Perspective

Spatiotemporal Dynamics of Frictional Systems: The Interplay of Interfacial Friction and Bulk Elasticity

Yohai Bar-Sinai 1,2, Michael Aldam 1, Robert Spatschek 3, Efim A. Brener 3,4 and Eran Bouchbinder 1,*

1 Chemical and Biological Physics Department, Weizmann Institute of Science, Rehovot 7610001, Israel; ybarsinai@gmail.com (Y.B.-S.); michael.aldam@weizmann.ac.il (M.A.)
2 John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
3 Institute for Energy and Climate Research, Forschungszentrum Jülich, D-52425 Jülich, Germany; r.spatschek@fz-juelich.de (R.S.); e.brener@fz-juelich.de (E.A.B.)
4 Peter Grünberg Institut, Forschungszentrum Jülich, D-52425 Jülich, Germany
* Correspondence: eran.bouchbinder@weizmann.ac.il

Received: 7 August 2019; Accepted: 14 October 2019; Published: 17 October 2019

Abstract: Frictional interfaces are abundant in natural and engineering systems, and predicting their behavior still poses challenges of prime scientific and technological importance. At the heart of these challenges lies the inherent coupling between the interfacial constitutive relation—the macroscopic friction law—and the bulk elasticity of the bodies that form the frictional interface. In this feature paper, we discuss the generic properties of a minimal macroscopic friction law and the many ways in which its coupling to bulk elasticity gives rise to rich spatiotemporal frictional dynamics. We first present the widely used rate-and-state friction constitutive framework, discuss its power and limitations, and propose extensions that are supported by experimental data. We then discuss how bulk elasticity couples different parts of the interface, and how the range and nature of this interaction are affected by the system’s geometry. Finally, in light of the coupling between interfacial and bulk physics, we discuss basic phenomena in spatially extended frictional systems, including the stability of homogeneous sliding, the onset of sliding motion and a wide variety of propagating frictional modes (e.g., rupture fronts, healing fronts and slip pulses). Overall, the results presented and discussed in this feature paper highlight the inseparable roles played by interfacial and bulk physics in spatially extended frictional systems.

Keywords: interfacial constitutive laws; rate-and-state friction; elastodynamics; onset of sliding motion; rapid slip nucleation; rupture fronts; slip pulses

1. Introduction

The dynamics of frictional systems have been the focus of intense scientific research in the last few decades [1–3], with applications ranging from atomic and nano-scale systems [4–6], through engineering and control systems [7,8], to geophysical systems [9–12]. Despite the progress made in understanding and controlling friction at different scales [13–16], various basic aspects of frictional systems are not yet well-understood. As frictional systems and processes span a wide range of spatiotemporal scales, the problem can be approached from different perspectives. The focus of the present feature paper is on spatially extended frictional systems approached from a macroscopic (continuum) perspective. Our major goal is to discuss the generic properties of minimal macroscopic interfacial constitutive relations, within the rate-and-state friction framework [9,14,17–21], and to
demonstrate that the overall dynamics of frictional systems emerge from the inherent coupling between interfacial tribology and the bulk elasticity of the bodies forming the interface. As such, the physics and dynamics of frictional systems are highlighted to result from the inherent interplay between the local, micro-scale tribological behavior of frictional interfaces and the non-local, long-range interaction between different parts of the interfaces, which are mediated by macro-scale bulk elasticity (cf. Figure 1).

To see the nature and origin of this inherent coupling let us consider two bodies in frictional contact and the associated slip displacement \( \delta \), which is the difference between the displacement of the upper and lower bodies \( u^\pm \), respectively, at the interface (cf. Figure 1). In spatially extended frictional systems, \( \delta \) generically varies along the interface, and hence should be in fact represented by a field \( \delta(x,t) \), where \( x \) is the coordinate along the interface that is located at \( y = 0 \) and \( t \) is the time (cf. Figure 1 and note that we do not consider the \( z \) direction, which is nevertheless marked in the figure). A coarse-grained constitutive law relates the frictional stress \( \tau(x,t) \) (sometimes termed the frictional strength/resistance) to quantities such as the slip displacement field \( \delta(x,t) \), the slip velocity field \( v(x,t) = \partial_t \delta(x,t) \), the interfacial normal (compressive) stress \( \sigma(x,t) \) and possibly other interfacial fields, and can be expressed as

\[
\tau(x,t) = f[\delta(x,t), \partial_t \delta(x,t), ...] \sigma(x,t) .
\]

Here the dots represent yet unspecified internal state fields, which satisfy their own evolution equations. The latter, together with the functional \( f[\cdot] \) that appears as a proportionality factor (“friction coefficient”) between \( \sigma(x,t) \) and \( \tau(x,t) \), define the coarse-grained constitutive relation. The purely local nature of \( f[\cdot] \) is manifested by the absence of spatial derivatives of \( \delta(x,t) \) in it. When bulk elasticity is not taken into account, i.e., when the bulks are considered infinitely rigid as in spring-block models [14,21,22], Equation (1) takes the form \( \tau_0 = f \sigma_0 \), where \( \tau_0 \) and \( \sigma_0 \) are the far-field applied shear stress and normal (compressive) stress, respectively (cf. Figure 1). In this case, the frictional problem has a single degree of freedom, namely \( \delta(t) \). A major goal of this feature paper is to understand how spatial degrees of freedom, emerging from bulk elasticity, change/extend the relation \( \tau_0 = f \sigma_0 \) and to demonstrate their far-reaching implications for frictional dynamics.

When the elasticity of the bodies forming the interface is taken into account, Equation (1) serves as a boundary condition that relates two components of the stress tensor field \( \sigma_{ij}(x,y,t) \) as the interface is approached, \( y \to 0 \). In particular, \( \tau(x,t) \) equals the bulk shear stress, \( \sigma_{xy}(x,y,t) \), and \( \sigma(x,t) \) equals the bulk normal (compressive) stress, \( -\sigma_{yy}(x,y,t) \), in the limit \( y \to 0 \). The interfacial shear stress and normal (compressive) stress, \( \sigma_{xy}(x,y=0,t) \) and \( -\sigma_{yy}(x,y=0,t) \), depend on the far-field applied shear stress and normal (compressive) stress, \( \tau_0 \) and \( \sigma_0 \) respectively, and on stresses that emerge from all processes that are mediated by the presence of the bulk of the bodies forming the interface. The latter depend on gradients of the slip displacement field, \( \partial_\delta \delta(x,t) \), and on the slip velocity field, \( v(x,t) = \partial_\delta \delta(x,t) \), and can be expressed as \( F_x[\partial_\delta \delta(x,t), \partial_t \delta(x,t)] \) and \( F_y[\partial_\delta \delta(x,t), \partial_t \delta(x,t)] \), respectively. In case the relevant bulk physics is faithfully represented by linear elastodynamics, as is widely assumed elsewhere and as is also adopted here, the interfacial stresses take the form

\[
\sigma_{xy}(x,y=0,t) = \tau_0 - F_x[\partial_\delta \delta(x,t), \partial_t \delta(x,t)] , \quad -\sigma_{yy}(x,y=0,t) = \sigma_0 - F_y[\partial_\delta \delta(x,t), \partial_t \delta(x,t)] .
\]

The linear functionals \( F_x[\partial_\delta \delta(x,t), \partial_t \delta(x,t)] \) and \( F_y[\partial_\delta \delta(x,t), \partial_t \delta(x,t)] \) generically describe long-range spatiotemporal interactions (i.e., note the explicit appearance of spatial gradient in \( \partial_\delta \delta(x,t) \), which schematically represents also higher order spatial derivatives and even a spatial integral) and interface-bulk radiation effects. They are affected by dimensionality, by the geometry of the bodies forming the interface and by their constitutive parameters (cf. Figure 1).
Figure 1. A schematic of a frictional system. Two macroscopic bodies are in contact, the interface between them is composed of a discrete collection of contact asperities (see zoom in). Here we plot the general case in which the bodies are made of different materials (hence the different colors used) and have different geometries. The system is loaded normally and tangentially ($\sigma_0$ and $\tau_0$, respectively) far from the interface. Stress conditions at the interface, which give rise to frictional motion, result from the interplay between local frictional rheology (see zoom in) and forces mediated by the elastic bulks, induced by the loading and by differential sliding in distant parts of the interface (see curved double-arrow). The height $H$ and length $L$ of one of the bodies are marked and a coordinate system is introduced. The local slip displacement $\delta$, the difference between the local displacement of the upper/lower bodies ($u_\pm$), is also schematically denoted.

Using the equalities $\tau(x, t) = \sigma_{xy}(x, y = 0, t)$ and $\sigma(x, t) = -\sigma_{yy}(x, y = 0, t)$, Equation (1) takes the form

$$\tau_0 - \mathcal{F}_\tau[\partial_x \delta(x, t), \partial_t \delta(x, t)]_\tau = f[\delta(x, t), \partial_t \delta(x, t), ...] (\sigma_0 - \mathcal{F}_\sigma[\partial_x \delta(x, t), \partial_t \delta(x, t)])$$

which sets the framework for the discussion in this paper. This basic equation clearly demonstrates that frictional dynamics, which correspond to solutions for $\delta(x, t)$, inherently depend on both the interfacial constitutive relation $f[\delta(x, t), \partial_t \delta(x, t), ...]$ (along with the internal state fields and their own evolution equations, to be discussed below), and on the bulk constitutive contributions $\mathcal{F}_\tau[\partial_x \delta(x, t), \partial_t \delta(x, t)]$ and $\mathcal{F}_\sigma[\partial_x \delta(x, t), \partial_t \delta(x, t)]$. Consequently, the dynamics of frictional systems inseparably emerge from the coupling between interfacial and bulk physics, as encapsulated in Equation (3).

In the following parts of the paper we first discuss the generic properties of the interfacial constitutive relation $f[\delta(x, t), \partial_t \delta(x, t), ...]$, including internal state fields and their evolution equations, and then consider the bulk constitutive contributions $\mathcal{F}_\tau[\partial_x \delta(x, t), \partial_t \delta(x, t)]$ and $\mathcal{F}_\sigma[\partial_x \delta(x, t), \partial_t \delta(x, t)]$, for various dimensionalities, physical situations and geometric configurations relevant for a broad range of frictional systems. While we build on extensive effort and progress accumulated in the literature over the last few decades, we are also strongly biased toward our own recent work. As such, this paper inevitably presents our personal perspective on the issues at stake. Yet, we hope that the emerging physical picture is sufficiently broad and general to be useful for scientists interested in a wide variety of frictional systems. Whenever possible, we try to highlight basic concepts and generic physical effects, while minimizing the technical complexity of the presentation and referring the readers to a more technically involved literature.
2. Generic Properties of Interfacial Constitutive Laws

A widely used class of interfacial constitutive laws, termed rate-and-state friction (RSF) models, originated from the seminal works of Dieterich [17,18,23] and Rice and Ruina [20,24] in the late 1970s, which elaborated on ideas formulated by Bowden and Tabor [1] and Rabinowicz [25] in the 1950s. These models introduce one or more coarse-grained internal state variables to account for the evolution of the interfacial structural state (and hence they carry memory of the interfacial history), and rate-dependence to account for interfacial rheology. In this section, we will briefly describe the conventional RSF model, discuss what we view as its power and possible shortcomings, and consequently propose modifications. Our ultimate goal is to provide an overview of the generic properties of interfacial constitutive laws.

