Journal of the Mechanics and Physics of Solids 59 (2011) 2442–2454

Contents lists available at SciVerse ScienceDirect

Journal of the Mechanics and Physics of Solids

journal homepage: www.elsevier.com/locate/jmps

Frictional energy dissipation in contact of nominally flat rough surfaces under harmonically varying loads

C. Putignano [a], M. Ciavarella [a,c,*], J.R. Barber [b]

[a] CEMEC-Politecnico di Bari, Bari, Italy
[b] Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109-2125, USA
[c] TUHH Univ., Eissendorfer Str.42, DE-20148 Hamburg, Germany

ARTICLE INFO

Article history:
Received 16 June 2011
Received in revised form
7 September 2011
Accepted 16 September 2011
Available online 1 October 2011

Keywords:
Contact problems
Fretting
Coulomb friction
Hysteretic damping
Rough surfaces

ABSTRACT

If the nominal contact tractions at an interface are everywhere below the Coulomb friction limit throughout a cycle of oscillatory loading, the introduction of surface roughness will generally cause local microslip between the contacting asperities and hence some frictional dissipation. This dissipation is important both as a source of structural damping and as an indicator of potential fretting damage. Here we use a strategy based on the Ciavarella-Jäger superposition and a recent solution of the general problem of the contact of two half spaces under oscillatory loading to derive expressions for the dissipation per cycle which depend only on the normal incremental stiffness of the contact, the external forces and the local coefficient of friction. The results show that the dissipation depends significantly on the relative phase between the oscillations in normal and tangential load—a factor which has been largely ignored in previous investigations. In particular, for given load amplitudes, the dissipation is significantly larger when the loads are out of phase. We also establish that for small amplitudes the dissipation varies with the cube of the load amplitude and is linearly proportional to the second derivative of the elastic compliance function for all contact geometries, including those involving surface roughness. It follows that experimental observations of less than cubic dependence on load amplitude cannot be explained by reference to roughness alone, or by any other geometric effect in the contact of half spaces.

1. Introduction

If a contact between two elastic components transmits a combination of normal and tangential loads, it is well known that *microslip* can occur in localized regions at load levels well below those required for complete slip or sliding to occur. If the loading is periodic, as in the case of a machine subject to vibratory loading, the resulting oscillatory microslip is a source of frictional energy dissipation (hysteretic damping) (Wentzel, 2006) and can lead to the initiation of local fretting fatigue cracks and eventual failure (Nowell et al., 2006).

The first solution to a problem of this kind is that due to Cattaneo (1938) (see also Mindlin, 1949), who considered the case where a Hertzian contact is loaded first by a normal load and then by a tangential load. If the tangential load is oscillatory, cyclic microslip is predicted in an annulus surrounding a central stick zone (Mindlin et al., 1952) and classical experimental evidence of the resulting fretting damage was presented by Johnson (1955).

* Corresponding author. Tel.: +39 080 5962811; fax: +39 080 5962777.
E-mail addresses: engineeringchallenges@gmail.com, mciava@poliba.it (M. Ciavarella).

0022-5096/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jmps.2011.09.005

If the contact problem is 'conforming', meaning that the bodies make contact over an extended area even before the loads are applied, then a sufficiently large normal preload can be defined such that no microslip is predicted at any point in the contact area during the oscillatory loading cycle. By contrast, in 'non-conforming' contact, the normal contact tractions will tend to zero at the edge of the contact region and this almost always implies the occurrence of a local region of cyclic microslip. However, all practical surfaces are rough on the microscale, implying that there are no truly conforming problems when the rough surface profile is included in the problem description. Thus, even in nominally conforming problems, we must anticipate some microslip on the scale of the surface roughness, including the attendant frictional dissipation.

Previous studies of the effect of roughness on hysteretic damping have generally depended on the use of a specific rough surface contact model. For example, Björkland (1997) and Eriten et al. (2011) combined the asperity model of Greenwood and Williamson (1966) with Mindlin's treatment of the tangential contact problem to define a statistical estimate of the resulting dissipation. A similar approach was used by Bureau et al. (2003) to interpret the results of experimental dissipation measurements under microslip conditions.

In this paper, we shall use some recent results (Barber et al., 2011) to show that the dissipation per unit area in a fairly general contact can be determined without reference to such a contact model, provided only that we have access to reliable data for the incremental *normal* contact problem and experimental measurements of the friction coefficient. The former might for example be obtained from a multiscale numerical model of the contact of rough surfaces (Hyun et al., 2004; Campañá et al., 2011; Putignano et al., under review). We also extend previous investigations to consider the effect of oscillations in both normal and tangential tractions, with particular reference to the significant effect of the relative phase of these components.

2. The tangential model

In Barber et al. (2011) a rationale was developed for determining the sequence of traction distributions experienced by a two-dimensional elastic contact between two half planes of arbitrary profiles subjected to fairly general far-field oscillatory loading, under the assumption of Coulomb's friction law. The method draws on Ciavarella and Jäger's extension of the Cattaneo–Mindlin superposition (Jäger, 1998; Ciavarella, 1998a). In particular, it was established that the initial conditions and the time-independent term in the tangential load affect the tractions locked in the 'permanent stick zone', but the extent of this zone and the time-varying terms in the tractions are unaffected by these parameters. As a consequence, the dissipation must also be so independent. The argument depends only on the assumptions that (i) the contacting bodies can be realistically represented as elastic half planes of similar materials, (ii) that the contact obeys Coulomb's law of friction in a pointwise sense with a constant coefficient f.

Here, the underlying contact problem is one between two nominally flat, but rough half spaces and this is inherently a three-dimensional problem. Cattaneo and Mindlin's original papers considered the three-dimensional Hertzian geometry and, relative to the two-dimensional case in Barber et al. (2011), this involves an approximation. The Cattaneo–Mindlin superposition defines a traction distribution that satisfies the condition for 'stick' (no relative tangential motion) in part of the contact area and the slip condition $q(x,y)=fp(x,y)$ outside the stuck region, but the three-dimensional Coulomb law that the direction of the relative microslip at each point must oppose the tractions is satisfied only approximately, except in the special case $\nu=0$ (Ciavarella, 1998b). However, the magnitude of the error involved in this approximation was assessed by Munisamy et al. (1994) and found to be extremely small, so we make no apology for using the same approximation here.

3. Formulation

We suppose that the contacting bodies are subjected to a normal compressive force P and a tangential force Q that vary cyclically in the form

$$P(t)=P_0+P_1\cos(\omega t),\quad Q(t)=Q_0+Q_1\cos(\omega t-\phi), \tag{1}$$

where t is the time and ϕ is a relative phase angle. This defines an ellipse in PQ-space, as shown in Fig. 1.

