

Frictional systems under periodic loads — History-dependence, non-uniqueness and energy dissipation

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2012 J. Phys.: Conf. Ser. 382 012002

(<http://iopscience.iop.org/1742-6596/382/1/012002>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 141.211.173.82

The article was downloaded on 25/06/2013 at 20:35

Please note that [terms and conditions apply](#).

Frictional systems under periodic loads — History-dependence, non-uniqueness and energy dissipation

J. R. Barber

Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109-2125,
U.S.A.

E-mail: jbarber@umich.edu

Abstract. Nominally static contacts such as bolted or shrink-fit joints typically experience regions of microslip when subjected to oscillatory loading. This results in energy dissipation, reflected as apparent hysteretic damping of the system, and also may cause the initiation of fretting fatigue cracks. Early theoretical studies of the Hertzian contact problem by Cattaneo and Mindlin were confirmed experimentally by Johnson, who identified signs of fretting damage in the slip annulus predicted by the theory.

For many years, tribologists assumed that Melan's theorem in plasticity could be extended to frictional systems — i.e. that if there exists a state of residual stress associated with frictional slip that is sufficient to prevent periodic slip in the steady state, then the system will shake down, regardless of the initial condition. However, we now know that this is true only if there is no coupling between the normal and tangential loading problems, as will be the case notably when contact occurs on a symmetry plane.

For all other cases, periodic loading scenarios can be devised such that shakedown occurs for some initial conditions and not for others. The initial condition here might be determined by the assembly protocol — e.g. the order in which a set of bolts is tightened — or by the exact loading path before the steady cycle is attained. This non-uniqueness of the steady state persists at load amplitudes above the shakedown limit, in which case there is always some dissipation, but the dissipation per cycle (and hence both the effective damping and the susceptibility to fretting damage) depends on the initial conditions. This implies that fretting fatigue experiments need to follow a well-defined assembly protocol if reproducible results are to be obtained. We shall also present results showing that when both normal and tangential forces vary in time, the energy dissipation is very sensitive to the relative phase of the oscillatory components, being greatest when they are out of phase.

With sufficient clamping force, 'complete' contacts (i.e. those in which the contact area is independent of the normal load) can theoretically be prevented from slipping, but on the microscale, all contacts are incomplete because of surface roughness and some microslip is inevitable. In this case, the local energy dissipation density can be estimated from relatively coarse-scale roughness models, based on a solution of the corresponding 'full stick' problem.

1. Introduction

Many engineering structures comprise a number of separate elastic components connected by nominally rigid connections such as bolted joints, shrink fits etc., which rely on friction to prevent relative motion. Typically these connections transmit a substantial normal 'clamping' force that

remains constant over time, and also time-varying forces that depend on the operation of the machine or the environmental conditions. The latter are often periodic in nature, and may result (for example) from machine vibration, operating cycles, or diurnal changes in environmental temperature. In this paper, we shall restrict attention to cases in which elastodynamic effects can be neglected and the contact problem regarded as quasi-static. This is not a serious restriction, since even in problems of impact, the quasi-static Hertzian contact theory has been shown to give good results [1].

Contact problems can be characterized as ‘complete’ or ‘incomplete’, depending on whether the contact area is determined by the geometry of the contacting components [complete], or whether it varies depending on the applied loads [incomplete]. In the latter case, the normal contact pressure usually tends to zero at the edge of the contact area, and a relatively small tangential force is sufficient to generate a region of ‘microslip’ — i.e. a region in which the deformability of the materials permits local tangential motion, even though the greater part of the contact area remains stuck. Under periodic loading, the microslip region will experience small amplitude reversed slip and work will be done against the frictional forces. If the Coulomb friction law applies, this will be reflected as apparent non-linear hysteretic (i.e. rate independent) damping in the overall behaviour of the structure [2]. Indeed, it has been estimated that in typical engineering structures, the energy dissipation in microslip at joints far exceeds that due to internal damping in the material [3]. Also, the extremely localized plastic deformation associated with this energy dissipation creates an environment favourable to the initiation and propagation of fretting fatigue failures that are a critical determining factor in the life of aeroengines [4, 5].

By contrast, if the contact is complete, it is theoretically possible to apply a sufficiently large clamping force to prevent microslip under any given tangential loading. However, we note that all real engineering surfaces are rough on the microscale, and hence all contact problems are incomplete in the sense that the ‘actual contact area’ (i.e. the area in which the surfaces are within the range of interatomic forces) always increases with increasing normal load. Thus, even in nominally complete contacts, we must anticipate some frictional energy loss and potential fretting damage [6].

2. Shakedown and Melan’s theorem

The slip that occurs during the first few cycles of loading causes a state of residual stress that generally tends to reduce the amplitude of subsequent periodic slip [7]. This is analogous with the behaviour of elastic-plastic structures under periodic loading, and for many years researchers in contact mechanics assumed that there should exist a frictional equivalent of Melan’s theorem [8] which states essentially that if an elastic-plastic structure can shake down (meaning that there exists an allowable state of residual stress that would be sufficient to prevent further plastic deformation under the given periodic loading), then it will do so.