2.1. Conventional RSF Models

It is now well-established that a crucial physical quantity affecting the dynamics of spatially extended dry frictional interfaces is the real contact area $A_r$, which is typically orders of magnitude smaller than the nominal (apparent) contact area $A_n$ [14,26–29], i.e., a spatially extended interface separating two macroscopic bodies in frictional contact is composed of a collection of discrete contact asperities (i.e., it is a multi-contact interface, cf. Figure 1) whose total area is typically much smaller than the nominal overlap of the bodies. A major implication of this physical picture is that the frictional stress $\tau$ can be written, following Bowden and Tabor [1], as $\tau = (A_r/A_n) \tau_s(v)$, where $\tau$ is the frictional stress of Equation (1), $\tau_s(v)$ is a rate-dependent asperity-level shear stress and $v$ is the interfacial slip velocity. The normalized real contact area, $A ≡ A_r/A_n$, is commonly expressed in terms of another state variable $\phi$, of time dimensions, as

$$A(\phi) = A_r/A_n = \frac{\sigma}{\sigma_n} \left[ 1 + b \log \left( \frac{\phi}{\phi^*} \right) \right]. \quad (4)$$

Here $\sigma$ is the interfacial normal stress of Equation (1), $\sigma_n$ is the material hardness [14,27,30] and $b$ is a dimensionless material parameter, typically of the order of $10^{-2}$. In this formulation, $\phi^*$ is a rather arbitrary time scale required for dimensional consistency. Below we discuss this point in detail.

Equation (4) properly describes the outcome of a broad range of experimental works that have demonstrated that the real contact area (or the static frictional resistance, which is proportional to $A$ according to the Bowden and Tabor relation $\tau = A(\phi)\tau_s(v)$) grows logarithmically with the time of contact in the absence of sliding [27,31–35]. This is the well-known logarithmic aging of the static frictional resistance. Thus, Equation (4) allows one to write the simple relation $\partial_t \phi = 1$ in the absence of sliding, $v = 0$. Within this physical interpretation, $\phi$ is the typical lifetime of a contact, and therefore, if the typical slip distance required to destroy an ensemble of contacts is $D$, we expect $\phi(v) = D/|v|$ under persistent (steady-state) sliding, $v \neq 0$. This implies that $A$ decreases logarithmically with increasing sliding velocity $v$, in agreement with the prevalent observation that the steady-state frictional resistance (to be discussed below), for a broad range of velocities, decreases logarithmically with $v$ [17,36–39]. To actually show this, one should consider $\tau_s(v)$, which we do next.

The rate-dependent contribution to the frictional resistance, $\tau_s(v)$, is usually associated with a stress-biased thermal activation process [14,38,40–43]. Within a simplified model of a single energy barrier $E$, linearly biased by the applied stress according to $E = E_0 - \Omega \tau$, where $\Omega$ is an activation volume and $E_0$ is the bare energy barrier, $\tau_s(v)$ takes the form [40,43]
Here \( k_B \) is Boltzmann’s constant, \( T \) is the temperature and \( v_c \) is a velocity scale setting the upper limit for the validity of the thermal activation model. Combining Equations (4) and (5), we obtain

\[
\tau(\phi, v) = \frac{k_B T}{\Omega} \sinh^{-1} \left( \frac{v}{2 v_c} \exp \left[ \frac{E_0}{k_B T} \right] \right). \tag{5}
\]

In the limit where \( E_0 \gg k_B T \), which is quite typical, it can be shown that this equation takes the well-known form of RSF friction \([40,43]\)

\[
f(\phi, v) \simeq f_0 + \alpha \log \left( \frac{v}{v_c} \right) + \beta \log \left( \frac{\phi}{\phi^*} \right), \tag{7}
\]

where \( f(\phi, v) = \tau(\phi, v)/\sigma \) as in Equation (1), higher order logarithmic terms have been omitted, and the dimensionless parameters \( f_0, \alpha \) and \( \beta \) can be expressed in terms of physical parameters defined above (see Appendix A.1).

The second logarithmic contribution on the right-hand side of Equation (7) emerges from the evolution of the real contact area, i.e., the state-dependent contribution of Equation (4). The first logarithmic contribution, on the other hand, emerges from the interfacial rheology, i.e., the rate-dependent contribution of Equation (5). This logarithmic \( v \)-dependent behavior is directly supported by experiments in which the slip velocity \( v \) changes abruptly, such that \( \phi \) does not change appreciably, and the change in the frictional stress is measured on a short time scale. Under steady-state conditions, \( \phi \) is slaved to the slip velocity according to \( \phi = D/|v| \), and the overall rate-dependence of \( f(\phi = D/|v|, v) = f_{\text{oa}}(v) \) is proportional to \( \log |v| \), as mentioned above. This logarithmic dependence is velocity-weakening for \( \alpha < \beta \) and velocity-strengthening for \( \alpha > \beta \). This issue will be further discussed in Section 2.2.2. Finally, the RSF model is fully defined once a dynamical evolution equation for \( \phi \), which connects the no-sliding \( \phi = t \) and sliding \( \phi = D/|v| \) fixed-points, is provided. This is commonly done through the popular “aging law” \([9,23]\)

\[
\dot{\phi} = 1 - \frac{\phi|v|}{D}, \tag{8}
\]

which together with Equation (7), constitute the conventional RSF model. Other \( \phi \) evolution equations, \( \dot{\phi} = \Phi(\phi|v|/D) \) with other functions \( \Phi(\cdot) \), are considered and studied in the literature \([20,44–46]\), but are not discussed here. Finally, note that a single internal state field model is certainly a minimal one, and in some cases additional internal states fields might be considered for a more physically faithful description of various tribological processes.

### 2.2. Limitations of Conventional RSF Models and Proposed Modifications

RSF models have been quite extensively used in macroscopic modeling of frictional phenomena. Yet, we believe that while these models are quite successful in capturing many physical aspects of frictional phenomenology, they are also deficient in various important ways. Our goal here is to highlight these limitations and to propose modifications to be incorporated into an extended friction model.
2.2.1. Linear Reversible Response and an Extended RSF Model

Conventional RSF models predict the frictional stress, cf. Equation (6), which should provide physically meaningful information both in the presence of sliding \((v \neq 0)\) and in its absence \((v = 0)\). Using conventional terminology, it should describe both static and dynamic friction. However, Equation (6) predicts \(\tau(\phi, v = 0) = 0\) in the limit \(v \to 0\), i.e., it predicts vanishing static friction. This happens because \(\tau_s(v)\) in Equation (5) corresponds to a Newtonian fluid behavior at small sliding velocities, i.e., \(\tau_s(v \to 0) \propto v\). One way in which this problem is circumvented in the literature is to change the \(v\)-dependent term in Equation (7) to be a \(\log(1 + v/v_c)\), such that it vanishes for \(v = 0\), together with interpreting the remaining part, \(f_0 + \beta \log(1/\phi^\alpha)\) (note that \(\phi = t\) for \(v = 0\)), as an upper threshold, i.e., in this case \(f(\phi, v = 0)\) does not correspond to the frictional stress, but rather to the threshold for the onset of sliding. Mathematically, it means that \(f(\phi, v = 0)\) is used to define an inequality imposed on a frictional stress that is calculated from other considerations (in fact, from the bulk shear stress \(\sigma_{xy}\) of Equation (2)). At the same time, \(f(\phi, v)\) of the very same Equation (7), is interpreted as the actual frictional stress for \(v \neq 0\), implying that \(f(\phi, v)\) has a different physical meaning for \(v = 0\) and for \(v \neq 0\).

Another common way to circumvent the vanishing static friction problem is by forcing the frictional interface to slide with a small background velocity \(v \neq 0\) and use Equation (7) as is \([45,47]\). While these ad hoc solutions to the stated problem might provide a practical way out under some circumstances, from a basic physics perspective we expect \(\tau(\phi, v)\), or \(f(\phi, v)\), to continuously describe the frictional resistance stress under both static and dynamic conditions. In particular, it should properly account for the transition between the two behaviors, e.g., for the onset of sliding and for sliding direction reversal (which is relevant for cyclic frictional motion, appearing in many engineering systems).

In our view, these problems originate from the fact that the conventional RSF model does not feature a linear reversible response at low stresses, corresponding to the elastic (reversible) deformation of contact asperities (note that as explained above, the model does feature a Newtonian viscous behavior, i.e., a linear irreversible response). The existence of a reversible response has been experimentally demonstrated \([22,48–51]\), see also Figure 2a here, and is expected on general grounds \([52]\), i.e., we expect the asperity-level stress \(\tau_s\) to be finite and reversible under small static shear loading conditions due to the elastic deformation of contact asperities. This should result in \(\tau(\phi, v)\) that is valid for both \(v = 0\) and \(v \neq 0\), as well as for \(v\) that changes sign dynamically, which should resolve the problems with conventional RSF models stated above. Indeed, some authors have incorporated an interfacial elastic response in their modeling efforts \([51,53–61]\) and noted its potential importance for various frictional phenomena, such as the onset of sliding \([22]\) and rupture mode selection \([55]\). Here we describe an extended RSF-type model that incorporates a linear reversible response at low stresses. The model has been briefly introduced and studied in previous publications \([58,60,61]\).

The basic idea is to incorporate into the RSF framework an elastic interfacial response at small stresses, within a Kelvin-Voigt-like viscoelastic model \([62,63]\) together with threshold dynamics. To that aim, we reinterpret the asperity-level stress in Equation (5) as a viscous-like stress \(\tau_{vis}(\phi, v)\), which is a partial stress constituting only one contribution to the total asperity-level stress \(\tau_s\). The complementary contribution is the asperity-level elastic stress \(\tau_{el}^s\), such that \(\tau_s = \tau_{vis} + \tau_{el}^s\). While this way of introducing asperity-level elasticity into the friction model is not unique, it suffices for our present purposes, as will be shown below. Using the Bowden and Tabor relation \(\tau = A(\phi)\tau_s(v)\) to translate the asperity-level stress to a coarse-grained stress, we obtain

\[
\tau(\phi, v, \tau_{el}) = \tau_{el} + \tau_{vis}(\phi, v),
\]

where, as explained above, \(\tau_{vis}(\phi, v) = A(\phi)\tau_s(v)\) with \(\tau_s(v)\) of Equation (5).
In Equation (9), $\tau^{el}$ is a dynamic variable by itself, in addition to $\phi$ and $v$. Therefore, we need to derive its evolution equation. As explained above, at small slip displacements we expect the interface to respond linear elastically. To capture this basic physical behavior, we treat the interface as a boundary layer of height $h$, much smaller than the macroscopic size $H$ of the system, which is characterized by an effective shear modulus $\mu_0$. This effective modulus may be different from the shear modulus $\mu$ of the bulk due to the existence of stress-free surfaces along the interface (i.e., surface roughness). Under deformation, the total shear accumulated in the boundary layer is $\delta/h$, producing a shear stress $\mu_0\delta/h$. Such an elastic response of the interface has been directly measured [48,50,64], as shown in Figure 2a, and has been predicted theoretically [52]. Defining a coarse-grained elastic stress $\tau^{el}$ through $A_h \tau^{el} = A_t \mu_0 \delta/h$ and differentiating it with respect to time, we obtain $\dot{\tau}^{el} = \mu_0 A \dot{v}/h$ (where we assumed that $A$ varies much slower than $\tau^{el}$ such that its time-derivative can be neglected). The term $\mu_0 A \dot{v}/h$ constitutes only one contribution to $\tau^{el}$; the other contribution is discussed next.