Notice that if we start from the unloaded condition, this implies the existence of some initial loading path such as OA in Fig. 1, but it was shown in Barber et al. (2011) that the time-varying terms in the steady state are unaffected by the exact nature of this path.

3.1. The Ciavarella–Jäger superposition

Suppose that the normal contact problem has been solved in some sense, so that we know the extent of the contact area $\mathcal{A}(P)$ and the contact pressure distribution $p(x,y,P)$ as functions of the normal load P. Notice that since the surfaces are rough, we anticipate that $\mathcal{A}(P)$ will comprise the collection of 'actual contact areas' at the points of contact between asperities of the two surfaces. The contact pressure is zero outside the contact area and hence

$$p(x,y,P)=0\quad (x,y)\notin\mathcal{A}(P). \tag{2}$$

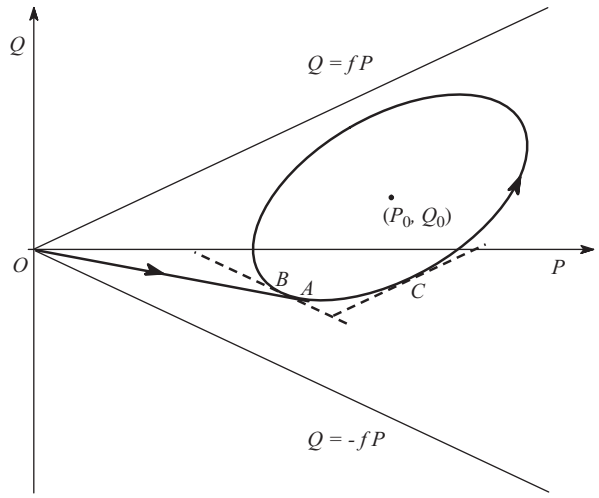

Fig. 1. Periodic loading cycle in PQ-space.

As part of the solution for the normal contact problem, we assume that we also know the rigid body approach of the bodies, which we denote by $v(P)$. This function will also be influenced by the surface roughness and will therefore generally be a strongly non-linear (and stiffening) function of P.

The contact area, contact pressure and normal approach are all non-decreasing functions of P in the sense that $v'(P) \geq 0$, $v''(P) \leq 0$, $\mathcal{A}(P_1) \in \mathcal{A}(P_2)$ if $P_2 > P_1$ and $p'(x,y,P) \geq 0$ for all x,y (Barber, 1974), where we introduce the notation

$$p'(x,y,P) \equiv \frac{\partial p(x,y,P)}{\partial P}, \quad v'(P) \equiv \frac{\partial v(P)}{\partial P}. \tag{3}$$

If the normal force is instantaneously equal to P and $v = 0$, Ciavarella (1998b) showed that the incremental distribution of shear tractions in the x-direction

$$q_x(x,y) = fp(x,y,P) - fp(x,y,P^*), \quad P^* < P \tag{4}$$

satisfies the condition for stick in the region defined by $\mathcal{A}(P^*)$ and for slip in the x-direction in the region $\mathcal{A}(P) \backslash \mathcal{A}(P^*)$. In Barber et al. (2011), it was shown how this superposition could be used in two-dimensional problems to define the traction distribution throughout the cycle defined by Eqs. (1) and Fig. 1.

3.2. Frictional dissipation

Two methods are available for the determination of the frictional dissipation. One is to determine the local slip displacement and integrate it over the slip area, and the other is to evaluate the work done by the external tangential force in moving through the relative rigid-body displacement of the contacting bodies during a complete cycle. In the two-dimensional problem (Barber et al., 2011), only the first of these approaches was practical, since in plane problems the rigid-body displacement due to contact loads is logarithmically infinite. In other words, we can only give a meaning to the relative motion of the distant regions of the bodies if these are finite in extent.

By contrast, the three-dimensional problem gives finite rigid-body displacements even when the contacting bodies are half spaces and this greatly facilitates the energy calculation. In particular, the input power can be written as

$$\dot{W} = Q\dot{u}, \tag{5}$$

where u is the tangential relative rigid-body displacement, which is also equal to the sum of the (uniform) elastic tangential displacements of the contacting bodies in the stick area. Notice that the external work goes to a combination of dissipation and strain energy of the contacting bodies, but if it is integrated over a complete cycle, the initial and final values of strain energy are the same, so the dissipation per cycle is given by

$$W = \oint Q\, du = \oint Q \frac{du}{dP}\, dP. \tag{6}$$

3.3. Tangential compliance

To determine the tangential displacement u in Eq. (6), we need to evaluate the rigid-body displacement in the stick region $\mathcal{A}(P^*)$ associated with the Ciavarella–Jäger traction distribution (4). In the special case where $v = 0$, this is equal to

the corresponding normal displacement $fv(P)-fv(P^*)$, but for other values of v we can obtain a better approximation by considering the relation between the normal and tangential compliance for surface loading of the half space.

The ratio between the tangential and normal compliance depends on the shape of the contact area, but this dependence is relatively weak. For example, Mindlin (1949) showed that the tangential elastic compliance C_T of all 'stuck' elliptical contacts lie in the range

$$C_N < C_T < \frac{C_N}{(1-v)}, \tag{7}$$

where C_N is the corresponding normal compliance and v is Poisson's ratio. The extreme values are only approached when the ellipticity is close to unity, the higher compliance corresponding to loading along the major axis and the lower to loading along the minor axis. For a circular contact area, the two values are of course equal by symmetry and satisfy the relation

$$C_T = \lambda C_N \quad \text{where } \lambda = \frac{(2-v)}{2(1-v)}. \tag{8}$$

It is clear that the difference between the maximum and minimum tangential compliance must correlate with the degree of anisotropy of the contact geometry and it seems probable that (i) all such geometries are likely to be bounded by the limiting expressions (7) and (ii) that all statistically isotropic geometries would exhibit a compliance at or close to the value (8). To the best of the authors' knowledge neither of these results have been rigorously proved, though Campañá et al. (2011) present a persuasive argument based on statistical considerations of the strain energy associated with the traction distributions. Also, Sevostianov and Kachanov (2008) established the relation (8) for a random distribution of elliptical contact areas. Based on these arguments, we therefore assume that the traction distribution (4) will produce a uniform relative tangential displacement in the stick area $\mathcal{A}(P^*)$ of

$$u(x,y) = \lambda f[v(P)-v(P^*)], \tag{9}$$

where λ is defined in (8).

