde{v})^{a/b} k_\star(\tilde{v}) \, d\tilde{v} \to \psi_\infty^{-1} \quad \text{as } v \to \infty,
\]

if we are to demand in addition that \( \phi_{ss} \) saturates at a positive value (instead of zero) at large velocity. This again breaks the monotonic behaviour of the steady-state Dieterich–Ruina expression (2) by introducing a local minimum as documented by Heslot et al. (1994) and Bureau et al. (2002), for instance.

Overall, a consistent expression of the equilibrium state variable \( \phi_{ss}(v) \) with the expression of the friction coefficient derived from the thermodynamic theory of section 2.2 is

\[
\phi_{ss}(v) = \frac{t_\phi}{t_{ss}}, \quad \text{where} \quad t_\phi(v) \equiv b\sigma / \int_0^v (v/\tilde{v})^{a/b} k_\star(\tilde{v}) \, d\tilde{v}
\]

(30)

corresponds to the interfacial state dynamic relaxation time. The equilibrium interfacial state is clearly a measure of the deviation of the magnitude of the dynamic interfacial rejuvenation of characteristic time scale \( t_\phi \) compared to the static ageing of the interface in stationary contact of characteristic time scale \( t_{ss} \). The smaller \( \phi_{ss} \) the faster the relaxation towards a steady state. We note that the interpretation of \( t_{ss} \) as a static ageing time scale lies upon the requirement that \( \phi_{ss}(v) \to 1 \) as \( v \to 0 \). In such a limit of quasi-stationary contact \( v \to 0 \), \( t_\phi \approx t_{ss} \) and \( \phi \to 1^- \) exponentially over \( t_{ss} \).

Moreover, the apparent memory length is just \( L_0(v) = v t_\phi(v) \). We see that an experimental measurement of the velocity dependence of \( L_0 \) will provide a direct measurement of the velocity dependence of \( \phi_{ss} \) independent from the information extracted from the measurement of the quasi-static critical stiffness \( k_\star \) and given by the method defined by (30).

### 4. Discussion

In this paper we discussed different microscopic theoretical arguments that underpin the rate-and-state formalism for friction. We compared two theories often quoted for deriving the rate-and-state dependence of the friction coefficient at low and moderate sliding velocities in order to clarify their microphysical and phenomenological foundations. Table 1 summarises the microphysical interpretations of the phenomenological material parameters of rate-and-state friction for the two theories.

In section 2 we focussed on the friction coefficient. In section 2.1 we discussed the constitutive theory of asperity deformation. This theory relies heavily on the Bowden and Tabor product decompositions (7)–(8) and the decomposition of the deformation of interfacial asperities into pure shear and simple shear thermally activated creep deformations. These are, respectively, held responsible for the logarithmic growth of the real
Figure 1: Experimental determination from the data of Heslot et al. of: (a) the equilibrium interfacial state \( \phi_{ss} = (V) \) from (25); (b) the apparent memory length (26). Only the six first experimental points from the left are relevant as they correspond to the quasi-static value of \( k_c \). A linear fit gives \( t_{ss} \phi_{ss} = 10^{0.436} \times (V/V_s)^{-0.723} \) and \( L_a = 10^{-5.564} \times (V/V_s)^{6.277} \) (dashed lines). Parameters: \( M = 1.2 \text{ kg}, \ g = 9.81 \text{ m s}^{-2}, \ a = 0.0349, \ b = 0.0489, \ V_s = 10^{-6} \text{ m}^{-1} \).

Contact area with time and for the logarithmic dependence of the asperity shear strength on velocity. The friction state dependence is consequently attributed to the time evolution of the real contact area, i.e. a first order correction of the initial real contact area, the interfacial state being interpreted as the corresponding contact age. The rate process involved in this point of view can be referred to when justifying the physical origin of motion of dislocations during a Nabarro–Herring type of creep.

In section 2.2 we discussed the thermodynamic theory for slip events. This theory makes use of transition-state theory and is based on the common experimental observation of the Arrhenius-type temperature dependence of the macroscopic frictional slip rate. Combined with the Bowden and Tabor postulate (9), the transition-state theory permits the expression of the macroscopic slip rate in terms of the friction coefficient and the energy barrier to surmount so that slip actually occurs. In contrast with the constitutive theory, it is this energy barrier which is state dependent and not the true contact area. Besides, the framework of the thermodynamic theory suggested that the state evolution of the energy barrier follows from the state evolution of the yield stress representing the asperity contact hardening promoted by the Nabarro–Herring creep flow of contacts initiated by the macroscopic interaction of the contacting solids.

From the point of view of the thermodynamic theory, it appears that the interpretation of the interfacial state variable in terms of a definite physical quantity is still too speculative with regard to the experimental data available about the microscopic evolution of frictional interfaces. We showed however that different theoretical proposals could lead in principle to the same logarithmic state evolution of the slip event energy activation. This result is of substantial interest, considering both the variety of physical and chemical processes and the span of time and length scales involved in tribology. It suggests indeed that the memory effect of friction could show the same state evolution despite having a variety of different microphysical origins depending on the system under consideration. In particular this would provide a possible justification for the ubiquity of the Dieterich–Ruina logarithmic velocity dependent behaviour of the friction coefficient in steady-state sliding.