Irreversible sliding drives contact annihilation and renewal, and hence is accompanied by a reduction in the average elastic stress stored in the contact asperities. For simplicity, we assume a linear form for this reduction, which is similar to that of $\phi$ in Equation (8). Thus, the evolution equation of $\tau^{el}$ is taken to be

$$\dot{\tau}^{el} = \frac{\mu_0}{h} A \dot{v} - \tau^{el}\frac{v}{D}\partial_{\tau^{el}} g(\tau, v),$$

(10)

where the function $g(\tau, v)$ controls the transition between the reversible and irreversible regimes. In fact, the very same function should appear in a modified version of Equation (8), which takes the form $\dot{\phi} = 1 - \phi|v| g(\tau, v)/D$ and is adopted here. When $g \ll 1$, the system responds elastically, i.e., $\tau^{el} = \mu_0 A \delta/h$. When $g(\cdot) \sim O(1)$, irreversible stress and area reduction take place. Therefore, $g(\cdot)$ plays the role of an effective threshold for the onset of irreversible slip, as in conventional elasto-plastic models that feature a yield stress. Explicit expressions for $g(\cdot)$ will be discussed below in relation to various applications.

To directly demonstrate the physical relevance and importance of interfacial elasticity, we focus on the experimental results of [50], which are reproduced here in Figure 2a. In these experiments, the multi-contact interface between two PMMA blocks has been subjected to two time-dependent shear loading-unloading protocols below the ‘static’ threshold for gross slip (which for this experimental system corresponded to a shear force of $F_{static} = 10N$). In the first protocol (the squares in Figure 2a, equally spaced in time), the shear force $F_s(t)$ is externally increased at a constant (and small) rate up to 4N (i.e., about 40% of the nominal ‘static’ friction threshold) and then decreased back to zero at the same small rate. The observed response, i.e., the slip displacement $\delta(t)$, is nearly linear in the driving force $F_s(t)$, featuring very little dissipation (i.e., hysteresis) and almost full recovery of the initial state. These observations explicitly demonstrate the existence of a linear reversible (elastic) response. In the second protocol (the circles in Figure 2a, equally spaced in time), the shear force $F_s(t)$ is externally increased at a constant (and small) rate up to 9N (i.e., about 90% of the nominal ‘static’ friction threshold), it is then maintained at this level for sufficiently long time and eventually it is decreased back to zero at the same small rate. Under this load-hold-unload protocol, the slip displacement $\delta(t)$ is no longer linear in the driving force $F_s(t)$ for values larger than $\sim 7N$. Moreover, when $F_s(t)$ is kept at its maximal value (here of 9N), $\delta(t)$ exhibits creep dynamics that may indicate arrest after a finite time (as suggested by the increasing density of the circles). Finally, as the shear force is reduced back to zero, a finite residual (irreversible) slip remains. This behavior corresponds to a viscoelastic response with threshold dynamics.
Figure 2. Frictional dynamics under small stresses and the existence of a linear reversible frictional response. (a) Experimental data, taken from Figure 1 of [50]. The shear force is shown as a function of slip displacement, with points equally spaced in time. The squares show the loading-unloading protocol at very low stresses, whereas the circles show the load-hold-unload protocol, still below the static friction threshold (see Section 2.2.1). The former exhibit almost entirely linear reversible response, with essentially no hysteresis. The latter exhibit a transition to irreversible (dissipative) slip, creep dynamics (during the hold phase) that may indicate arrest after a finite time (as suggested by the increasing density of the circles) and a residual slip upon unloading. (b) Our friction model, which includes a linear reversible frictional response, semi-quantitatively reproduces the experimental observations (see text and Appendix A.2 for details).

The friction model discussed above, which incorporates an interfacial elastic response at small driving forces, naturally and semi-quantitatively reproduces these experimental observations. To see this, we neglect all bulk degrees of freedom, whose role has been deliberately minimized in the experiments of [50] in order to isolate the interfacial response (this is the only place in this paper where bulk degrees of freedom are neglected), and focus on the equation

\[ F_s(t) = F_N f(\phi, \delta, \tau_{el}) = A_n \tau(\phi, \delta, \tau_{el}) . \] (11)

Here \( F_s(t) \) is the applied shear force (i.e., corresponding to each one of the two protocols discussed above), \( F_N \) is the constant applied normal force and \( A_n \) is the nominal (macroscopic) contact area (such that \( c_0 = F_N / A_n \) is the normal stress, not used here). \( \tau(\phi, \delta, \tau_{el}) \) of Equation (11) is given in Equation (9), \( \tau_{el} \) evolves according to Equation (10) and \( \phi \) according to \( \dot{\phi} = 1 - |v|g(\tau, v) / D \) (the choice of \( g(\cdot) \) is discussed in Appendix A.2). The solutions to these equations for the two experimental protocols discussed in relation to Figure 2a are presented in Figure 2b (see additional details in Appendix A.2), reproducing all of the experimental observations semi-quantitatively (no attempt has been made to quantitatively reproduce the experimental data). This example concludes the discussion of the first limitation we identify with RSF models, i.e., the absence of a linear reversible response at small stresses.

2.2.2. Short Time Cutoff and the Steady-State Friction Curve

The second problem we identify with conventional RSF models is related to Equation (4). Imagine that a fresh interface is formed by bringing two bodies into frictional contact. In the absence of irreversible slip, for which \( \phi = 1 \), the age \( \phi \) can be replaced by the time \( t \) elapsed from the formation of the interface. Equation (4) then predicts that the contact area will grow logarithmically with time,
as observed in a broad range of experiments [35,37,65,66]. However, when \( t \to 0 \) it becomes singular, which is clearly unphysical. Evidently, a short time cutoff exists and must be incorporated. A simple way to do so is to replace Equation (4) by

\[
A(\phi) = \frac{\sigma}{\sigma_\text{ss}} \left[ 1 + b \log \left( 1 + \frac{\phi}{\phi^*} \right) \right].
\]  

(12)

With this modification, \( \phi^* \) is no longer an arbitrary time scale—it is an intrinsic interfacial/material parameter which can be interpreted as a typical cutoff time for the onset of logarithmic aging. The existence of this cutoff has been directly verified experimentally [35,65,67]; it has been suggested by Dieterich [17,18], mentioned by Baumberger, Caroli and coworkers [14,34,38] and extensively discussed recently in [43,60].

While this modification might appear at first glance rather innocent, it has major implications for persistent sliding because the short time cutoff is translated, through the steady-state relation \( \phi = D/|v| \), to a high-velocity cutoff. To see this, note that Equation (12) predicts that \( A \) decreases logarithmically with the steady-sliding velocity \( v \). The presence of a short time cutoff implies that this decrease saturates at velocities \( \gtrsim D/\phi^* \). At higher velocities, the real contact area approaches a constant, and the velocity dependence of friction is dominated by \( \tau_\text{r}(v) \). Since the latter is intrinsically an increasing function of \( v \) (cf. Equation (5)), friction becomes \textit{velocity-strengthening} in this range of velocities. Put differently, the short time cutoff makes the steady-sliding frictional resistance a non-monotonic function of the sliding velocity, reaching a minimum at \( v_{\text{min}} \sim D/\phi^* \). As will be thoroughly discussed below, the crossover from velocity-weakening to velocity-strengthening has important implications on many aspects of the spatiotemporal dynamics of frictional systems.

We conclude this discussion of conventional RSF models and their extensions (the complete set of equations defining the constitutive law is collected in Appendix A.1) by stating what we perceive to be a generic picture of the steady-sliding friction curve. The latter, denoted by \( f_{\text{ss}}(v) \) and derived in Appendix A.1, is schematically plotted in Figure 3. The function \( f_{\text{ss}}(v) \) has a generic “N-like” shape, composed of three distinguishable regimes. First, at very small slip velocities, the frictional stress increases linearly with \( v \), corresponding to a viscous-like creep motion [30]. We note that at these low velocities, dynamically reaching the true steady state might be very difficult/long because the relaxation time scales as \( D/|v| \) (though there exist experimental measurements of this regime [68]). This velocity-strengthening branch (i.e., an increase in the frictional resistance with increasing steady-state sliding velocity) corresponds in practical terms to ‘stick conditions’, i.e., to situations in which the frictional interface is essentially locked-in; the reason for this is either the extremely low slip velocities in this regime or that it takes so long to reach this low-velocity steady state that the interfacial response is in fact predominantly elastic (representing an extremely long transient). This regime is terminated by a maximum of the curve, followed by a logarithmic velocity-weakening regime (corresponding to the \( \beta < \alpha \) case discussed in Section 2.1 in relation to Equation (7)). Finally, the curve reaches a minimum around \( v_{\text{min}} \sim D/\phi^* \) and then increases logarithmically (i.e., follows a velocity-strengthening behavior, as discussed above).

The validity of the N-shaped steady-state friction curve of Figure 3 is supported by experimental data for a broad range of materials [43] and its implications for the dynamics of frictional systems will be reviewed below. We note that while we consider here logarithmic velocity-strengthening above the minimum of the steady-state curve, strengthening can in fact become stronger than logarithmic, as discussed in [14,43,69]. We also note that at yet higher slip velocities (compared to those shown in Figure 3), the heat generated by frictional dissipation might not have enough time to diffuse away from the interface, leading to thermal softening accompanied by a substantial drop in the frictional resistance [70,71]. This regime is not shown in Figure 3 and is not further discussed in this paper. Finally, note that the velocity-weakening and post-minimum velocity-strengthening regimes of the steady-state friction curve are still approximately described by Equation (7) once \( \beta \log(\phi/\phi^*) \) is replaced by \( \beta \log(1 + \phi/\phi^*) \) (following Equation (12)), but with redefined \( f_0, \alpha \) and \( \beta \) (see Appendix A.1).
For $\alpha > \beta$, the steady-state friction curve is velocity-strengthening for all slip velocities (again, if extremely high-velocity thermal softening is excluded from the discussion).

Figure 3. The generic form of steady-state frictional resistance $f_{ss}(v)$ (see text and Appendix A.1 for details). At extremely low sliding velocities (yellow-shaded), steady-state friction is velocity-strengthening. This regime, associated with $v \approx 0$ in the figure, corresponds to stick conditions (effectively no-sliding) either because the slip velocity is minute or because steady state is dynamically inaccessible due to extremely long relaxation timescales (see text for additional discussion). At intermediate velocities (purple-shaded), friction is velocity-weakening. Finally, at higher velocities (green-shaded), above a local minimum occurring at $v_{\text{min}}$ (corresponding to a frictional resistance $\tau_{\text{min}}/\sigma_0$), friction becomes velocity-strengthening again. When loaded at stresses higher than this minimal value, $\tau_0 > \tau_{\text{min}}$ (horizontal line), three homogeneous solutions (fixed-points) exist; two stable (on the velocity-strengthening branches, marked by squares) and one unstable (on the velocity-weakening branch, marked by a circle), see text for additional discussion.

It is important to note that the tribological processes that give rise to the generic N-shaped steady-state friction curve of Figure 3 are diverse and may vary from one frictional system to another. These processes may include wear and debris formation, multi-scale plastic deformation, chemical effects that involve both intrinsic processes and interactions with the environment, among others. Studying these system-dependent processes is needed to understand quantitative aspects of the steady-state friction curve and its detailed functional form. Moreover, such studies are important for elucidating the possible need for additional internal state fields. Yet, we believe that a lot can be learned about the behavior of frictional systems by studying generic N-shaped steady-state friction curves coupled to a small number of internal state fields that are characterized by a memory length $D$, which is an underlying assumption of this paper. With these comments, we conclude our discussion of the generic properties of interfacial constitutive laws.