3.4. Evolution of the traction distribution

The determination of the tangential traction distribution during the steady-state loading cycle of Fig. 1 is fully described in Barber et al. (2011) and will only be summarized here. The cycle falls into four segments *BC*, *CH*, *HE*, *EB*, where the points *B*,*C*,*H*,*E* are defined in Fig. 2. The dashed lines defining these points all have slope of $\pm f$ and the ellipse is traversed in the anticlockwise direction, as shown.

In *BC*, the contact area is growing ($dP/dt > 0$) and $d|Q|/dP < f$, thus satisfying the condition for the newly established contact to remain stuck. In *CE* there is a growing forward slip zone, but the equations governing this process change at the point *H*, where a line of slope *f* drawn through *B* cuts the other side of the ellipse. The system then sticks instantaneously at *E*, after which a reverse slip zone grows from the outside of the contact until once again the system sticks instantaneously at *B*.

The tangential traction distribution in these segments can be written down as the superposition of Ciavarella-Jäger distributions, except in the stick phase *BC* where we need to use an incremental (integral) representation. We obtain:

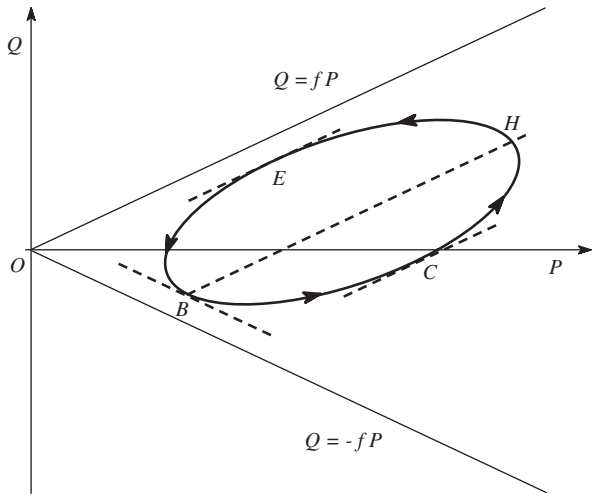

Fig. 2. Definition of the four segments of the ellipse.

In EB:

$$q(x,y) = -fp(x,y,P) + 2fp(x,y,P_S) + q_0(x,y),$$

$$u = \lambda f[-v(P) + 2v(P_S)], \tag{10}$$

where

$$P_S = \frac{1}{2}\left(\frac{Q}{f} + P - \frac{Q_E}{f} + P_E\right) + u_0 \tag{11}$$

and $q_0(x,y)$, u_0 represent time-invariant terms that persist throughout the cycle. These terms depend upon the details of the initial loading path *OA* in Fig. 1 and on the value of the mean tangential load Q_0, but they have no influence on the steady-state dissipation. In particular, the tractions $q_0(x,y)$ are confined to the permanent stick zone. Eq. (11) and similar relations for the remaining segments come from the requirement that the total tangential force be equal to the instantaneous value of Q.

In BC:

$$q(x,y) = -fp(x,y,P_B) + 2fp(x,y,P_T) + \int_{P_B}^{P} p'(x,y,P)\frac{dQ}{dP}\,dP + q_0(x,y),$$

$$u = \lambda f[-v(P_B) + 2v(P_T)] + \lambda \int_{P_B}^{P} v'(P)\frac{dQ}{dP}\,dP + u_0, \tag{12}$$

where

$$P_T = \frac{1}{2}\left(\frac{Q_B}{f} + P_B - \frac{Q_E}{f} + P_E\right) \tag{13}$$

and $\mathcal{A}(P_T)$ defines the permanent stick zone—i.e. the set of points that are in a state of stick throughout the steady-state cycle (Barber et al., 2011).

In CH:

$$q(x,y) = fp(x,y,P) - fp(x,y,P_B) + 2fp(x,y,P_T) - fp(x,y,P_Y) + \int_{P_B}^{P_Y} p'(x,y,P)\frac{dQ}{dP}\,dP + q_0(x,y),$$

$$u = \lambda f[v(P) - v(P_B) + 2v(P_T) - v(P_Y)] + \lambda \int_{P_B}^{P_Y} v'(P)\frac{dQ}{dP}\,dP + u_0, \tag{14}$$

where Y is a point on the ellipse where

$$Q_Y - fP_Y = Q - fP \tag{15}$$

and can be found from the construction in Fig. 3.

In HE:

$$q(x,y) = fp(x,y,P) - 2fp(x,y,P_K) + 2fp(x,y,P_T) + q_0(x,y), \tag{16}$$

$$u = \lambda f[v(P) - 2v(P_K) + 2v(P_T)] + u_0, \tag{17}$$

where

$$P_K = P_T + \frac{1}{2}\left(P - \frac{Q}{f} - P_E + \frac{Q_E}{f}\right). \tag{18}$$

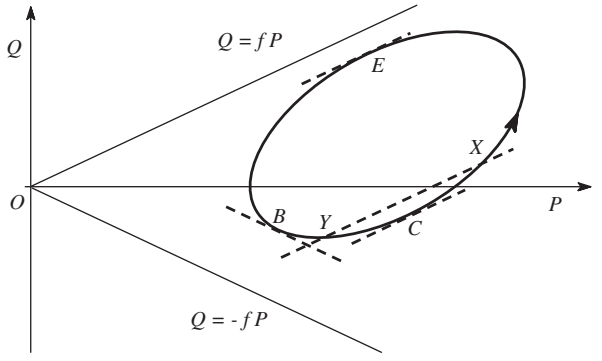

Fig. 3. Geometric construction for determining the point Y and hence P_Y. The instantaneous values of P and Q define the point X.

If these expressions are substituted into (6), the resulting integral, partitioned into the four segments EB,BC,CH,HE can be evaluated numerically for any given normal compliance function $v(P)$. The integrals can be performed parametrically, using t as a parameter, or alternatively, t can be eliminated between Eqs. (1) to yield Q as a function of P, after which (6) can be evaluated directly as an integral with respect to P.