The equivalent frictional theorem can be enunciated as “*If a set of time-independent tangential displacements at the interface can be identified such that the corresponding residual stresses when superposed on the time-varying stresses due to the applied loads cause the interface tractions to satisfy the conditions for frictional stick throughout the contact area at all times, then the system will eventually shake down to a state involving no slip, though not necessarily to the state so identified.*” This theorem was eventually proved by Klarbring *et al.*, both for discrete (e.g. finite element) systems [9] and continuous systems [10], but only under the rather restrictive condition that the normal and tangential contact problems should be uncoupled. In other words, the normal contact tractions must be unaffected if a tangential slip is imposed on the system at any part of the contact interface. The basis of both proofs is to define a scalar norm representing the difference between the instantaneous state and that which would occur at the same point in the cycle in the identified ‘safe shakedown state’. In the discrete case, this is expressed as a sum over the set of contact nodes, whereas in the continuum case, it is

represented by an integral over the contact area. It can then be shown that *if the system is uncoupled*, this norm changes only during periods (and at locations) where there is slip and that any such changes in the norm are necessarily negative by virtue of the conditions defining the Coulomb friction law.

For all other (i.e. coupled) cases, counter-examples to the frictional Melan's theorem can be found, in which there exist loading scenarios for which the system may shake down or not shake down, depending on the initial conditions [9, 10]. A simple heuristic argument for this conclusion is that when the system is coupled, any point that remains stuck throughout the loading cycle could be subjected to a slip displacement sufficiently small to avoid violating the friction law, and if the system is coupled, this may affect the normal tractions at other points sufficiently to permit or inhibit cyclic slip.

3. Behaviour of coupled systems

The frictional Melan theorem shows that we must anticipate history-dependence in the steady-state response of coupled systems. In fact, the above heuristic argument suggests that the theorem is a special case of a more general result (as yet unproven) that the steady-state frictional response to periodic loading in an elastic structure can depend on the initial conditions if and only if there exists coupling between the normal and tangential contact problems [13].

There is some anecdotal evidence for this conjecture, notably:

- (i) Fretting fatigue tests are very consistent for smooth 'Hertzian-like' contact geometries, but much more erratic when a rectangular indenter is pressed against a plane surface. The former geometry is reasonably approximated by two half planes, which involves no coupling when the materials are similar, whereas the latter involves significant normal-tangential coupling.
- (ii) Experimental measurements of the effective damping of nominally identical bolted joints exhibit significantly different behaviour, and even the same system, if disassembled and then reassembled, can give different results. Depending on the assembly protocol employed, this might be equivalent to a change in initial conditions.

To understand the reasons for this history-dependence, it is helpful to consider a two-dimensional discrete system, which might comprise a finite element solution of a two dimensional frictional elastic contact problem. The internal nodal displacements can be eliminated by standard static reduction procedures, leaving a system with two degrees of freedom (displacement components) at each contact node. The nodal forces (contact reactions) can then be expressed as

$$\begin{aligned} q_j &= q_j^w + A_{ji}v_i + B_{ij}w_i \\ p_j &= p_j^w + B_{ji}v_i + C_{ji}w_i, \end{aligned} \quad (1)$$

where p_j, q_j are the normal and tangential reactions respectively at node j , v_i, w_i are the tangential (slip) and normal (separation) displacements at node i , p_j^w, q_j^w are the nodal reactions that would be obtained if all the nodes were constrained in the position $v_i = w_i = 0$, and the summation convention is implied. Notice that the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ represent a partitioning of the stiffness matrix reduced to the contact nodes and hence \mathbf{A} and \mathbf{C} must be positive definite, but no such restriction applies to \mathbf{B} .

We consider the case where the time-varying external loads $\mathbf{p}^w(t), \mathbf{q}^w(t)$ can be written in the form

$$\mathbf{p}^w(t) = \mathbf{p}_0 + \lambda \mathbf{p}_1(t); \quad \mathbf{q}^w(t) = \mathbf{q}_0 + \lambda \mathbf{q}_1(t), \quad (2)$$

where $\mathbf{p}_0, \mathbf{q}_0$ represent the time-independent mean loads, $\mathbf{p}_1(t), \mathbf{q}_1(t)$ are periodic vectors with mean value zero and λ is a non-negative scalar load factor characterizing the magnitude of the periodic loading.