3. Bulk Elasticity in Frictional Systems

In the previous section we focused on the generic properties of the interfacial constitutive law, i.e., on $f[\delta(x,t), \partial_t \delta(x,t), ...]$ in Equation (3). Our goal here is to discuss the properties of the bulk-mediated interaction contributions to this basic equation, i.e., $F_c[\partial_x \delta(x,t), \partial_t \delta(x,t)]$ and $F_c[\partial_x \delta(x,t), \partial_t \delta(x,t)]$. The latter, $F_c[\partial_x \delta(x,t), \partial_t \delta(x,t)]$, represents the coupling between interfacial slip and normal stress variations along the interface. It is now well-established that this coupling exists only when reflection symmetry with respect to the frictional interface is broken [40,72–74]. This symmetry can be broken by material contrast (i.e., with two blocks made of different materials, cf. Figure 1), by geometric asymmetry (i.e., with two blocks of identical materials, but with different shapes, cf. Figure 1) or by the loading configuration. In the first part of the discussion in this section,
we assume perfect material and geometric symmetry, and perfectly anti-symmetric loading, such that \( \mathcal{F}_x[\partial_x \delta(x,t), \partial_t \delta(x,t)] = 0 \). Consequently, we first focus on \( \mathcal{F}_x[\partial_x \delta(x,t), \partial_t \delta(x,t)] \).

The shear stress contribution \( \mathcal{F}_x[\partial_x \delta(x,t), \partial_t \delta(x,t)] \) generally features complex spatiotemporal properties. The spatial properties, in particular the range of bulk-mediated interaction between different parts of the interface, is determined by the system dimensions. We assume that the system length \( L \) is much larger than its height \( H \) (cf. Figure 1 for a visual definition of \( L \) and \( H \)), and focus on the effect of the latter on \( \mathcal{F}_x[\partial_x \delta(x,t), \partial_t \delta(x,t)] \). Consider first two infinite blocks, i.e., \( H \to \infty \), for which we have [75–77]

\[
\mathcal{F}_x[\partial_x \delta(x,t), \partial_t \delta(x,t)] = \frac{H}{2c_s} \partial_t \delta(x,t) - s(x,t),
\]

where \( \mu \) is the shear modulus (already defined above) and \( c_s \) is the shear wave speed (\( s(x,t) \) will be defined and briefly discussed below). The right-hand-side of Equation (13) represents the infinite half-space linear elastodynamic Green’s function [75–77], corresponding to the momentum balance equation \( \rho \ddot{u} = \nabla \cdot \sigma \) (where \( \rho \) is the mass density and the stress tensor \( \sigma \) is related to displacement gradients \( \nabla u \) through Hooke’s law), and contains two contributions. The contribution \( \mu \partial_t \delta(x,t)/2c_s \) is ‘instantaneous’, i.e., it depends on the slip velocity \( \partial_t \delta(x,t) \) alone, and corresponds to wave radiation from the interface into the bulks forming it [45,47,78,79]. It is known as the ‘radiation-damping’ term because mathematically it appears as an effective linear drag/viscous stress that the bulks apply on the interface.

The contribution \( s(x,t) \) in Equation (13) is infinitely long-ranged in space, i.e., it involves integration over the whole interface with power-law kernels, and involves time integration that is limited by causality (formally it is a functional of \( \delta(x,t) \), not just a function of \( x \) and \( t \)). In general, \( s(x,t) \) does not admit an analytic real-space representation and even its spectral (Fourier) representation may be highly complex [75–77]. To nevertheless gain some physical intuition into \( s(x,t) \) and to support the statements just made about it, we consider situations in which the frictional dynamics are significantly slower compared to elastodynamic effects, i.e., we consider the quasi-static limit in which the inertia of the bulks plays no role. In this quasi-static limit, the radiation-damping term is also negligibly small and we have (see, for example, [80])

\[
\mathcal{F}_x[\partial_x \delta(x,t), \partial_t \delta(x,t)] \simeq -\frac{\mu}{2\pi} \int_{-\infty}^{\infty} \frac{\partial_x \delta(x',t)}{x' - x} \, dx'.
\]

The infinitely long-range bulk-mediated interaction between different parts of the interfaces and the associated power-law kernel are evident, as is also schematically illustrated in Figure 1. Mathematically handling the infinitely long-range interaction in the \( H \to \infty \) limit generally raises significant difficulties and in many cases a variety of numerical methods are invoked, e.g., spectral techniques [75–77].

The spatiotemporal range of interaction encapsulated in \( \mathcal{F}_x[\partial_x \delta(x,t), \partial_t \delta(x,t)] \) diminishes as the system height \( H \) is reduced. A particularly useful and analytically tractable expression is obtained in the limit of small \( H \), i.e., if one assumes that \( H \) is sufficiently smaller than any other relevant external/ geometrical lengthscale in the problem, a systematic expansion (reminiscent of the commonly used shallow water approximation in fluid mechanics [81]) yields [60,61,82]

\[
\mathcal{F}_x[\partial_x \delta(x,t), \partial_t \delta(x,t)] \simeq \rho H \partial_t \delta(x,t) - \rho H \partial_x \delta(x,t).
\]

Here \( \rho \) is proportional to \( \mu \), with a prefactor that depends on Poisson’s ratio [61]. In this thin-systems approximation, sometimes termed the quasi-1D limit, the spatiotemporal interaction becomes short-ranged (i.e., captured by second order differential operators, with no integral terms) and the bulk—described by a scalar wave equation with \( H \)-dependent coefficients—is effectively inseparable from the interface. Precisely for this reason, the radiation-damping term of Equation (13) also disappears (i.e., there is no distinct bulk to radiate energy into) and \( \mathcal{F}_x = 0 \). While the short-range spatiotemporal interaction is a limitation, the relative simplicity of the small-\( H \) formulation allows a comprehensive, quantitative and predictive analytical treatment of complex spatiotemporal
frictional dynamics, in addition to extremely efficient numerical calculations [60,61,69], as will be demonstrated below.

When $H$ is neither infinitely large nor small, interesting new effects related to the finite system height emerge, as will be discussed below in relation to various physical phenomena. Moreover, when the perfect symmetry assumption is relaxed, and reflection symmetry with respect to the frictional interface is broken by material contrast, geometric asymmetry or the loading configuration, $\mathcal{F}_c[\partial_x \delta(x,t), \partial_t \delta(x,t)]$ does not vanish. In such situations, which are quite prevalent, one needs to consider both $\mathcal{F}_c[\partial_x \delta(x,t), \partial_t \delta(x,t)]$ and $\mathcal{F}_c[\partial_x \delta(x,t), \partial_t \delta(x,t)]$, and rich dynamical effects emerge from Equation (3). While the analysis of finite systems can be pursued analytically in some cases, numerical methods such as the Finite Element Method (FEM) are required in many other cases. Finally, we note that finite-$H$ systems, i.e., realistic systems, are well described by the infinite-$H$ formulation discussed above on time scales shorter than the typical wave reflection time from outer boundaries (which is determined by $H/c_0$). In the next section we will see how the various forms of bulk elasticity discussed in this section couple to the interfacial constitutive law to give rise to rich spatiotemporal frictional dynamics.

4. Basic Phenomena in Spatially Extended Frictional Systems

Our goal in this section is to demonstrate, through a broad range of examples, how basic phenomena in spatially extended frictional systems emerge from the inherent coupling between the interfacial constitutive law (discussed in Section 2) and bulk elasticity (discussed in Section 3).

4.1. The Stability of Homogeneous Sliding

A central question in relation to spatially extended frictional systems concerns the stability of homogeneous sliding [19,40], i.e., suppose that two bodies in frictional contact slide at a constant homogeneous slip velocity $v_0$. When this sliding motion is unstable, interesting phenomena such as squeaky door hinges [83], squealing car brakes [84] or earthquakes along geological faults [85,86] might emerge. What are the conditions for instability? How are they influenced by the interplay between interfacial friction and bulk elasticity? In this section, we scratch the surface of these important questions through two examples that highlight the interplay between interfacial friction and bulk elasticity in relation to the stability of frictional systems. Mathematically, one considers small spatiotemporal perturbations $\delta v(x,t)$ on top of the homogeneous sliding state $v_0$, i.e., $v(x,t) = v_0 + \delta v(x,t)$, and expand all relevant equations to leading (linear) order in $\delta v(x,t)$. Due to linearization, one can consider each Fourier mode separately, i.e., express the perturbation as $\delta v(x,t) \sim \text{exp}(\Lambda t - ikx)$, where $\Lambda$ is the complex frequency and $k$ is the wavenumber. An instability emerges whenever the real part of $\Lambda$ becomes positive, i.e., $\Re[\Lambda] > 0$, which implies that perturbations exponentially grow in time.

The complex frequency $\Lambda$ and the wavenumber $k$ can be related by perturbing Equation (5) to linear order in $\delta v(x,t)$. The result takes the schematic form

$$-\delta \mathcal{F}_\tau = \sigma_0 \delta f - f \delta \mathcal{F}_\sigma,$$

where $\delta \{ \cdot \}$ (not to be confused with the slip displacement $\delta(x,t)$) represents a variation with respect to $v$.

The first example we consider in this context concerns the stability of homogeneous sliding in the small $H$ limit in the quasi-static regime, i.e., when Equation (15) reads $\mathcal{F}_c \simeq -H \Lambda_{xx} \delta(x,t)$. In the small $H$ limit there exists no coupling between slip and normal stress variations, hence $\mathcal{F}_c = 0$ in Equation (16). Finally, $v_0$ is assumed here to be on the velocity-weakening branch of the steady-state friction curve in Figure 3, i.e., $\frac{df(v)}{dv} < 0$ at $v_0$. In this case, $\delta f$ in Equation (16) is destabilizing, i.e., from the perspective of the friction law alone, perturbations are expected to be amplified. However, as will be shown next, the bulk contribution $\delta \mathcal{F}_\sigma$ is stabilizing and consequently it generates a stability regime that depends on an emerging elasto-frictional lengthscale. Performing the analysis [19,87], one finds
that homogeneous sliding under the stated conditions is unstable for $0 < k < k_c$, where $k_c = 2\pi / L_c$ is determined by a critical length of the form [61]

$$L_c \sim \sqrt{\frac{\mu H D}{-\delta f_{ss}(v) d f_{ss}(v) / d \log v}}.$$

This result shows that indeed stability is enhanced (i.e., the range of instability shrinks) as the bulk combination $\mu H$ increases compared to the interfacial friction combination $\delta f_{ss}(v) / D$, and that the competition between the two (along with the applied normal stress $\sigma_0$) gives rise to a characteristic lengthscale. These properties are demonstrated in Figure 4a. Finally, note that a similar physical picture emerges in the $H \to \infty$ limit, where the critical length of Equation (17) is replaced by an $H$-independent length of the form $L_c \sim \mu D - \sigma_0 / d f_{ss}(v) / d \log v$ [88]. Results for intermediate $H$ values will be presented in Section 4.2.