3.5. Power-law compliance relation

Following Jang and Barber (2011), we can develop a more compact representation by defining the dimensionless loading parameters

$$\hat{P}=\frac{P}{P_0},\quad \hat{Q}=\frac{Q}{fP_0}. \tag{19}$$

A further simplification arises if the normal compliance relation has the power-law form

$$v'(P)=CP^{\beta}, \tag{20}$$

where C,β are constants. In this case, we write

$$\hat{u}=\frac{u}{C\lambda f P_0^{\beta+1}} \tag{21}$$

and the dissipation integral (6) then takes the form

$$\hat{W}\equiv\frac{W}{C\lambda f^2 P_0^{\beta+2}}=\oint \hat{Q}\,d\hat{u}=\oint \hat{Q}\,\frac{d\hat{u}}{d\hat{P}}\,d\hat{P}. \tag{22}$$

The dimensionless tangential compliance function $d\hat{u}/d\hat{P}$ can then be summarized in the equations

$$\begin{aligned}\frac{d\hat{u}}{d\hat{P}}&=-\hat{P}^{\beta}+2\hat{P}_S^{\beta}\frac{d\hat{P}_S}{d\hat{P}} \quad \text{in } EB\\ &=\hat{P}^{\beta}\frac{d\hat{Q}}{d\hat{P}} \quad \text{in } BC\\ &=\hat{P}^{\beta}-\left[1-\left(\frac{d\hat{Q}}{d\hat{P}}\right)_{\hat{P}_Y}\right]\hat{P}_Y^{\beta}\frac{d\hat{P}_Y}{d\hat{P}} \quad \text{in } CH\\ &=\hat{P}^{\beta}-2\hat{P}_K^{\beta}\frac{d\hat{P}_K}{d\hat{P}} \quad \text{in } HE.\end{aligned}$$

Of course, the power-law compliance relation (20) reduces the generality of the solution, but the fractal nature of typical rough surfaces suggests that a power law should represent a good approximation, based on Barenblatt's (1996) concept of incomplete similarity. Also, numerous experimental and theoretical treatments of the mathematically analogous problem of thermal or electrical contact resistance exhibit behaviour close to power-law form (Paggi and Barber, 2011).

More specifically, the widely used theories of Persson (e.g. Persson et al., 2002; Yang and Persson, 2008) are based on the concept that the statistical population of asperity contacts remains unchanged as the normal force is increased, leading to Eq. (20) with $\beta=-1$. In other words, the normal stiffness (reciprocal of compliance) is assumed to be proportional to the normal load. We shall use this function in the numerical results that follow, but we emphasize that there is no particular difficulty in using a more general function $v'(P)$, such as one obtained from a numerical study of the normal contact of two prescribed rough surfaces.

4. Results

In this section, we shall explore the effect of the three independent dimensionless parameters, $\hat{P}_1$,$\hat{Q}_1$,ϕ on the dissipation $\hat{W}$ for the special case $\beta=-1$. Notice that this paper is restricted to cases where the entire load loop is contained between the lines $Q=\pm fP$ in Fig. 1, so that the permanent stick zone $\mathcal{A}(P_T)$ is non-null. Thus, we do not consider cases where the contacting bodies experience periods of gross slip (sliding) or separation. Most of the following results are obtained by numerical integration of the preceding equations, but analytical expressions can be obtained for the case $\phi=0$, and the corresponding analysis is summarized in the Appendix. The numerical procedure tends to become ill-conditioned near $\phi=0,\pi$, but the results obtained agree well with the analytical limits, which lends some confidence to the accuracy and numerical convergence of the results. For example, Fig. 4 compares numerical and analytical expressions for the 'Ciavarella–Jäger' case $\hat{P}_1=0$, for various values of $\hat{Q}_1$ below the gross slip limit.

Since all previous studies of frictional dissipation due to microslip have considered either pure tangential oscillations or in-phase oblique loading (e.g. Johnson, 1961), we first examine the effect of phase on the dissipation at fixed values of $\hat{P}_1$,$\hat{Q}_1$. Fig. 5 shows $\hat{W}$ as a function of ϕ for some representative cases. In all cases, the dissipation increases with ϕ to a

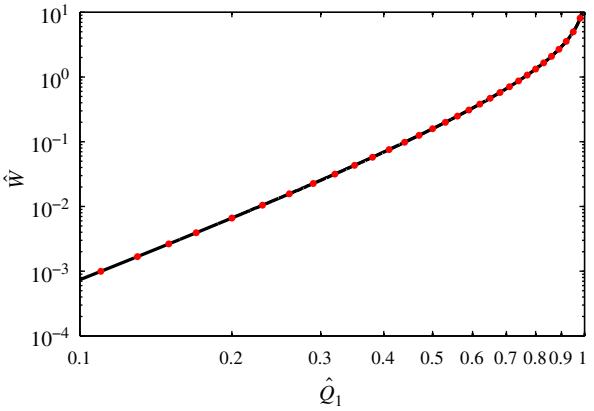

Fig. 4. Dissipation as a function of $\hat{Q}_1$ for the case $\hat{P}_1 = 0$. The points were obtained by numerical integration, using the results of Section 3 and the line was obtained using the analytical expressions from the Appendix.

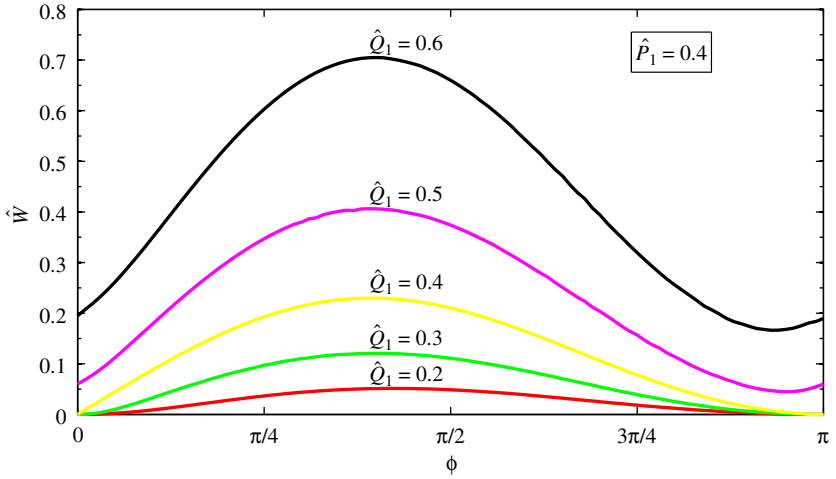

Fig. 5. Effect of relative phase for $\hat{P}_1 = 0.4$ and various values of $\hat{Q}_1$.

maximum near $\phi = \pi/2$. This result agrees with the conclusions of Jang and Barber (2011), but that paper considered mostly cases involving gross slip and/or separation so comparisons are limited. The dependence on ϕ is substantial, implying that it is critical to estimate the phase relation of the loading in practical applications and take it into account in any estimate of effective damping.