In the simple case where there are only two contact nodes and they are both in contact, $w_1 = w_2 = 0$ and the Coulomb friction inequalities $|q_j| \leq fp_i$ with equations (1) demand that

$$\begin{aligned} (A_{11} - fB_{11})v_1 + (A_{12} - fB_{12})v_2 &\leq fp_1^w - q_1^w && \text{I} \\ (A_{11} + fB_{11})v_1 + (A_{12} + fB_{12})v_2 &\geq -fp_1^w - q_1^w && \text{II} \\ (A_{21} - fB_{21})v_1 + (A_{22} - fB_{22})v_2 &\leq fp_2^w - q_2^w && \text{III} \\ (A_{21} + fB_{21})v_1 + (A_{22} + fB_{22})v_2 &\geq -fp_2^w - q_2^w && \text{IV} \end{aligned} \quad (3)$$

where f is the coefficient of friction and the equalities in these relations control the motions

$$\text{I: } \dot{v}_1 < 0; \quad \text{II: } \dot{v}_1 > 0; \quad \text{III: } \dot{v}_2 < 0; \quad \text{IV: } \dot{v}_2 > 0, \quad (4)$$

with the dot indicating the derivative with respect to time. We restrict attention to cases where the coefficient of friction is sufficiently small for the rate problem so defined to be well-posed. The limiting coefficient imposed by this condition can be determined from Klarbing’s P-matrix condition [11]. For higher coefficients of friction, conditions can be reached where the system becomes wedged, or where sudden elastodynamic state changes are precipitated even when the loading rate is extremely small [12].

Each of the constraints (3) excludes the region on one side of a certain straight line in v_i -space, suggesting that tracking the motion of the instantaneous state in this space might be a useful way of characterizing the evolution of the system [14]. If the nodes are to remain in contact, there must at all times t exist a region that is not excluded by any of the constraints. Furthermore, as the external loads p_i^w, q_i^w vary in time, the constraint lines (3) move whilst retaining the same slope, and hence cause the instantaneous state P to move either vertically or horizontally in accordance with (4), as illustrated in Figure 1.

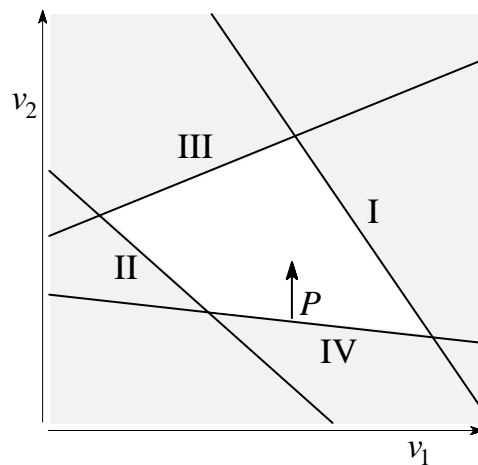


Figure 1: Motion of the instantaneous state P due to the advance of constraint IV ($\dot{v}_2 > 0$).

If shakedown is to be possible there must exist a safe shakedown region that is not excluded by any of the constraints *at any time t in the periodic loading cycle*. Ahn *et al.* [14] showed that the system would always shake down if the safe shakedown region so defined was a quadrilateral, but that the occurrence of shakedown depends on the initial conditions if it is a triangle.

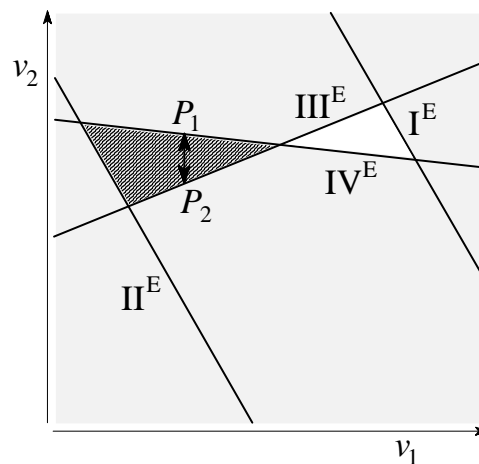


Figure 2: Cyclic slip limit cycle in the case where the safe shakedown region is triangular..

The reason for this is apparent from Figure 2, which shows the extreme positions I^E , II^E , III^E , IV^E of the four constraints — i.e. the positions that exclude the maximum amount of space in the diagram. These extreme positions are not generally reached at the same time, in which case it is possible for the system to oscillate along the line P_1P_2 implying cyclic slip at node 2, because the constraint II never moves far enough to the right to push P into the safe (unshaded) triangle. On the other hand, if we start from a position in the safe triangle, the constraints will never push us out of it and shakedown will be assured. This criterion is easily extended to the N -node system and is equivalent to the requirement that if all $2N$ extremal frictional constraints are active in defining the safe region in v_i -space, then the safe region is always reached. If one constraint or more is inactive in defining this region, then the steady state depends on the initial conditions and may comprise shakedown or cyclic slip with the node(s) associated with the inactive constraint(s) permanently stuck.

3.1. System memory

For the steady-state behaviour under periodic loading to depend on the initial conditions, the system must in some sense possess a ‘memory’ of these initial conditions. Now at any node that is instantaneously separated or slipping, we have two equations to determine the degrees of freedom v_i, w_i , namely $p_i = q_i = 0$ for separation and $w_i = 0, q_i = \pm fp_i$ for slip (the sign taken depending on the direction of slip). Thus, if an N -node system experiences a period when all nodes are simultaneously either slipping or separated, there will be $2N$ equations for $2N$ unknowns and the state is uniquely determined by the *instantaneous* values of $p_i^w(t), q_i^w(t)$. In this case, the subsequent evolution of the system cannot depend on the initial conditions. By contrast, if a node is instantaneously stuck, we have only one equation, $w_i = 0$, and the friction *inequality* $-fp_i < q_i < fp_i$. It follows that at any given time, the system memory must reside in the slip displacements v_i at those nodes that are instantaneously stuck.