The second example we consider concerns the stability of homogeneous sliding of two finite height bodies made of the same material, where fully inertial dynamics are taken into account (unlike the quasi-static dynamics discussed in the previous example). This time, the homogeneous slip velocity $v_0$ is assumed to be on the velocity-strengthening branch of the steady-state friction curve in Figure 3, i.e., $\delta f_{ss}(v) > 0$ at $v_0$. In this case, $\delta f$ in Equation (16) is stabilizing, i.e., from the perspective of the friction law alone perturbations are expected to be attenuated, and $\delta F_\tau$ is still stabilizing. Consequently, any possible instability can arise only from a destabilizing normal stress contribution $\delta F_\sigma$. Let us denote the height of the upper body by $H$ and that of the lower one by $\eta H$ [73]. For $\eta = 1$, i.e., for a perfectly symmetric system, we have $\delta F_\sigma = 0$, and homogeneous sliding under the stated conditions is categorically stable [88]. This is demonstrated in Figure 4b. Consider then the case in which $\eta \gg 1$, i.e., there exists significant geometric asymmetry in the system, giving rise to a destabilizing normal stress contribution $\delta F_\sigma$ that competes with the velocity-strengthening stabilizing friction contribution. Depending on the values of the problem’s parameters the former can even overcome the latter, implying an instability. This scenario is demonstrated in Figure 4b, where an instability emerges at a critical geometric asymmetry level $\eta_c > 1$ and a finite range of unstable modes with wavenumbers $k \sim H^{-1}$ exists for larger geometric asymmetry levels, showing that bulk geometry and elasticity can destabilize frictional sliding that is otherwise stable [40,73].

There are other physical ways in which the interplay between interfacial friction and bulk elasticity affects the stability of frictional systems. For example, the combined effect of bimaterial contrast, finite geometry and inertial dynamics, which strongly affect $\delta F_\sigma$, has been shown to play crucial roles in the development of instabilities of velocity-strengthening interfaces (i.e., when the frictional contribution $\delta f$ is stabilizing [89]). In particular, it has been shown that there exists a universal instability with a wavenumber $k \sim H^{-1}$ and a maximal growth rate $\Re[\Lambda] \sim f_{cL} / H$, in the presence of strong bimaterial contrast (e.g., a deformable body sliding on top of a rigid one). This universal instability has been shown to be mediated by waveguide-like modes that are intrinsically coupled to the system height $H$. Furthermore, it has been shown that in the limit of large systems, $H \to \infty$, there exist—in addition to quasi-static instability modes that are relevant at small sliding velocities [40]—also fully dynamic instabilities. These are mediated by propagating plane-wave-like modes, featuring interesting directionality effects with respect to the sliding direction [89]. Overall, these examples show how the interplay between interfacial friction and bulk elasticity determines the stability of homogeneous frictional sliding.
Figure 4. Linear stability spectrum of homogeneous steady sliding. Plotted is the growth rate of perturbations $\Re[\Lambda]$ as a function of wavenumber $k$ (both properly non-dimensionalized). When $\Re[\Lambda] > 0$, sliding is unstable. (a) In the small-$H$ limit, sliding on top of a rigid substrate becomes unstable when $k$ is smaller than a critical value, determined by the critical nucleation length $L_c$ of Equation (17). Curves are shown for two parameter sets, demonstrating that stability is promoted (i.e., the range of unstable modes shrinks) when the body is stiffer (large $\bar{\mu}$) and when friction is less velocity-weakening (smaller $\sigma_0 df_{ss}/d \log(v)$). Details of the derivation of the linear stability spectrum can be found in the Supporting Information of [87], though similar calculations appeared earlier [19,40]. Please note that the kink in the spectrum is related to its complex-plane structure. (b) For a system of two elastic bodies of different height (quantified by the height contrast $\eta$), velocity-strengthening sliding can be destabilized by geometric contrast/asymmetry alone (adapted from Figure 5 of [73]). In the absence of geometric contrast ($\eta = 1$), sliding is categorically stable. At a critical contrast ($\eta_c \approx 2$), an instability emerges (first zero crossing). Finally, for large contrast ($\eta \gg 1$), a range of unstable modes with $k \sim H^{-1}$ exists (see text for additional details). The insets show the geometric configuration in each panel.

4.2. The Onset of Sliding Motion: Creep Patches and Their Stability

The discussion in the previous subsection focused on homogeneous sliding that results from homogeneous loading conditions. However, in many engineering applications and laboratory experiments the loading on a frictional system is inhomogeneous, e.g., it is applied to the trailing edge of the bodies involved. To address this situation, we consider a body of height $H$ in frictional contact with an infinitely rigid substrate that is being pushed sideways, say at $x = 0$, with a velocity $v_d$ [61,87]. The latter is assumed to be on the velocity-weakening branch of the steady-state friction curve in Figure 3, i.e., $\frac{df_{ss}(v)}{d \log(v)} < 0$ at $v_d$. The inhomogeneity of this loading configuration, where at time $t = 0$ the slip velocity is $v = v_d$ at $x = 0$ and $v = 0$ for $x > 0$, gives rise to the propagation of quasi-static creep patches, shown in Figure 5a based on a FEM calculation [87]. Creep patches progressively transfer stress and slip into the interface with increasing time $t$. Their quasi-static nature is explicitly revealed when the time dependence of their length, $l(t)$, is considered; in the small $H$ limit it takes the form $l(t) \sim \sqrt{H v_d t}$ and in the limit of large $H$ it reads $l(t) \sim v_d t$ [61]. In both limits, and for intermediate $H$ values, $l(t)$ depends on $v_d$ and in particular stops growing when the loading is halted, $v_d = 0$. 

Figure 5. Loss of stability in transient, spatially inhomogeneous frictional dynamics under sideways loading (adapted from Figure 3 of [87]). A long elastic body of height $H$, whose trailing edge $(x=0)$ is pushed at a velocity of 10 $\mu$m/s, is considered. (a) A creep patch, manifested through different temporal snapshots of the slip velocity field $v(x,t)$ (see legend), is observed to propagate from the trailing edge into the interface as time progresses. When the sliding region reaches a critical size $L_c(H)$ (see text for details), the system loses stability and $v$ rapidly grows (note the time difference between the two upper curves). The subsequent dynamics feature a rupture front that rapidly propagates further into the interface $(x > L_c(H))$, which is not shown. (b) Dependence of critical length $L_c$ on system height $H$. The lines show analytical predictions for a 2D analysis (solid line) and a small-$H$ limit analysis (dashed line). The two agree for small $H$, while the former saturates at a finite value for large $H$. The circles show numerical results from 2D finite-$H$ FEM simulations [87]. Reasonable and non-trivial agreement between the theory, which is based on homogeneous linear stability of steady sliding, and the numerics, based on transient inhomogeneous simulations, is demonstrated (see [87] for additional discussion).

Do the creep patches propagate indefinitely as long as $v_d$ is maintained finite? The stability analysis of homogeneous sliding presented above suggests that this is not the case. While creep patches are characterized by inhomogeneous slip velocity distributions (cf. Figure 5a), varying from $v(x=0,t)=v_d$ to $v(x>l(t),t)=0$ at any time $t$, the values themselves are predominantly on the velocity-weakening branch of the steady-state friction curve; overall, the physical situation is not expected to dramatically differ from that of a homogeneous patch characterized by an average slip velocity $\bar{v}(t)=\left[l(t)^{-1}\int_0^{l(t)}v(x,t)dx\right]v_d$ that belongs to the velocity-weakening branch. Consequently, the stability analysis discussed above predicts that creep patches undergo a linear instability when their size surpasses the critical (minimal) length for instability, i.e., when $l(t=\tau_c)=L_c(H)$, i.e., when they become large enough to contain the smallest unstable mode. This prediction is tested in Figure 5a and is fully supported by the FEM calculation, i.e., when $l(t)$ surpasses $L_c(H)$ (for the $H$ employed in this calculation), an instability characterized by an exponential growth of slip velocities is observed.

To test this prediction over a broad range of $H$ values, $L_c(H)$ has been theoretically calculated using the homogeneous linear stability analysis discussed in Section 4.1 [87]. Note, though, that the present problem is more complicated than the case discussed in Section 4.1 due to the existence of bimaterial contrast, i.e., the present problem involves a deformable body sliding on top of a rigid one. The scaling properties of the solutions reported on in Section 4.1 are nevertheless preserved, i.e., $L_c(H) \sim \sqrt{H}$ for small $H$ and $L_c(H) \rightarrow \text{const.}$ for large $H$, as shown in Figure 5b. The comparison between the size in which creep paths lose their stability, measured in dynamical simulations of propagating creep patches for various $H$ values, and the theoretical prediction for $L_c(H)$ is also shown in Figure 5b. The dynamics of creep patches and their accompanying instabilities determine the onset of global frictional motion, occurring when slip first reaches the leading edge of the body. These rich and realistic spatiotemporal dynamics crucially depend on the interplay between interfacial friction and
bulk elasticity (as well as other related analyses in the literature, e.g., [82,90–93]), and are qualitatively different from the description of the onset of frictional motion using spring-block models that lack spatially extended degrees of freedom. In fact, the transfer of slip to the leading edge of the body also involves dynamically propagating modes that are excited by the instabilities of creep patches (e.g., the one shown in Figure 5a), not discussed up until now. Such propagating frictional modes are discussed next.

4.3. Propagating Frictional Modes

Propagating frictional modes play crucial roles in the dynamics of frictional systems [3,10,11,27,28,93,94]. For example, following the previous subsection, propagating frictional modes that are generated by instabilities of creep patches play important roles in the onset of global frictional motion, relevant to a wide variety of engineering systems. Our goal in this subsection is to understand how such propagating frictional modes emerge from the interplay between interfacial friction and bulk elasticity and what forms can they take. To this aim, we consider a frictional system subjected to homogeneous shear stress and normal (compressive) stress loading, \( \tau_0 \) and \( \sigma_0 \) respectively (cf. Figure 1).

Suppose then that \( \tau_0 \) is larger than \( \tau_{\text{min}} \equiv \sigma_0 f_{\text{min}} \), where \( f_{\text{min}} \) corresponds to the minimum of the steady-state friction curve in Figure 3. Under these conditions, there exist 3 spatially homogeneous fixed-points (already mentioned above) that correspond to the intersection points of the \( \tau_0/\sigma_0 \) line with \( f_{ss}(v) \), cf. Figure 3. The first and last intersection points, on the velocity-strengthening branches, are stable fixed-points from the perspective of the friction law, while the intermediate one, on the velocity-weakening branch, is an unstable fixed-point (it has been discussed in the context of creep patches in Section 4.2). Generally speaking, frictional modes that propagate at a steady-state speed \( c \) spatially connect the two stable fixed-points. The steady nature of these propagating objects imply that they can be described in a co-moving frame defined through \( \xi \equiv x - ct \), and consequently that \( \partial_t = -c \partial_\xi \) and \( \partial_x = \partial_\xi \) such that all of the relevant equations can be expressed in terms of the co-moving coordinate \( \xi \).