Fig. 6 shows the effect of obliquity of loading for the in-phase case $\phi = 0$ and is plotted using the analytical expressions (32) from the Appendix. This is the loading treated experimentally by Johnson (1961), but recall that his system comprises a smooth axisymmetric Hertzian contact, whereas the present results are for a nominally flat rough surface. The plot is presented on a logarithmic scale, so the approximately linear behaviour indicates a power-law dependence on the amplitude $\hat{Q}_1$ and the slope at small $\hat{Q}_1$ is asymptotically cubic. The normal compliance relation for Johnson's problem can also be expressed in the power-law form (20) with $\beta = -1/3$ and he also predicted cubic dependence of dissipation on shear load amplitude, suggesting that this behaviour is independent of β. In fact, we shall demonstrate in Section 5.1 that any problem involving the contact of elastic half spaces must exhibit dissipation with cubic dependence on the load amplitude at sufficiently small oscillatory loads for arbitrary but specified values of ϕ and the ratio $\hat{P}_1/\hat{Q}_1$, subject only to the condition that $v'(P)$ be differentiable in the vicinity of the mean normal load P_0. Notwithstanding this general result, we should remark that many authors (including Johnson, 1961) have reported experimental results in which the dependence of dissipation on load amplitude is less than cubic and may even approach quadratic. The present results rule out the possibility of explaining these results in terms of surface roughness or indeed any geometrical effects.

In interpreting Fig. 6, the reader should recall that $\hat{Q}_1$ contains the coefficient of friction f and hence dissipation is zero for $\hat{P}_1/\hat{Q}_1 \geq 1$. Also, the curves terminate at the point where a further increase in $\hat{Q}_1$ would involve a period of gross slip.

Fig. 7 shows a similar plot of dissipation against $\hat{Q}_1$ for $\phi = \pi/2$. The curves again exhibit a substantial linear region with a slope again indicating a cubic dependence on load amplitude. However, in contrast to Fig. 6, we observe (i) that the

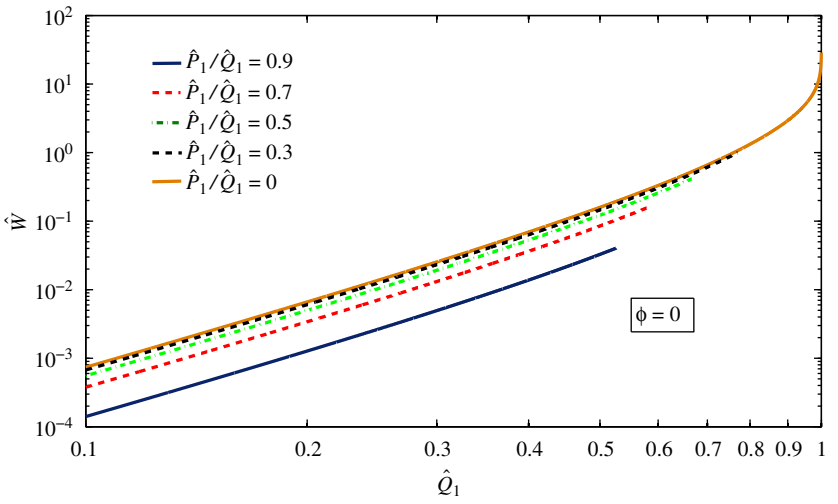

Fig. 6. Effect of obliquity of loading on dissipation for in-phase loading ($\phi = 0$).

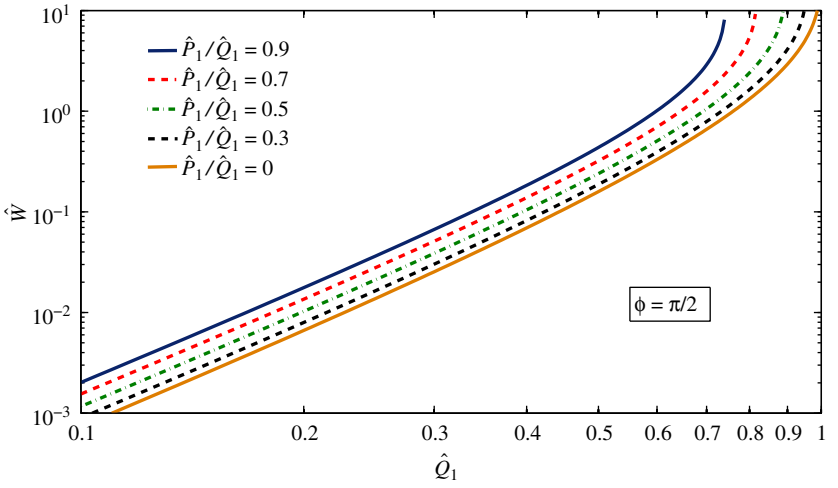

Fig. 7. Effect of load amplitude on dissipation for out-of-phase loading $\phi = \pi/2$.

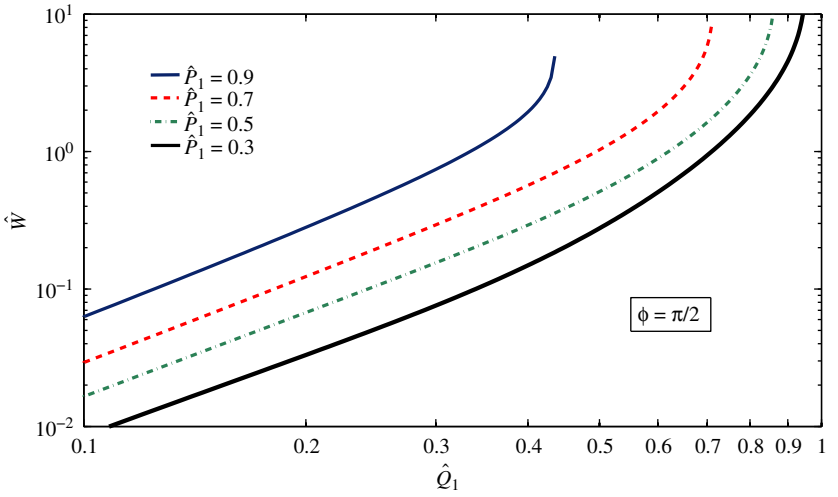

Fig. 8. Effect of $\hat{Q}_1$ on dissipation for out-of-phase loading ($\phi = \pi/2$) and various constant values of $\hat{P}_1$.

dissipation is consistently significantly larger than the corresponding values for in-phase loading, confirming the conclusions of Fig. 5, and (ii) they bend upwards as we approach the gross slip limit.