Although it may thus be premature to firmly identify a microphysical mechanism responsible for the frictional memory effect, the formalism of the thermodynamic theory combined with the phenomenology of rate-and-state friction indicates that the transition-state theory may also be invoked to justify the state
evolution of the slip event energy of activation which determines the memory of frictional sliding. Overall the state dependence of the slip event activation energy results from microphysical mechanisms which combine themselves as a thermally activated process in such a way that the resulting collective behaviour mimics a state dependent frictional behaviour (equation (20)) setting the level of the yield stress which determines the probability of macroscopic slip events. We showed that a possible candidate for such a thermally activated process could follow from assuming a Nabarro-Herring creep for the asperities.

The fading memory effects of friction illustrated by the relaxation phenomena observed in response to velocity jumps indicate the time evolution of the interface state. A microscopic theory for the derivation of the interfacial state evolution law is still lacking. Despite this, a possible remedy is to model the interfacial state evolution by a first order kinetic equation which determines the relaxation of the interfacial state towards its equilibrium value \( \phi_{ss} \) in steady-state sliding. Such a kinetic process is in general defined by equation (22) and naturally introduces a characteristic time scale \( t_{ss} \) that we interpreted as the interfacial static ageing. Most importantly, using equation (25), we showed that the velocity dependence of the equilibrium interfacial state \( \phi_{ss} \) and the static ageing time \( t_{ss} \) can be identified experimentally from the measurement of the steady-state friction force and the critical stiffness of stick-slip oscillations emergence in a typical spring–block experiment performed for different driving velocity. Moreover, combined with the Dieterich–Ruina approximation of the friction coefficient deduced from the thermodynamic theory, the first order kinetics hypothesis allowed us to interpret the equilibrium interfacial state as the ratio of the kinetic rejuvenation characteristic time over the static ageing time. If the interfacial state were to be interpreted as a strain rate, \( \phi_{ss} \) would then compare the relative magnitude of the asperity strain rate \( \dot{\varepsilon}_{ss} \) in stationary contact with the dynamic asperity strain rate \( \dot{\varepsilon}_{s}(v) \) during sliding. As a result we define

\[
\phi_{ss}(v) \equiv \frac{\dot{\varepsilon}_{ss}}{\dot{\varepsilon}_{s}(v)} \quad \text{with} \quad \dot{\varepsilon}_{s}(0) = \dot{\varepsilon}_{ss},
\]

where

\[
\dot{\varepsilon}_{s}(v) = \int^{v}(v/\bar{v})^{n/b} k_{s}(\bar{v}) \, d\bar{v}/(b\sigma).
\]

This last expression suggests an interpretation of the integral as the power dissipated by the normal force: \( \sigma \dot{\varepsilon}_{s}(v) \).

Finally we believe the thermodynamic theory for slip events provides a more robust framework to justify the microphysical foundations of rate-and-state friction than the constitutive theory of asperity deformation. We point out that the thermodynamic theory avoids the reference to the Bowden and Tabor hypotheses (7)-(8) which appear to us a strong constraint on the actual mechanisms of asperity interactions. Instead the thermodynamic theory relies on the Bowden and Tabor postulate (9) which is supported by the experimental observations of Berman et al. (1998). Besides, the thermodynamic theory is also favoured by our recent theoretical findings (Putelat et al., 2010), according to which friction laws derived from the constitutive theory predict, at low driving velocity, a subcritical Hopf bifurcation for the onset of stick-slip oscillations for a spring–block system, instead of the supercritical Hopf bifurcation as experimentally reported by Heslot et al. (1994). Nevertheless, the reasoning of the constitutive theory is interesting and could maybe be relevant for deriving friction laws in other tribological contexts.

In the meantime, we suggest that careful experimental studies à la Heslot and direct optical observations of frictional interfaces, as in Dieterich and Kilgore (1994), would be strongly desirable in order to assess further the validity of the thermodynamic theory of slip events. In particular, the macroscopic and phenomenological identification of the state evolution law is crucial in order to constrain the microphysical foundations of rate-and-state friction. It is an important matter of current debate in the geophysical community in relation to earthquake source mechanics. We have shown here that assuming a first order kinetic equation allows a clear definition of the equilibrium interfacial state and the frictional memory length that can be experimentally determined from the nonlinear dynamics of a spring–block system and the methods described in section 3.

Acknowledgements

JHPD gratefully acknowledges financial support from the Royal Society through a University Research Fellowship.
<table>
<thead>
<tr>
<th>Constitutive theory</th>
<th>thermodynamic theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>$b$</td>
</tr>
<tr>
<td>$\mu_s k_B T/E_s$</td>
<td>$\mu_s k_B T/E_L$</td>
</tr>
<tr>
<td>$k_B T/(\Omega_L \sigma_Y)$</td>
<td>$\mu_s k_B T/(\Omega_L \sigma_Y)$</td>
</tr>
<tr>
<td>$k_B T/(\Omega_s \sigma_c)$</td>
<td>$k_B T/(\Omega_s \sigma_c)$</td>
</tr>
<tr>
<td>$\mu_l k_B T/E_s$</td>
<td>$\mu_l k_B T/E_r$</td>
</tr>
</tbody>
</table>

Table 1: The microphysical interpretations of the phenomenological material parameters of rate-and-state friction. For the two theories, the first line corresponds to the definitions of the coefficients as they naturally appear from the formalism, the second line corresponds to alternative definition obtained from combinations between definitions. We point out that the energy of activation for slip events $E_s$ and the yield stress $\tau_Y$ are state dependent in the thermodynamic theory whereas there are constant in the constitutive theory. We denote $\mu_l = \mu_s + b \ln(\phi/\phi_s)$.

References