We discuss below 3 types of steady-state propagating frictional modes. First, we consider modes that connect the very low-\( v \) stable fixed-point (\( v \approx 0 \), termed the ‘stick state’) with the high-\( v \) stable fixed-point (\( v > 0 \), the ‘sliding/slipping state’). If the latter is realized at \( \xi \to -\infty \) and the former at \( \xi \to +\infty \), the modes propagate from left to right with \( c > 0 \), and are characterized by a lengthscale \( \ell \) around \( \xi = 0 \) over which the ‘stick state’ changes to the ‘sliding/slipping state’. These propagating frictional modes are sometimes termed frictional rupture fronts. Second, when the stick state invades the sliding state, we obtain healing (‘inverse’) frictional fronts. Finally, when the stick state exists both at \( \xi \to \pm \infty \), with a spatially compact region of \( v > 0 \) around \( \xi = 0 \), we obtain slip pulses. As will be explained below, slip pulses can be constructed from the interaction of rupture and healing frictional fronts (we note in passing that there exists a related frictional mode, composed of a periodic train of pulses [55,74,95], which is not discussed here). In all cases, propagating frictional modes involve the self-selection of both the characteristic transition length \( \ell \) and the propagation speed \( c \). Our goal below is to understand how they are determined by the interplay between interfacial friction and bulk elasticity.

As a first step, we show that \( \ell \) and \( c \) satisfy a generic scaling relation. To see this, note that the time it takes for a propagating frictional mode to travel over a given point on the interface is \( \ell/c \). During this time, the slip accumulated at this point scales as \( v(\ell) \ell/c \sim v_p \ell/c \), where we estimated the slip velocity in the transition region \( v(\ell) \) by the peak (maximal) slip velocity \( v_p \). Finally, since the accumulated slip scales with the typical slip distance \( D \) (while the accumulated slip itself is quite significantly larger than \( D [96] \)), we obtain

\[
\frac{c}{\ell} \sim \frac{v_p}{D}.
\]
This relation can be somewhat more formally rationalized using Equation (8) (the same reasoning applies to Equation (10)). Please note that the relation between \( \ell \) and \( c \) in Equation (18) is independent of the loading conditions and dimensionality as a scaling relation, though an implicit dependence emerges through \( v_p = v_p (0, H) \), as will be discussed below.

To see how these physical quantities emerge in realistic frictional systems, we first discuss steady-state propagating rupture modes. To that aim, we consider a long deformable body of height \( H \) in frictional contact with a rigid substrate, subjected at \( y = H \) to homogeneous shear and normal stress, \( \tau_0 \) and \( \sigma_0 \) respectively (cf. Figure 1). We then obtain steady-state propagating rupture modes using a treadmill FEM routine detailed in Appendix A.3. The results, demonstrating steady-state rupture modes propagating from left to right for a broad range of \( H \) values, are presented in Figure 6. In panel (a) we present \( \tau (\xi) / \sigma_0 \), focusing on the transition region of size \( \ell \) that clearly emerges near \( \xi = 0 \). It is observed that far ahead of the transition region \( \tau (\xi) \) simply equals to the applied shear stress \( \tau_0 \), as expected. Far behind the transition region, \( \tau (\xi) \) approaches a residual stress \( \tau_r \), which identifies here with the loading shear stress \( \tau_r \rightarrow \tau_0 \); this is expected since the two stable fixed-points that the rupture mode is connecting feature the same stress \( \tau_0 \) (cf. the intersection points marked by black squares in Figure 3). The fact that the residual and applied shear stresses are identical for steady-state rupture modes will be further discussed below. Finally, in the transition region, \( \tau (\xi) \) goes through a maximum, which we denote by \( \tau_p \).

The corresponding slip velocity \( v (\xi) / v_{\text{min}} \) is shown in Figure 6b. The residual slip velocity left behind the propagating rupture mode, \( v_r \), corresponds to the high-\( v \) stable fixed-point in Figure 3, \( \tau_0 / \sigma_0 = f_{\text{ss}} (v_r) \), while far ahead of it the slip velocity is vanishingly small, \( v \approx 0 \) (formally also a solution of \( \tau_0 / \sigma_0 = f_{\text{ss}} (v) \)). In between, the two are smoothly connected over a lengthscale \( \ell \), where the connection is monotonic for small \( H \) and non-monotonic for larger \( H \), going through a maximum denoted by \( v_p \). When the transition is monotonic, we simply have for the maximal slip velocity \( v_p = v_r \).

In panel (c) the normalized real contact area \( A(\xi) \) is presented, directly justifying the term ‘rupture front’ as a low \( A \) value state (‘partially ruptured contact’) is observed to invade a high \( A \) value state (reference real contact area). Finally, in panel (d) we present \( \sigma (\xi) / \sigma_0 \), which is simply unity for small \( H \) (since \( \delta F_c = 0 \) in this limit) and deviates from unity as \( H \) increases due to the bimaterial effect manifested in \( \delta F_c \neq 0 \) (recall that we are considering a deformable body sliding on top of a rigid one).

To gain deeper insight into \( \ell \) and \( c \), going beyond the scaling relation in Equation (18), we follow the scaling analysis of [60] in the small-\( H \) limit (\( H \ll \ell \), to obtain

\[
\ell \sim \sqrt{\frac{\mu H D}{\Delta \tau_{p-r}}} \quad \text{and} \quad c \sim v_p \left( \frac{\tau_0}{\sigma_0} H \right) \sqrt{\frac{\mu H}{\Delta \tau_{p-r}} D} \sim v_p \left( \frac{\tau_0}{\sigma_0} H \right) \sqrt{\frac{\mu H D}{\Delta \tau_{p-r}}} \quad \text{for} \quad H \ll \ell , \quad (19)
\]

where the so-called dynamic stress drop is defined as \( \Delta \tau_{p-r} \equiv \tau_p - \tau_r \). Please note that strongly inertial effects, which emerge when \( c \) approaches the elastic wave speeds, have been neglected, that we used the fact that \( v_p \left( \frac{\tau_0}{\sigma_0} H \right) = v_r \left( \frac{\tau_0}{\sigma_0} H \right) \) in the small-\( H \) limit, and that \( \ell \) and \( c \) indeed satisfy the scaling relation in Equation (18). The relations in Equation (19) clearly demonstrate how \( \ell \) and \( c \) emerge from the interplay between interfacial friction and bulk elasticity. The prediction \( \ell \sim \sqrt{H} \) is tested in Figure 6 through rescaling \( \xi \) by \( \sqrt{H} \). It is observed that for small \( H \), but over a large range (here \( H \) varies 16-fold, from 625 \( \mu \) to 1cm), the fields nearly collapse on \( H \)-independent curves, verifying the \( \ell \sim \sqrt{H} \) prediction. For yet larger \( H \) values, some deviations are observed, as will be discussed next. The scaling prediction for \( c \) in Equation (18) will also be tested below.

With increasing \( H \), we expect the small-\( H \) predictions in Equation (19) to break down gradually, eventually being replaced by their large-\( H \) counterparts (\( H \gg \ell \)). The procedure to obtain the latter from Equation (19) is in principle straightforward; one should just replace \( H \) in the first relation in Equation (19) by \( \ell \) and solve for the latter to obtain \( \ell \sim \mu D / \Delta \tau_{p-r} \) (the small-\( H \) shear modulus \( \mu \) is replaced here by \( \mu \)). Then the latter is used to replace \( H \) in the second relation in Equation (19), leading to \( c \sim v_p \mu / \Delta \tau_{p-r} \). The only remaining issue is the \( H \)-dependence of \( v_p \left( \frac{\tau_0}{\sigma_0} H \right) \), which is
clearly observed for the largest $H$ values used in Figure 6b. This pronounced increase of $v_p$ with $H$ is nothing but a signature of the build-up of the famous square root singularity of crack-like objects [97], which is absent in the quasi-one-dimensional limit of small $H$. Consequently, using conventional fracture mechanics in long systems (recall that we assumed $L \to \infty$, i.e., our system features a ‘strip geometry’ $H \ll L$) [97], the residual slip velocity $v_r$ is amplified in the transition region according to $v_p(\tau_0/\sigma_0, H) \sim v_r(\tau_0/\sigma_0)\sqrt{H/\ell}$. Taken together, we obtain for the large-$H$ limit the following predictions

\[ \ell \sim \frac{\mu D}{\Delta \tau_{p,r}} \quad \text{and} \quad c \sim v_r(\tau_0/\sigma_0) \sqrt{\frac{\mu H}{\Delta \tau_{p,r} D}} \quad \text{for} \quad \ell \ll H. \quad (20) \]

![Figure 6](image)

Figure 6. Spatial profiles of steady-state rupture fronts for varying $H$ and fixed $\tau_0/\sigma_0$. The solid lines correspond to 2D treadmill FEM calculations (see Appendix A.3 for details and the legend in panel (b) for the $H$ values used). The dashed line corresponds to a quasi-1D small-$H$ limit calculation (for the smallest $H$). The agreement between the two different calculations in the small-$H$ limit is evident. Shown are the (a) shear stress, (b) slip velocity, (c) real contact area and (d) normal stress (all properly normalized). All fields are shown as a function of the co-moving coordinate $\xi$, rescaled by $\sqrt{H}$ in order to highlight the collapse of all curves for small $H$, as predicted by Equation (19). At larger $H$, fronts develop a peak in the slip velocity, cf. panel (b), and a variation in the normal stress, cf. panel (d). All quantities defined in the text are marked on the figure, note in particular the transition length $\ell$ and the propagation speed $c$ (these, as with all other quantities marked, depend on $H$).

This analysis seems to suggest that the propagation speed $c$ follows the scaling relation $c \sim \sqrt{H}$ in both the small and large $H$ limits, cf. Equations (19) and (20). While at present we cannot comprehensively test this prediction, simply because very large $H$ values are computationally inaccessible, we can consider the available results of Figure 6, where deviations from the small-$H$ limit are observed (e.g., the lack of collapse of the curves shown in panel (a) for the largest $H$ values used and the appearance of a significant overshoot in the slip velocity profiles shown in panel (b) there). Consequently, we plot in Figure 7a $c(H)$ for the results presented in Figure 6. It is observed that the scaling prediction $c \sim \sqrt{H}$ is accurately followed up to the largest $H$ value considered here. Upon further increasing $H$, $c$ cannot increase without bound and we expect inertial effects to intervene, leading to the saturation of $c$ at the elastic wave speed. As just explained, probing this limit using a
treadmill FEM routine is practically impossible due to the computational costs associated with large $H$ values. Yet, we do demonstrate the inertial saturation effect by considering the dependence of $c$ on the applied shear stress $\tau_0$, a dependence that has not been addressed up to now, using numerical calculations in the small-$H$ limit. In fact, note that results of such calculations (using $\mathcal{F}_r$ of Equation (15)) and efficient numerics $[60,61,69]$) have already been included in Figures 6 and 7a, demonstrating perfect agreement with the small-$H$ FEM results. According to Equations (19) and (20), the $\tau_0$ dependence of $c$ is fully contained in $v_s(\tau_0/\sigma_0)$, which is a property of the high-$v$ velocity-strengthening branch of the steady-state friction curve (determined through $\tau_0/\sigma_0 = f_{ss}(v_s)$).

The ‘spectrum’ of propagation speeds $c(\tau_0)$ is shown in Figure 7b. Please note that $c$ is well-defined for $\tau_0 > \tau_{\text{min}}$, as only in this regime steady-state rupture fronts exist $[60]$. Interestingly, as discussed extensively in $[60]$, in the limit $\tau_0 \rightarrow \tau_{\text{min}}$ the propagation speed $c$ attains a finite value that is much smaller than the elastic wave speed $c_s$ (cf. inset), a result that might be related to the long-debated problem of slow earthquakes/slip phenomena $[27,98–100]$. As $\tau_0$ is increased away from $\tau_{\text{min}}$, the propagation speed $c$ increases according to the theoretical prediction of Equation (19), which corresponds to the dashed line. However, inasmuch as $c$ cannot increase indefinitely with $H$, it cannot also increase indefinitely with $\tau_0$. In particular, we expect that once $c$ becomes comparable to $c_s$, inertial effects become dominant and $c \rightarrow c_s$. This expectation is fully supported by the results of Figure 7b. In the inertia-dominated regime other effects that are not discussed here emerge. For example, the characteristic length $\ell$ undergoes relativistic (Lorentz) contraction that follows a $\sqrt{1 - c^2/c_s^2}$ (in the large $H$ limit, $c_s$ is replaced by the Rayleigh wave speed $[94,97]$).