Finally, in Fig. 8, we show the dissipation for out-of-phase loading as a function of $\hat{Q}_1$ for several fixed values of $\hat{P}_1$ (in the previous figures the ratio $\hat{P}_1/\hat{Q}_1$ was held constant). Notice that in this figure, the slope of the initial linear portion is lower, indicating a power of about 2.2. It is tempting to identify this result with published experimental observations of lower than cubic dependence of dissipation on load amplitude, but there is no indication in these experiments that a constant out-of-phase normal load oscillation was involved. Plots similar to Fig. 8 were also developed for other values of the phase lag ϕ, but no slopes lower than 2.2 were obtained. In particular, the slope is also about 2.2 for $\phi = 0.4\pi$, which is close to the maximum energy dissipation point in Fig. 5.

5. Discussion

5.1. Small amplitude oscillations

If the $\hat{P}_1, \hat{Q}_1 \ll 1$, the oscillating loads will be small compared with the mean normal load P_0 and the ellipse in Fig. 1 will be small and relatively distant from the slip lines $Q = \pm P$. To explore the behaviour in this range, it is convenient to define a new variable

$$\tilde{P} = \hat{P} - 1 = \frac{P - P_0}{P_0} \ll 1,$$

and it is easily verified from Eqs. (11) and (18) and the construction in Fig. 3 that the corresponding values of $\tilde{P}_S, \tilde{P}_Y, \tilde{P}_K \ll 1$ at all points in the cycle. If $v'(P)$ is a continuous function, we can therefore expand it as a Taylor series

$$v'(P) = v'(P_0) + v''(P_0)(P - P_0) + \frac{v'''(P_0)(P - P_0)^2}{2!} + \cdots \tag{23}$$

and for sufficiently small load amplitudes, only contributions from the first two terms will be significant. Furthermore, the first term makes no contribution to the integral (6). This can be verified by a rather tedious evaluation of the integrals in the four regions *EB*,*BC*,*CH*,*HE*, but a more direct proof comes from the observation that a constant value of $v'(P)$ can arise only if $\mathcal{A}(P)$ is independent of P, in which case we know from the Ciavarella–Jäger superposition that the entire contact area must be in a state of stick as long as we do not reach the gross slip limit. Thus, at small amplitudes, we need only consider the second (linear) term in (23).

Suppose we now perform a second normalization on the integrals in Section 3.4 by defining

$$P^* = \frac{P - P_0}{P_1}, \quad Q^* = \frac{Q}{fP_1}, \tag{24}$$

so that (for example) in *EB*,

$$\frac{du}{dP} = \lambda f P_1 v''(P_0)\left[-P^* + 2P_S^* \frac{dP_S^*}{dP^*}\right].$$

Using this result and (24) in the integral (6) over the segment *EB*,we then have

$$W_{EB} = f^2 P_1^3 v''(P_0) \int_{P_E^*}^{P_B^*} Q^* \left[-P^* + 2P_S^* \frac{dP_S^*}{dP^*}\right] dP^* \tag{25}$$

and the load amplitude enters the expression only through the explicit multiplier P_1^3 (or equivalently Q_1^3), since the integral, including the limits, is the same for all load paths of a given shape (prescribed Q_1/fP_1 and ϕ). Similar results are obtained for each of the four segments, showing that dissipation is proportional to the cube of the load amplitude.

We also note that for small load amplitudes, the dissipation due to surface roughness is linearly proportional to the second derivative of the compliance function at the mean load, $v''(P_0)$. This suggests that dissipation may be sensitive to microscopic changes in surface topography due to wear, which may be a contributory factor in regard to Johnson's observation that experimental measurements of dissipation show significant history dependence (Johnson, 1961).

Mathematically, the lack of dependence on the constant term $v'(P_0)$ has the incidental advantage of eliminating complications associated with the infinite displacements in two-dimensional, half-plane solutions of the contact problem and also permits one to use numerical models of sufficiently large finite bodies without regard to the conditions at the distant boundaries.

5.2. Bounds on the dissipation

For the present theory to yield quantitative results, we need to measure or estimate the incremental normal compliance function $v(P)$. Barber (2003) showed that the inverse 'load-compliance' function $P = P(v)$ can be rigorously bounded between the curves

$$F(v) > P(v) > F(v + s),$$

where $F(v)$ is the load-compliance relation for the system in the absence of roughness and s is the maximum peak-to-valley roughness height. Furthermore, these bounds can be tightened at the cost of transferring some of the long wavelength roughness features to the 'deterministic' part of the profile. Bounds on the incremental stiffness $P'(v)$ [and hence on $v'(P)$] can then be obtained, using the fact that $P'(v)$ is a non-decreasing function (Barber, 1974).

Unfortunately, there is no proof that the *second* derivative $P''(v)$ is monotonic, so this strategy cannot be directly extended to establish rigorous bounds on the function $v''(P)$ required in dissipation equations like (25), but one might reasonably use this strategy to obtain an order of magnitude estimate of the expected dissipation. For example, Eqs. (6.12) and (6.14) of Barber (2003) show that the incremental compliance function for a rough flat punch is bounded in the range

$$\frac{C_0}{E^*} < v'(P) < \frac{C_0}{E^*} + \frac{s}{P},$$

where E^* is the plane strain modulus and C_0 is a constant depending only on the planform of the punch and given by $1/2a$ for a circular punch of radius a.

If we estimate the actual relation as the average of these two bounds, we would have

$$v'(P) = \frac{C_0}{E^*} + \frac{s}{2P} \quad \text{and hence } v''(P) = -\frac{s}{2P^2}.$$

Notice that this expression is rather surprisingly independent of the elastic modulus, as is therefore the dissipation at small amplitudes calculated from integrals such as (25).

A similar conclusion follows from the assumption that increase of nominal contact pressure increases the number of actual contact areas, without changing their statistical distribution, as in the Greenwood and Williamson (1966) exponential asperity distribution theory and Persson's theory (Persson et al., 2002). For the latter case, Akarapu et al. (2011) quote a relation equivalent to

$$v'(P) = \frac{\gamma\sigma}{P} \quad \text{and hence } v''(P) = -\frac{\gamma\sigma}{P^2},$$

where σ is the standard deviation of the height distribution and γ is a dimensionless constant of order unity.

Notice that both these arguments lead to expressions of the form of Eq. (20) with $\beta = 1$, and in each case the multiplying constant C is linearly proportional to the RMS roughness amplitude—a result that can be rigorously established for elastic systems by pure dimensional arguments, assuming that the shape of the power spectral density (PSD) remains unchanged. In fact, this proportionality applies for any value of the exponent β and it follows from Eq. (22) that for given loading parameters, the dissipation W must vary linearly with σ.