We conclude that if all the nodes slip *at some time during the cycle*, but not all at the same time, the memory must somehow be exchanged between nodes during each cycle and we would expect this to lead to a degradation of memory and hence the attraction of the system to a unique steady state. Behaviour of this kind is illustrated in Figure 3, where the constraints are assumed to advance and recede in the sequence I, III, II, IV. Starting from two distinct points A, B , we find that the trajectory is attracted to a unique rectangular orbit.

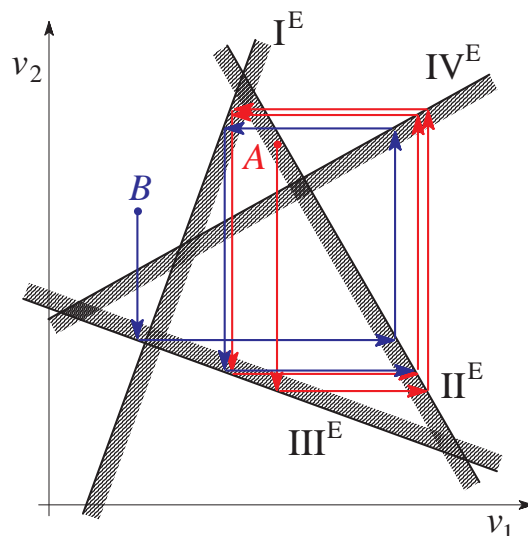


Figure 3: Attraction of a coupled system to a unique periodic steady state.

This issue has been explored for the two-node system by Andersson [15], who shows that for more complex loading patterns, conditions can arise where a steady periodic state is *unstable*, so that small perturbations from it grow with successive cycles until the sequence of stick/slip/separation states is altered, at which point we anticipate a qualitatively distinct stable (attractive) steady state. In some many cases, two such attractive orbits exist and the state eventually reached depends on which side of the unstable orbit the initial condition is located. However, in other cases, the partition of the initial condition space into catchment areas for the two stable orbits becomes fractal in the neighbourhood of the unstable orbit.

3.2. Effect of load factor

Consider now a system where the loading is defined by equation (2). If the scalar load factor λ is sufficiently small, the system will shake down for all initial conditions and we can identify a critical value of $\lambda = \lambda_1$ at which one of the extreme constraints ceases to be active and a larger value (λ_2) at which the safe shakedown region becomes null. In the range $\lambda_1 < \lambda < \lambda_2$ the system may or may not shake down depending on the initial conditions. For $\lambda > \lambda_2$, shakedown is impossible, but the steady state (and hence the frictional dissipation per cycle) still exhibits some dependence on initial conditions as long as a subset of nodes remains permanently stuck. However, at some even larger value λ_3 , we anticipate reaching a condition where all the nodes slip at least once during the cycle, so the system is attracted asymptotically to a unique steady state. Above a still larger value λ_4 there will exist at least one point in the cycle at which all the nodes either slip or separate, in which case the unique steady state is reached in the first one or two cycles. This kind of behaviour under gradually increasing periodic loading was documented by Jang and Barber [16] for a system comprising a two-dimensional elastic body containing a set of plane cracks. Of course, for systems that are nominally stuck, such as a bolted joint, there will usually be a permanently stuck region, implying that $\lambda < \lambda_3$.

4. Uncoupled systems

Conditions are considerably simpler if the system is uncoupled, and the most practically useful class of systems for which this condition is satisfied comprise those for which the two contacting bodies can be approximated by half spaces of similar materials [17].

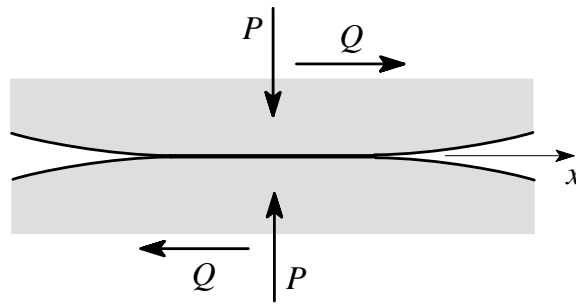


Figure 4: The Cattaneo-Mindlin problem.

4.1. Cattaneo and Mindlin's problem

The earliest theoretical study of microslip in a nominally static contact is that due to Cattaneo [18], who considered the problem of the Hertzian contact of Figure 4, comprising two quadratic surfaces loaded first by a normal compressive force P , which is then held constant whilst a tangential force Q is applied.