As mentioned above, one can consider in addition to frictional rupture fronts also frictional healing fronts. In principle, the discussion and procedures presented above in relation to rupture fronts can be applied to healing fronts as well, where the stick state invades the sliding state instead of vice versa. Here we focus on just one aspect of the results concerning healing fronts, the one that addresses the spectrum of propagation speeds $c(\tau_0)$. The major difference compared to the results presented in Figure 7b, where $c(\tau_0)$ for rupture fronts is shown to be an increasing function of $\tau_0$, is that for healing fronts $c(\tau_0)$ is a decreasing function of $\tau_0$ $[101,102]$. This implies that the two propagation speed spectra intersect at some loading level, say $\tau^*$. The latter is of physical significance because it is related to the third type of steady-state propagating frictional modes mentioned above, i.e., to slip pulses (sometimes also termed self-healing pulses) $[101,103,104]$. The point is that when the system is driven at $\tau_0 = \tau^*$, rupture and healing fronts propagate at the same speed and hence can be superimposed without interaction to form a steady-state pulse of infinite width. As $\tau_0$ is increased above $\tau^*$, the two fronts interact, leading to pulses of finite width $\mathcal{L}(\tau_0)$, with $\mathcal{L} \rightarrow \infty$ for $\tau_0 \rightarrow \tau^*$. An example of such a finite-width propagating slip pulse in the small-$H$ limit (i.e., using $\mathcal{F}_r$ of Equation (15)) is presented in Figure 8a $[102]$. Up until now, we focused on steady-state propagating frictional modes, i.e., on solutions that depend only on the co-moving coordinate $\xi = x - ct$. These modes can be also studied in fully dynamic numerical simulations in which steady-state conditions are not enforced. A snapshot from such a simulation in the infinite-$H$ limit (i.e., using $\mathcal{F}_r$ of Equation (13)) is shown in Figure 8b $[102]$. The figure presents a slip pulse that appears to propagate over long distances without appreciably changing its shape and propagation speed. Such pulses are sometimes termed sustained pulses $[102]$, though determining whether they are truly steady-state objects or not is difficult based on dynamic simulations. The quite pronounced differences in the shape of the slip pulses presented in Figure 8a,b are related to dimensionality (i.e., the former is obtained in the quasi-1D limit of small $H$ and the latter in the 2D infinite-$H$ limit).
Figure 7. Steady-state rupture front speeds. (a) The propagation speed $c$, corresponding to the 2D steady-state rupture fronts shown in Figure 6, as a function of $H$ (circles). Results using quasi-1D calculations for various $H$ values are added (diamonds). The solid line is the theoretical prediction for $c(H)$ in Equations (19) and (20), with a prefactor of order unity, demonstrating favorable agreement with the numerical results. Please note that both $c$ and $H$ are dimensional, and that the presented results are obtained in the quasi-static regime where Equations (19) and (20) are valid (the wave speed in the 2D calculations has been assigned a large value, see Appendix A.3). (b) Rupture propagation speed as a function of loading stress $\tau_0$ for fixed $H$ in fully inertial quasi-1D calculations (solid line), highlighting inertial effects. The quasi-static approximation (dashed line) agrees with the inertial solutions as long as $c$ is sufficiently smaller than the wave speed $c_s$, while strong deviations are observed when $c$ saturates at $c_s$. Inset: same data in logarithmic vertical axis (axis labels are not shown again), highlighting the existence of a finite minimal propagation speed $c_{\text{min}}$ and the exponential growth of $c$ with $\tau_0$ in the quasi-static regime.

Looking for steady-state solutions by using the co-moving coordinate $\xi = x - ct$ has other implications. Most notably, by so doing one excludes from the discussion the possibility that some of the obtained steady-state solutions are in fact unstable. It turns out that small-$H$ steady-state pulse solutions, an example of which is presented in Figure 8a, are actually unstable [102]. In particular, it has been shown that small-$H$ steady-state pulse solutions can be used to sort out perturbations based on their spatial extent; suppose that a homogeneous state corresponding to the stick fixed-point of the steady-state friction curve at a given $\tau_0 > \tau^*$ is introduced with a slip velocity perturbation. If this perturbation is narrower than the characteristic width of the corresponding steady-state slip pulse (i.e., its spatial extent is smaller than $L(\tau_0)$), then it decays in the subsequent dynamics (where steady state is not enforced). On the other hand, perturbations characterized by a spatial extent larger than $L(\tau_0)$ are amplified and develop into propagating rupture fronts [102].

The unstable nature of small-$H$ steady-state slip pulses has led to the idea that they might in fact serve as critical nuclei for rapid slip propagation mediated by rupture fronts, in analogy to first order phase transitions, i.e., it has been suggested that slip pulses, and in particular their characteristic size $L(\tau_0)$, determine the minimal size of perturbations needed for a homogeneous stick state to develop rupture fronts that bring the system to the high-$v$ velocity-strengthening fixed-point for the same $\tau_0$, cf. Figure 8b. An example of a rapid rupture front emerging from introducing a perturbation larger than $L(\tau_0)$ into a homogeneous stick state, in the infinite-$H$ limit, is shown in Figure 8c. This phase transition-like nucleation process of rupture fronts at locked-in (i.e., stick state) frictional interfaces is qualitatively different from the linear frictional instability nucleation process at velocity-weakening interfaces, discussed in Section 4.2 (also compare to the nucleation scenario proposed in [105]).
Figure 8. (a) A steady-state slip pulse of characteristic size \( L \) obtained in the small-\( H \) limit using \( F_{\tau} \) of Equation (15) (adapted from Figure 2 in [102]). Shown is the slip velocity profile as a function of the co-moving coordinate \( \xi \) used in steady-state calculations (normalized by \( L_c \) of Equation (17)).

(b) A sustained slip pulse obtained through a spectral boundary integral method calculation in the 2D infinite-\( H \) limit using \( F_{\tau} \) of Equation (13) (extracted from supplementary movie S4 in [102]). The curve is truncated in the vertical axis for visual clarify. (c) A non-steady 2D infinite-\( H \) rapid rupture front generated through a spectral boundary integral method calculation (adapted from Figure 3 in [106]). The curve is truncated in the vertical axis for visual clarify. For all panels, see text for additional details.

The 2D infinite-\( H \) rapid rupture front shown in Figure 8c highlights another interesting bulk elasticity effect in frictional systems, which concludes our discussion. Comparing the rupture front in Figure 6a to the one in Figure 8c reveals a qualitative difference; the former leaves behind it a residual stress \( \tau_r \) that equals the applied shear stress \( \tau_0 \) (achieved ahead of the front), while in the latter we observe \( \tau_r < \tau_0 \). Put differently, the former features a vanishing stress drop \( \Delta \tau_{r-s} \equiv \tau_0 - \tau_r \rightarrow 0 \) (\( \Delta \tau_{r-s} \) should be distinguished from the dynamic stress drop \( \Delta \tau_{p-s} \), discussed above), while the latter features a finite stress drop \( \Delta \tau_{r-s} \). The existence of stress drops has important implications for frictional rupture [3,106–108]. The origin of this qualitative difference is that \( F_{\tau} \) of Equation (13), whose value behind the rupture fronts is simply \( \Delta \tau_{r-s} \) (cf. Equation (3)), is finite for transient, out of steady-state rapid rupture (due to both the radiation-damping term and the long-ranged interaction term \( s(x,t) \) in Equation (13)), while it vanishes under steady-state conditions in finite-\( H \) systems [106]. These findings demonstrate that even a quantity such as the residual stress \( \tau_r \), which is traditionally regarded as a property of interfacial friction law [109,110], is in fact strongly affected by bulk elasticity and bulk dynamics.

5. Concluding Remarks

Frictional systems pose great conceptual, physical and mathematical challenges, with far-reaching scientific and technological implications. A frictional system is generically composed of a low-dimensional part—the contact interface—and a higher-dimensional part—the bulks that surround the interface. The major theme of this feature paper is that in order to understand the behavior of frictional systems one has to take into account both the interfacial physics that characterizes the frictional contact and the bulk-mediated interactions between different parts of the interface, and that this inherent interface-bulk coupling is indispensable.

In the context of macroscopic interfacial constitutive laws, we discussed the widely used rate-and-state friction modeling framework. We highlighted the power of this coarse-grained approach and some of its limitations, and then proposed physically motivated modifications to overcome these limitations. This part of the paper emphasized the generic properties of interfacial constitutive relations, where the discussion focused on spatially extended, multi-contact dry frictional interfaces. We did not explicitly take into account third-body effects, wear or various lubrication processes that may give rise to additional important and interesting interfacial physics. It is worthwhile mentioning that the steady-state friction curve of Figure 3 features all the salient properties of the Strubeck curve [111], which generically describes fluid-lubricated contact interfaces. Consequently, we expect some aspects of our results to be relevant also to lubricated interfaces, possibly offering some unifying concepts.
In the context of bulk-mediated interactions, we focused on a linear elastodynamic bulk constitutive relation, highlighting how macroscopic geometry and material contrast affect the range and nature of the emerging spatiotemporal interactions. While we believe that linear elastodynamics is generally relevant and applicable to frictional systems, other bulk physics might be relevant and important in various situations. The latter may include bulk viscoelasticity (e.g., \([112–114]\)), poroelasticity (i.e., the interaction between fluid flow and solid deformation in porous materials, e.g., \([74,115]\)), plasticity and distributed/offset damage (e.g., \([57]\)), which can affect the overall behavior of some frictional systems.

Finally, we have shown how basic phenomena in frictional systems—including instabilities of homogeneous sliding, the onset of sliding motion and a wide variety of propagating frictional modes (e.g., rupture fronts, healing fronts and slip pulses)—emerge from the inherent interplay of interfacial and bulk physics. We believe that these phenomena are generic and essential for understanding the behavior of a broad range of frictional systems. We therefore hope that the approach discussed and advocated in this feature paper will be useful for solving problems in fields ranging from engineering tribology to geophysics.

**Author Contributions:** All authors contributed to the research and to the writing of the manuscript.

**Funding:** This research was partially funded by the Israel Science Foundation grant number 295/16 (E.B. and M.A.), by the William Z. and Eda Bess Novick Young Scientist Fund (E.B.), by the Harold Perlman Family (E.B.) and by a James S. McDonnell post-doctoral fellowship for the study of complex systems (Y.B.-S.).

**Conflicts of Interest:** The authors declare no conflicts of interest.

**Appendix A**

The goal of the appendix is to provide additional technical details about the results presented above, with a special focus on results that have not appeared elsewhere. The latter include the small stress frictional dynamics summarized in Figure 2b, and the steady-state frictional modes results presented in Figures 6 and 7a, involving treadmill FEM calculations.