5.3. Effect of plastic deformation

The results obtained in this paper depend on the assumption that the deformation is elastic. In the contact of rough surfaces, very high contact pressures are to be expected at the contacting asperities and this might be expected to cause some local plastic deformation at least during the first application of the load. The extent of plastic deformation in rough surface contact can be assessed by calculating a plasticity index such as $\psi = E^*\sigma_m/H$, where σ_m is the RMS surface slope and H is the hardness of the softer material (Johnson, 1985, Section 13.4), values in the range $\psi < 1$ indicating that the contact is predominantly elastic. Notice that σ_m is a scale-dependent parameter, but the bounds established in the preceding section show that dissipation is determined largely by the coarse scale features of the surface, where $\sigma_m \ll 1$. Even in cases where significant plastic deformation occurs on first loading, work hardening and residual stresses are likely to decrease plasticity effects under cyclic loading—a conclusion that is supported by experimental evidence in the related case of steady sliding (Archard, 1957).

5.4. Scale separation

In practice, the nominal contact problem – i.e. the problem in which roughness is ignored – may involve a variation in contact tractions with position. The present results can be used for this case, provided the length scale over which this variation occurs is significantly larger than the longest wavelength in the roughness profile. In this case, we first interpret the quantities P,Q as being forces per unit nominal area—i.e. as the local contact tractions $p(x,y),q(x,y)$ respectively, and the dissipation W as the dissipation per unit area $w(x,y)$. Assuming the 'smooth' contact problem has been solved and that the tractions satisfy the 'stick' condition $|q(x,y)| < fp(x,y)$ for all x,y, the present results can then be used to determine the local value of dissipation per unit area, which might be expected to correlate with the initiation of fretting damage.

6. Conclusions

The analysis presented here permits one to calculate the frictional dissipation per cycle for the contact between two rough surfaces with any surface profile provided only that (i) the contact is elastic, (ii) the local contact condition obeys the Coulomb friction law with a constant coefficient of friction and (iii) we have access to information about the *normal*

incremental stiffness function $v'(P)$. In contrast to previous studies of the problem, the method does not require any further assumptions about the rough surface contact process (such as an asperity model theory). It is also completely rigorous within the above constraints in the limiting case where Poisson's ratio $v=0$ and the extent of possible error for finite v is quantifiable (see Section 3.3) and likely to be well below other sources of uncertainty in the physical description. We also note that the required incremental stiffness is a quantity that is determined by the coarser levels of the multiscale surface roughness profiles (Barber, 2003) and that tight bounds can therefore be placed on it using a relatively unsophisticated numerical model.

Results are presented for the special case where the incremental stiffness is linear with normal load, but there is no special difficulty in evaluating the integrals for other cases. At small amplitudes, the dissipation for a given ratio $\hat{P}_1/\hat{Q}_1$ and given relative phase varies with the cube of the load amplitude for all possible contact geometries, and is linearly proportional to the second derivative of the compliance function $v''(P)$ evaluated at the mean normal load. It follows that experimental observations of less than cubic dependence on load amplitude cannot be explained by reference to surface roughness alone, or by any other geometric effect in the contact of half spaces. Dissipation is significantly larger when the normal and tangential components are out of phase.

Acknowledgements

M.C. is grateful to the Humboldt foundation for sponsoring his visit at Hamburg TUHH University.

Appendix A. Dissipation for the in-phase case $\phi=0$

If $\phi=0$, the ellipse in Fig. 1 degenerates to a straight line and we can eliminate t in Eqs. (1) to obtain

$$Q=\frac{Q_1(P-P_0)}{P_1}. \tag{26}$$

If $|Q_1|<fP_1$, no slip occurs during loading or unloading and the dissipation is identically zero. For $|Q_1|>fP_1$, microslip will occur during both loading segments. We suppose that $Q_1>0$ and we denote the end points of the line as $A(P_0+P_1,Q_1)$ and $B(P_0-P_1),-Q_1)$. Notice that we can set $Q_0=0$ without loss of generality since we know that this parameter does not affect the cyclic dissipation.

Routine applications of the Ciavarella–Jäger distribution show that during the segment BA, the tangential displacement u satisfies the equation

$$\frac{du}{dP}=\lambda f\left[v'(P)-2v'(P_Y)\frac{dP_Y}{dP}\right],$$

where

$$P_Y=\frac{1}{2}\left(P+(P_0-P_1)-\frac{(Q_1+Q)}{f}\right). \tag{27}$$

The work done during the segment BA can therefore be written as

$$W_{BA}=\int_{P_0-P_1}^{P_0+P_1} Q\frac{du}{dP}dP=W_1+W_2,$$

where

$$W_1=\lambda f\int_{P_0-P_1}^{P_0+P_1} Qv'(P)\,dP,\quad W_2=-2\lambda f\int_{P_0-P_1}^{P_0+P_1} Qv'(P_Y)\frac{dP_Y}{dP}\,dP.$$

Substituting for Q in W_1, using (26) we have

$$W_1=\frac{\lambda fQ_1}{P_1}\int_{P_0-P_1}^{P_0+P_1}(P-P_0)v'(P)\,dP. \tag{28}$$

For W_2 it is convenient to change the variable of integration to P_Y. This necessitates that we first use (26) to eliminate Q in (27) and then solve the resulting equation for P as a function of P_Y. The manipulations are routine and lead to the expression

$$W_2=\frac{2\lambda fQ_1}{(Q_1/f-P_1)}\int_{P_0-P_1}^{P_0-Q_1/f}[2P_Y-2P_0+P_1+Q_1/f]v'(P_Y)\,dP_Y. \tag{29}$$

An essentially similar procedure is used for the segment AB, where

$$\frac{du}{dP}=\lambda f\left[-v'(P)+2v'(P_Z)\frac{dP_Z}{dP}\right]$$

and

$$P_Z = \frac{1}{2}\left(P+(P_0+P_1)-\frac{(Q_1-Q)}{f}\right).$$

We obtain

$$W_{AB} = W_3 + W_4,$$

where

$$W_3 = -\frac{\lambda f Q_1}{P_1}\int_{P_0+P_1}^{P_0-P_1}(P-P_0)\nu'(P)\,dP, \tag{30}$$

$$W_4 = \frac{2\lambda f Q_1}{(Q_1/f+P_1)}\int_{P_0+P_1}^{P_0-Q_1/f}[2P_Z-2P_0-P_1+Q_1/f]\nu'(P_Z)\,dP_Z. \tag{31}$$