This loading scenario is illustrated in Figure 5, where we also identify the limiting lines $Q = \pm fP$ beyond which gross slip (sliding) would occur. Cattaneo showed that the tangential contact tractions $q(x, y)$ can be written as the superposition

$$q(x, y) = fp(x, y) - fp_S(x, y) , \tag{5}$$

where $p(x, y)$ are the corresponding (Hertzian) normal tractions and $p_S(x, y)$ are the normal tractions at a smaller fictitious load P_S defined by

$$P_S = P - \frac{Q}{f} . \tag{6}$$

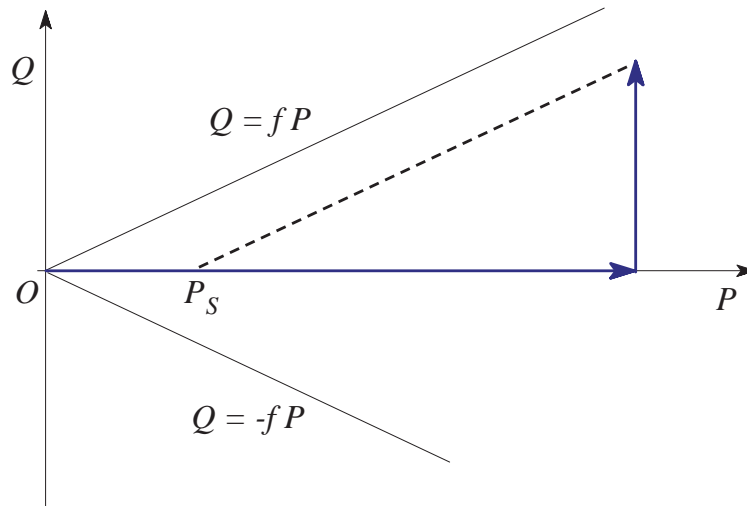


Figure 5: Load path for the Cattaneo-Mindlin problem.

Figure 5 shows a geometrical construction from which the load P_S can be determined. We also note that if $\mathcal{A}(P)$ denotes the contact area at a normal load P , microslip will occur in the region $\mathcal{A}(P) - \mathcal{A}(P_S)$ and stick in the region $\mathcal{A}(P_S)$.

It should be noted that this form of superposition was to some extent prefigured in a much earlier solution by Carter [19] to the related problem of a cylinder rolling on a plane. Cattaneo's result was 'rediscovered' apparently independently by Mindlin [20], who also considered the case

where the tangential force varies periodically in time [21, 22]. For the contact of two spheres, they predicted that microslip would occur in an annulus at the edge of the circular contact area, a result that was verified experimentally by Johnson [23] by observing the fretting damage to the surfaces in this region after a period of oscillation.

4.2. *The Ciavarella-Jäger theorem*

The superposition defined by equations (5.6) applies to a much broader class of contact problems, as was shown, again apparently independently, by Ciavarella [24, 25] and Jäger [26]. If the loading is as shown in Figure 5, it applies exactly to all two-dimensional contact problems between similar materials for which the bodies may be approximated by half planes, and in an approximate sense to corresponding three-dimensional problems. The approximation involved in the latter case is that the slip directions predicted deviate somewhat from the direction of the local frictional traction, particularly near the stick-slip boundary. This approximation is also inherent in the Cattaneo and Mindlin solutions and its magnitude was assessed by Munisamy *et al.* [27] and found to be extremely small. We shall see that Ciavarella and Jäger’s result is extremely useful in estimating the frictional energy dissipation in fairly general contacts, including those involving rough surfaces.

4.3. *Periodic loading*

The simplest form of periodic loading for the system of Figure 4 can be written

$$P(t) = P_0 + P_1 \cos(\omega t) ; \quad Q(t) = Q_0 + Q_1 \cos(\omega t - \phi) ,$$

which comprises a constant mean load (P_0, Q_0) and a sinusoidal periodic load in which the normal and tangential components (P_1, Q_1) may not necessarily be in phase. The resulting trajectory would then be elliptical as shown in Figure 6. The periodic cycle must be reached by some preliminary loading phase, which we here identify by the segment OA . In other words, the contact must be ‘assembled’ in some way.

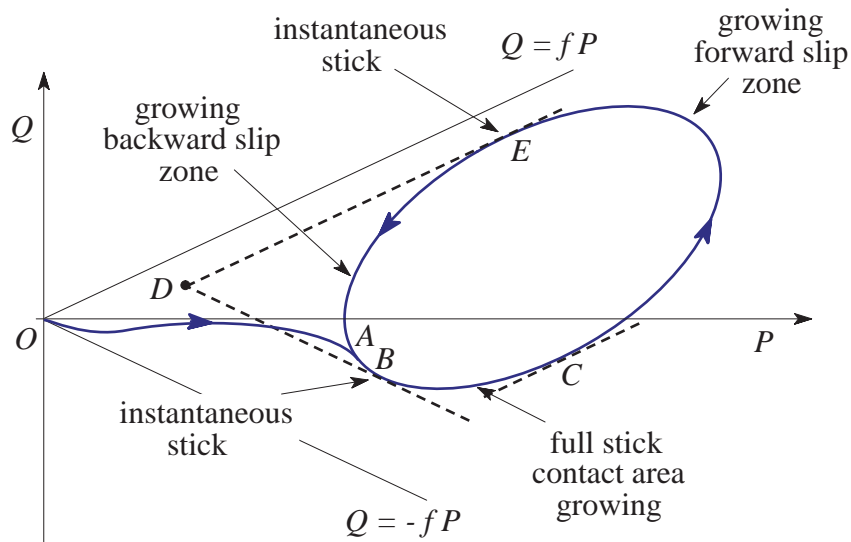


Figure 6: Periodic loading path.