**Appendix A.1. Summary of the Mathematical Formulation of the Extended Friction Law**

The extended RSF friction law discussed above is formulated in Equations (9), (10) and (12), where the real contact area \(A(\phi)\) of Equation (12) is used in Equations (10) and (6), and the latter is interpreted as \(\tau^{\text{vis}}(\phi, v)\) in Equation (9). Finally, the function \(g(\cdot)\) is introduced in Equation (8) as done in Equation (10). Collecting all these equations, we obtain

\[
f(\phi, v, \tau^\text{el}) = \frac{\tau^\text{el} + \tau^{\text{vis}}(\phi, v)}{\sigma} \left[ \frac{\tau^\text{el}}{\sigma} + k_BT \frac{1}{\Omega v_i} \left[ 1 + b \log \left( 1 + \frac{\phi}{\phi^*} \right) \right] \sinh^{-1}\left( \frac{v}{2v_c} \exp \left[ \frac{E_\text{D}}{k_BT} \right] \right) \right], \quad (A1)
\]

\[
\phi = 1 - \frac{\phi|v|}{D} g(\cdot) \quad \text{and} \quad \tau^\text{el} = \frac{\mu_0 \sigma}{h \sigma_{in}} \left[ 1 + b \log \left( 1 + \frac{\phi}{\phi^*} \right) \right] v - \tau_0^\text{el} |v| D g(\cdot). \quad (A2)
\]

This model contains the following parameters (in the order they appear above): \(\alpha \equiv \frac{k_BT}{h \sigma_{in}}, b, \phi^*, \overline{\phi} \equiv v_c \exp[-\frac{E_\text{D}}{k_BT}], D, \overline{f}_D \equiv \frac{D_{\text{thm}}}{h \sigma_{in}}\) and whatever additional parameters that appear in the threshold function \(g(\cdot)\). Please note that a defined here is the same one used in Equation (7). The values of these parameters are discussed below in Appendices A.2 and A.3.

In most of the studies discussed in this work (exceptions are mentioned in Section 2.2.1 and in Appendix A.2), \(g(v) = \sqrt{1 + (v^*/v)^2}\) is used. This choice, which introduces a single additional parameter \(v^*\), ensures that \(|v|g(v) \to v^*\) in Equation (A2) is extremely small for \(|v| \ll v^*\), once an extremely small \(v^*\) is chosen, and that \(|v|g(v) \to |v|\) for \(v \gg v^*\). For this choice of \(g(\cdot)\), the emerging predictions are insensitive to the precise value of \(v^*\), as long as it is significantly smaller than all other velocity scales in a given problem. In some applications of the model, most notably when large slip
displacements are involved, one assumes that \( \tau^e \) quickly relaxes to its steady state. Under these conditions, we set \( \tau^e = 0 \) inside Equation (A2) and substitute the resulting \( \tau^e \) in Equations (A1) and (A2) to obtain

\[
f(\phi, v) = \left[ 1 + b \log \left( 1 + \frac{\phi}{\phi^*} \right) \right] \left[ \frac{f_0}{\text{sgn}(v)g(\cdot)} + a \sinh^{-1} \left( \frac{v}{2\vartheta} \right) \right]
\quad \text{and} \quad \phi = 1 - \frac{\phi^*}{D} g(\cdot) \quad \text{(A3)}
\]

The steady-state friction curve \( f_{ss}(v) \), which is sketched in Figure 3, is obtained once we set \( \phi = 0 \) and substitute the solution for \( \phi \) in \( f(\phi, v) \) of Equation (A3).

Finally, note that the resulting \( f_{ss}(v) \) reduces to the conventional RSF expression in Equation (7) once the two constitutive extensions introduced in this paper are removed, i.e., \( \log(1 + \phi/\phi^*) \rightarrow \log(\phi/\phi^*) \) and \( \mu_0 \rightarrow 0 \) (the latter implies \( f_0 = 0 \)) are substituted in Equation (A3), and then the steady-state of \( \phi \) is used. To see this, set \( g(\cdot) = 1 \) (\( g(\cdot) \) accounts for the low-\( v \) velocity-strengthening branch in Figure 3, which is not described by Equation (7)), expand \( \sinh^{-1}(x) \approx \log(2x) \) for large argument \( x \) and drop the resulting \( O(\log^2) \) term. The outcome identifies with Equation (7), with the parameters \( f_0 \equiv \frac{E}{\mu_0} \) and \( \beta \equiv b f_0 \), where \( a \) is the same as above (and note that \( \delta \) is not used here, i.e., \( \vartheta \equiv \nu_e \exp[-\frac{f_0}{kg_T}] \) has been substituted in Equation (A3) before \( \sinh^{-1}(\cdot) \) has been expanded).

Appendix A.2. Theoretical Predictions for Frictional Dynamics under Small Stresses

The set of equations used to obtain the results reported on in Figure 2b is listed and discussed in detail in Section 2.2.1. The only additional input needed to reproduce Figure 2b is the function \( g(\cdot) \) and the determination of the material parameters. The experimental data in Figure 2a clearly indicate that the frictional response up to shear forces of \( 5-6N \) is predominantly linear elastic, i.e., that \( \tau^{\text{vis}} \) is negligible in Equation (9) (and consequently also in Equation (A1)) in this range of loading forces. While the onset of irreversibility is not sharp, it is clear that significant deviations from a linear elastic behavior emerge only for shear forces larger than \( 7N \). To capture these observations in a phenomenological manner, we set \( g(\cdot) = \Theta(\tau - A(\phi)\tau_c) \), where \( \Theta \) is the Heaviside step function. Here \( \tau_c \) is a characteristic interfacial yield stress, whose value is chosen such that the onset of irreversibility takes place around \( 7N \) (in particular, given \( A(\phi(t=0)) \) of Equation (12) as detailed below, we set \( \tau_c = 70 \) MPa). With this choice of \( g(\cdot) \), which has also been used in [60], the \( \nu = 0 \) regime in the steady-state friction curve of Figure 3 becomes a strictly vertical (elastic) \( \nu = 0 \) line (as shown in Figure 1a of [60]).

The friction parameters for the PMMA plates used in the experiments of [50], and which enter Equations (A1) and (A2), are the same as those reported on in Table A1 below (note that \( \nu^* \) there is irrelevant for the present analysis as \( g(\cdot) \) here is characterized by \( \tau_c \), as just discussed). The external parameters are as reported explicitly in [50] (specifically we have \( A_0 = 0.01m^2 \), \( F_v = 24N \) and \( |dF_v/dt| = 0.5N/s \) and the initial conditions are \( \phi(t = 0) = 100s \) (which together with \( b \) of Table A1 allows determination of \( A(\phi(t = 0)) \) mentioned above), \( \delta(t = 0) = 0 \) and \( \tau^e(t = 0) = 0 \). The only missing parameter is the interfacial elasticity ratio \( \mu_0/h \), which is directly extracted from the initial linear slope in the \( F_v(\delta) \) experimental data (dashed line in Figure 2a) according to \( dF_v/d\delta = \frac{\mu_0 F}{h\nu} \left[ 1 + b \log(1 + \phi(t = 0)/\phi^*) \right] \), resulting in \( f_0 = \frac{D\mu_0}{h
u} = 0.209 \) (which is not the same as the one listed in Table A1, though the two are quite close to each other). Then Equations (11), (A1) and (A2) are solved for \( \phi(t), \delta(t) \) and \( \tau^e(t) \) (for the two applied loading protocols \( F_v(t) \)). Finally, \( \delta(t) \) of Figure 2b is obtained (for the two protocols) by integrating \( \delta(t) \) with \( \delta(t = 0) = 0 \).

Appendix A.3. A Treadmill FEM Routine for Obtaining Steady-State Frictional Modes

The FEM results reported above have been obtained using the FEM software package FreeFem++ [116]. The partial differential equations are solved in a monolithic scheme on a triangular mesh using a real-space implementation. The fully inertial evolution equations, which include the linear elastodynamic equations (momentum balance with Hooke’s law), are expressed in weak form in the spirit of the FEM. The time derivatives are expressed via finite differences. The frictional boundary
condition is implemented via a semi-implicit integration scheme. Details about the results reported on in Figure 5 can be found in [87], while here we focus on the results reported on in Figures 6 and 7a.

In these calculations, a deformable body of height \( H \) and length \( L \gg H \) sliding on top of a rigid substrate has been considered. A shear stress of magnitude \( \tau_0 \), which is larger than the minimum of the steady-state friction curve in Figure 3, and a normal stress \( \sigma_{yy} = -\sigma_0 \) are applied at \( y = H \) (cf. Figure 1). The lateral boundaries are traction free and the rigid substrate implies that we have \( u_y = 0 \) at the interface (which ensures no interpenetration into and detachment from the substrate). To efficiently reach steady-state conditions, a “treadmill” routine in which the system is shifted such that the propagating mode always remains in the lateral center of the sliding body has been employed. Consequently, we effectively obtain a description in a co-moving reference frame. This requires imposing additional boundary conditions on the fields \( \phi \) and \( \tau_{el} \) (cf. Equation (A2)) at the leading edge, which are chosen according to the stable low-velocity fixed-point corresponding to the intersection of \( \tau_0 \) the steady-state friction curve in Figure 3. The treadmill scheme also allows a logical increase of the spatial resolution near the front position, a main region of interest, without the need for remeshing during propagation, as the front always remains at the same position.

For the largest simulated systems with \( H = 16 \) cm we used a slider of length up to \( L = 9.6 \) m. Altogether, we typically used non-uniform discretizations with up to \( 800 \times 20 \) grid points, and checked that the results are insensitive against further mesh refinement. To avoid numerical instabilities, the dynamics is suppressed near the trailing and leading edge of the slider, and its length \( L \) is chosen to be sufficiently large to avoid their influence on the front dynamics. We used an initially large timestep, \( \Delta t = 10^{-4} \) s, which allows for rapid progress in the initial stage of the simulations. This large initial timestep also stabilizes the dynamics, but leads to reduced accuracy. Consequently, we varied the timestep such that it relaxes to a final value of \( \Delta t = 10^{-7} \) s in the course of the simulations, and we checked that the results are then insensitive against further reduction of the timestep. The simulations typically run until the front has traversed the system length at least once to get rid of sound waves and perturbations which result from the initialization of the system, and which interfere with the front dynamics. After this time, steady-state conditions are reached, and the spatial profiles of all relevant fields are extracted. The results in Figures 6 and 7a involve the friction law of Equations (A1) and (A2) (together with \( g(v) = \sqrt{1 + (v^* / v)^2} \)), with the bulk and interfacial parameters listed in Table A1. The same interfacial parameters are relevant also for Figure 2b, see the caption of Table A1 and Appendix A.2 for details.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu )</td>
<td>3.1 GPa</td>
</tr>
<tr>
<td>( \nu )</td>
<td>1/3</td>
</tr>
<tr>
<td>( \rho )</td>
<td>60 kg/m³</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.005</td>
</tr>
<tr>
<td>( b )</td>
<td>0.075</td>
</tr>
<tr>
<td>( \phi^* )</td>
<td>( 3.3 \times 10^{-4} ) s</td>
</tr>
<tr>
<td>( \dot{\vartheta} = v^* )</td>
<td>( 10^{-7} ) m/s</td>
</tr>
<tr>
<td>( \dot{D} )</td>
<td>( 0.5 \times 10^{-6} ) m</td>
</tr>
<tr>
<td>( \tilde{f}_0 )</td>
<td>5/18</td>
</tr>
<tr>
<td>( \sigma_{nt} )</td>
<td>( 5.4 \times 10^8 ) Pa</td>
</tr>
</tbody>
</table>
References


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).