Notice incidentally that W_1 and W_3 are identical, since the sign change is negated by the reversal of order of the limits. The cyclic dissipation is then obtained as

$$W = W_{BA} + W_{AB} = W_1 + W_2 + W_3 + W_4.$$

The above expressions are given for a general compliance function $\nu'(P)$, but for the power-law form of Eq. (20) we can write the dimensionless expressions

$$\hat{W}_1 = \hat{W}_3 = \frac{\hat{Q}_1}{\hat{P}_1}\int_{1-\hat{P}_1}^{1+\hat{P}_1}(\hat{P}-1)\hat{P}^{\beta}\,d\hat{P},$$

$$\hat{W}_2 = \frac{2\hat{Q}_1}{(\hat{Q}_1-\hat{P}_1)}\int_{1-\hat{P}_1}^{1-\hat{Q}_1}[2\hat{P}-2+\hat{P}_1+\hat{Q}_1]\hat{P}^{\beta}\,d\hat{P},$$

$$\hat{W}_4 = \frac{2\hat{Q}_1}{(\hat{Q}_1+\hat{P}_1)}\int_{1+\hat{P}_1}^{1-\hat{Q}_1}[2\hat{P}-2-\hat{P}_1+\hat{Q}_1]\hat{P}^{\beta}\,d\hat{P}, \tag{32}$$

where we have dropped the suffices in P_Y,P_Z, since at this stage these are merely dummy variables of integration.

References

Akarapu, S., Sharp, T., Robbins, M.O., 2011. Stiffness of contacts between rough surfaces. Phys. Rev. Lett. 106, Art.204301.

Archard, J.F., 1957. Elastic deformation and the laws of friction. Proc. Roy. Soc. (London) A 243, 190–205.

Barber, J.R., 1974. Determining the contact area in elastic indentation problems. J. Strain Anal. 9, 230–232.

Barber, J.R., 2003. Bounds on the electrical resistance between contacting elastic rough bodies. Proc. Roy. Soc. (London) A 459, 53–66.

Barber, J.R., Davies, M., Hills, D.A., 2011. Frictional elastic contact with periodic loading. Int. J. Solids Struct. 48, 2041–4047.

Barenblatt, G.I., 1996. In: Scaling, Self-similarity and Intermediate AsymptoticsCambridge University Press, Cambridge.

Björkland, S., 1997. A random model for micro-slip between nominally flat surfaces. ASME J. Trib. 119, 726–732.

Bureau, L., Caroli, C., Baumberger, T., 2003. Elasticity and onset of frictional dissipation at a non-sliding multi-contact interface. Proc. R. Soc. (London) A 459, 2787–2805, doi:10.1098/rspa.2003.1146.

Campañá, C., Persson, B.N.J., Müser, M.H., 2011. Transverse and normal interfacial stiffness of solids with randomly rough surfaces. J. Phys.: Condens. Matter 23, 085001 (doi: 10.1088/0953-8984/23/8/085001).

Cattaneo, C., 1938. Sul contatto di due corpi elastici: distribuzione locale degli sforzi. Rend. Accad. Naz. Lincei 27, 342–348 434–436, 474–478 (in Italian).

Ciavarella, M., 1998a. The generalized Cattaneo partial slip plane contact problem. I—Theory, II—Examples. Int. J. Solids Struct. 35, 2349–2378.

Ciavarella, M., 1998b. Tangential loading of general three-dimensional contacts. ASME J. Appl. Mech. 65, 998–1003.

Eriten, M., Polycarpou, A.A., Bergman, L.A., 2011. Surface roughness effects on energy dissipation in fretting contact of nominally flat surfaces. ASME J. Appl. Mech. 78, 021011 doi: 10.1115/1.4002433.

Greenwood, J.A., Williamson, J.B.P., 1966. Contact of nominally flat surfaces. In: Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences, pp. 300–319.

Hyun, S., Pei, L., Molinari, J.-F., Robbins, M.O., 2004. Finite-element analysis of contact between elastic self-affine surfaces. Phys. Rev. E 70, 026117.

Jäger, J., 1998. A new principle in contact mechanics. ASME J. Trib. 120, 677–684.

Jang, Y.H., Barber, J.R., 2011. Effect of phase on the frictional dissipation in systems subjected to harmonically varying loads. Eur. J. Mech. A/Solids 30, 269–274.

Johnson, K.L., 1955. Surface interaction between elastically loaded bodies under tangential forces. Proc. Roy. Soc. (London) A 230, 531–548.

Johnson, K.L., 1961. Energy dissipation at spherical surfaces in contact transmitting oscillating forces. J. Mech. Eng. Sci. 3, 362–368.

Johnson, K.L., 1985. In: Contact MechanicsCambridge University Press, Cambridge.

Mindlin, R.D., 1949. Compliance of elastic bodies in contact. ASME J. Appl. Mech. 16, 259–268.

Mindlin, R.D., Mason, W.P., Osmer, T.F., Deresiewicz, H., 1952. Effects of an oscillating tangential force on the contact surfaces of elastic spheres. In: Proceedings of the 1st US National Congress of Applied Mechanics, 1951, pp. 203–208.

Munisamy, R.L., Hills, D.A., Nowell, D., 1994. Static axisymmetrical Hertzian contacts subject to shearing forces. ASME J. Appl. Mech. 61, 278–283.

Nowell, D., Dini, D., Hills, D.A., 2006. Recent developments in the understanding of fretting fatigue. Eng. Fract. Mech. 73, 207–222.

Paggi, M., Barber, J.R., 2011. Contact conductance of rough surfaces composed of modified RMD patches. Int. J. Heat Mass Transfer 54 (21–22), 4664–4672.

Persson, B.N.J., Bucher, F., Chiaia, B., 2002. Elastic contact between randomly rough surfaces: comparison of theory with numerical results. Phys. Rev. 65 (18), 184106.

Putignano, C., Afferrante, L., Carbone, G., Demelio, G. A new efficient numerical method for contact mechanics of rough surfaces. Int. J. Solids Struct., under review.

Sevostianov, I., Kachanov, M., 2008. Normal and tangential compliances of interface of rough surfaces with contacts of elliptic shape. Int. J. Solids Struct. 45, 2723–2736.
Wentzel, H., 2006. Modelling of Frictional Joints in Dynamically Loaded Structures: A Review. Technical Report, KTH Solid mechanics, Royal Institute of Technology 〈www-old.hallf.kth.se/forskning/publikationer/rapport_419.pdf〉.
Yang, C., Persson, B.N.J., 2008. Contact mechanics: contact area and interfacial separation from small contact to full contact. J. Phys. Condens. Matter 20, 215214.