The evolution of the traction distribution can be tracked by considering the incremental problem in which the normal and tangential loads $P(t), Q(t)$ increase by $\Delta P, \Delta Q$ during the small time increment Δt . If the (compressive) normal load is increasing ($dP/dt > 0$), and if $|dQ/dP| < f$, this incremental problem involves complete stick, so there is no microslip during

the phase BC in Figure 6. Once we pass C , a microslip zone begins to develop at the edge of the contact area and elementary calculations [28] show that the change in the traction distribution can be described by a superposition similar to that of Ciavarella and Jäger.

This state continues until we reach the point E , where the derivative dQ/dP is again of magnitude f , but the normal load is decreasing ($dP/dt < 0$). At this point the contact sticks everywhere instantaneously and a reverse microslip zone starts to grow from the edges of the contact area. As before, the resulting incremental tractions can be defined by superposed Ciavarella-Jäger distributions. One result of some interest is that the minimum extent of the stick region occurs just before the points E, B are reached and is equal to the contact area at a reduced load P_D , where the point D is identified by the geometrical construction shown in Figure 6. The tractions in the permanently stuck region $\mathcal{A}(P_D)$ depend on the initial load path OA , and this constitutes the memory of the system. However, these stored tractions do not influence those in the rest of the contact area, and hence they do not affect the frictional energy dissipation or the damage due to fretting, which are therefore independent of the initial assembly protocol.

4.4. In-phase loading

Figure 6 is presented for the case where the tangential and normal loads are out of phase. In the special case where the phase lag $\phi = 0$, the ellipse will condense to a straight line which is traversed back and forth during the cycle. The behaviour then depends critically on the magnitude of the ratio Q_1/fP_1 . If $Q_1/fP_1 < 1$, the entire loading phase ($dP/dt > 0$) occurs in a state of stick. Furthermore, during unloading ($dP/dt < 0$) the system passes through the same states as during loading, but in reverse order, so there is no tendency for microslip and no energy is dissipated at any point in the cycle. By contrast, if $Q_1/fP_1 > 1$, microslip occurs during both the loading and unloading phases and the tractions are given by Ciavarella-Jäger distributions, with the entire contact area becoming instantaneously stuck at the two extreme points.

4.5. Frictional energy dissipation

The frictional energy dissipation during the loading cycle can be calculated either in a local sense, using the instantaneous local tractions and slip displacements, or in a global sense, using the applied tangential force $Q(t)$ and the corresponding relative rigid-body displacement of the contacting bodies. In the former case, for two-dimensional geometries, we can obtain the dissipation *per unit area* as

$$W(x) = \oint q(x, t) \dot{v}(x, t) dt ,$$

where $v(x, t)$ is the local slip displacement, and the integral is performed around one complete loading cycle ($t = 0..2\pi/\omega$). The distribution of $W(x)$ through the contact area provides a good measure of the severity of fretting and its maximum is likely to correlate with the location of the first fretting fatigue cracks. For a Hertzian geometry, this maximum is found to be about half way between the edge of the contact area and the edge of the permanent stick zone [29]. The total dissipation per cycle \mathcal{W} can then be obtained by integrating $W(x)$ over the maximum extent \mathcal{A} of the interface — i.e.

$$\mathcal{W} = \int_{\mathcal{A}} W(x) dx .$$

However, a more direct approach is to evaluate the integral

$$\mathcal{W} = \oint Q(t) \dot{V}(t) dt , \quad (7)$$

where V represents the relative tangential rigid-body displacement of the two contacting bodies. The integrand in this case represents the incremental work done by the load $Q(t)$ during a time increment dt , but not all this work constitutes frictional dissipation — there is also a contribution to the elastic strain energy of the bodies. However, if the integral is performed over a complete cycle, the strain energy at the two limits will be identical and the integral will therefore represent the frictional dissipation per cycle.

The integral (7) can also be written

$$\mathcal{W} = \oint Q(t) \frac{dV}{dQ} dQ, \quad (8)$$

and hence can be evaluated if we can determine the *incremental tangential compliance* dV/dQ . Also, since the traction distribution corresponding to dQ is at all points in the cycle the derivative of one or more expressions of the form of equation (5), dV/dQ can be written down by a similar superposition associated with the contact areas $\mathcal{A}(P)$ and $\mathcal{A}(P_S)$. In the two dimensional problem, it is necessary to define a finite reference point in the elastic bodies to avoid mathematical problems associated with logarithmically unbounded displacements at infinity, but the choice of this reference point does not affect the final result and the compliances are found to be identical with the corresponding normal compliances dU/dP , where U is the rigid-body approach. Thus, the integral (8) can be determined provided we know the relation between normal force P and the corresponding normal approach for the particular profiles of the contacting bodies. In three-dimensional problems, this equivalence is not exact except for certain contact geometries such as the ellipse, but the ratio between normal and tangential compliance is quite tightly bounded, and for a wide range of three-dimensional contact geometries is closely approximated by

$$\frac{dV}{dQ} = \lambda \frac{dU}{dP} \quad \text{where} \quad \lambda = \frac{(2 - \nu)}{2(1 - \nu)} \quad (9)$$

and ν is Poisson's ratio [30, 31].

4.6. Effect of surface roughness

We have already remarked that most engineering surfaces are rough and hence the contact problem is incomplete, even in cases where the large scale geometry makes it appear complete. Thus, for example, if two coextensive rectangular blocks are pressed together, the contact area $\mathcal{A}(P)$ will comprise a set of 'actual contact areas' at the peaks of the asperities on the two surfaces, and these contact areas will grow with increasing normal load. This implies that the stick zones in the corresponding Ciavarella-Jäger distributions are defined by the set of actual contact areas $\mathcal{A}(P_S)$ at the fictitious load P_S . It also implies that there will be dissipation and the potential for fretting damage even in contacts where the solution of the 'smooth' contact problem predicts complete stick.

Roughness wavelengths are typically short compared with the dimensions of the nominal contact area, so a convenient way of incorporating the effect of roughness is first to solve the equivalent smooth problem, find the spatial distribution of the mean and alternating tractions $p(x, y, t), q(x, y, t)$, and then approximate the local dissipation per unit area due to roughness as being that obtained in a nominally uniform case where two plane surfaces with the given roughness are subjected to the same nominal traction cycle.

Many theories have been advanced to describe the contact of rough surfaces, the most widely used being those of Greenwood and Williamson [32] and Persson [33]. Though approaching the problem in very different ways, both authors conclude that the relation between macroscale quantities such as the normal force P and the incremental stiffness dP/dU is approximately linear because the population of microscale contact events retains approximately the same distribution

over a wide range of load levels. As P is increased, the number of actual contact areas increases. Small areas get larger, but a sufficient number of new smaller areas are established to preserve an approximately similar size distribution. Using Persson's theory, Akarapu *et al.* [34] argue that

$$\frac{dP}{dU} = \frac{P}{\gamma\sigma}, \tag{10}$$

where σ is the RMS roughness height and γ is a numerical constant of order unity. More generally, the incremental normal stiffness dP/dU due to roughness can be bounded, based only on the peak-to-valley height [35], and the bounding procedure implies that a relatively coarse numerical model of the surface is sufficient to give good estimates. In other words, only the long wavelength terms in the surface profile spectrum make significant contributions to the incremental stiffness.

Using (9, 10), we then have

$$\frac{dV}{dQ} = \frac{\lambda\gamma\sigma}{P}$$

and the integral in (8) can be performed numerically for the Ciavarella-Jäger distributions developed during the load cycle of Figure 6. Putignano *et al.* [36] showed that the resulting expression can be presented dimensionlessly in the form

$$\hat{W} \equiv \frac{\mathcal{W}}{\lambda\gamma\sigma f^2 P_0} = f(\hat{P}_1, \hat{Q}_1, \phi) \quad \text{where} \quad \hat{P}_1 = \frac{P_1}{P_0}; \quad \hat{Q}_1 = \frac{Q_1}{fP_0},$$

and for $\hat{Q}_1 < 0.5$ they found that $\hat{W} \sim \hat{Q}_1^3$. Notice that if the mean force P_0 is replaced by the mean normal traction p_0 , the corresponding expression for \mathcal{W} will then define the dissipation per unit area of the nominal contact interface. The expressions \hat{P}_1, \hat{Q}_1 are already normalized and will remain unchanged if forces are replaced by tractions.

We recall that for in-phase loading the dissipation is zero for $Q_1/fP_1 < 1$, which here corresponds to $\hat{Q}_1 < \hat{P}_1$. For larger values of \hat{Q}_1 , dissipation occurs for all phase angles, but is significantly larger when $\phi = \pm\pi/2$, as shown in Figure 7.

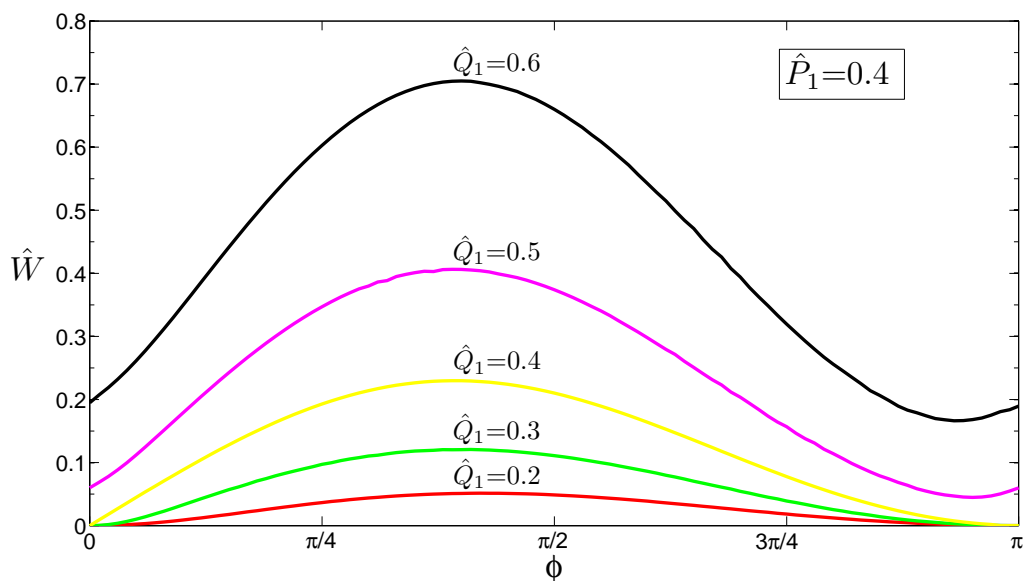


Fig.7. Dissipation as a function of relative phase ϕ for $\hat{P}_1 = 0.4$ and various values of \hat{Q}_1 .

5. Conclusions

The behaviour of elastic systems with frictional interfaces is qualitatively influenced by elastic coupling between the normal and tangential tractions and displacements. When coupling is present, the behaviour under cyclic loading, including frictional dissipation and susceptibility to fretting damage, will generally depend on the initial conditions, which might include the initial loading or assembly protocol. By contrast, for uncoupled systems, including the contact of bodies of similar materials that can be modelled as half spaces, the steady-state response is independent of initial conditions, and simple procedures, based on a theorem due to Ciavarella and Jäger, can be used to calculate the dissipation. In particular, it is possible to estimate the dissipation associated with surface roughness in nominally tight joints, and this is found to depend only on the long wavelength components in the surface roughness.

References

- [1] Johnson, K L 1985 *Contact Mechanics* Cambridge University Press, Cambridge
- [2] Wentzel H 2006 Modelling of frictional joints in dynamically loaded structures: a review Technical report KTH Solid mechanics Royal Institute of Technology (2006) http://www-old.hallf.kth.se/forskning/publikationer/rapport_419.pdf
- [3] Gaul L and Lenz J 1997 Nonlinear dynamics of structures assembled by bolted joints *Acta Mechanica* **125** 169–181 doi:101007/BF01177306
- [4] Murthy H Harish G and Farris T N 2004 Efficient modeling of fretting of blade/disk contacts including load history effects *ASME Journal of Tribology* **126** 56–64
- [5] Nowell D Dini D and Hills D A 2006 Recent developments in the understanding of fretting fatigue *Engineering Fracture Mechanics* **73** 207–222
- [6] Björkland S 1997 A random model for micro-slip between nominally flat surfaces *ASME Journal of Tribology* **119** 726–732
- [7] Churchman C M and Hills D A 2006 General results for complete contacts subject to oscillatory shear *Journal of the Mechanics and Physics of Solids* **54** 1186–1205
- [8] Melan E 1936 Theorie statisch unbestimmter Systeme aus ideal-plastischem Baustoff *Sitzungsber d Akad d Wiss Wien* **2A(145)** 195–218
- [9] Klarbring A Ciavarella M and Barber J R 2007 Shakedown in elastic contact problems with Coulomb friction *International Journal of Solids and Structures* **44** 8355–65
- [10] Barber J R Klarbring A and Ciavarella M 2008 Shakedown in frictional contact problems for the continuum *Comptes Rendus Mecanique* **336** 34–41.
- [11] Klarbring A 1999 Contact friction discrete mechanical structures and mathematical programming In: Wriggers P Panagiotopoulos P (Eds) *New Developments in Contact Problems* Springer 55–100
- [12] Ahn Y J 2010 Discontinuity of quasi-static solution in the two-node Coulomb frictional system *International Journal of Solids and Structures* **47** 2866–71
- [13] Barber J R 2011 Frictional systems subjected to oscillating loads *Annals of Solid and Structural Mechanics* **2** 45–55
- [14] Ahn Y J Bertocchi E and Barber J R 2008 Shakedown of coupled two-dimensional discrete frictional systems *Journal of the Mechanics and Physics of Solids* **56** 3433–40
- [15] Andersson L-E 2012 Private communication
- [16] Jang Y H and Barber J R 2011 Frictional energy dissipation in materials containing cracks *Journal of the Mechanics and Physics of Solids* **59** 583–594
- [17] Hills D A Nowell D and Sackfield A 1993 *Mechanics of elastic contact* Butterworth-Heinemann Oxford
- [18] Cattaneo C 1938 Sul contatto di due corpi elastici: distribuzione locale degli sforzi *Rendiconti dell'Accademia Nazionale dei Lincei* **27** 342–348 434–436 474–478 (In Italian)
- [19] Carter F C 1926 On the action of a locomotive driving wheel *Proceedings of the Royal Society of London* **A112** 151–157
- [20] Mindlin R D 1949 Compliance of elastic bodies in contact *ASME Journal of Applied Mechanics* **16** 259–268
- [21] Mindlin R D Mason W P Osmer T F and Deresiewicz H 1952 Effects of an oscillating tangential force on the contact surfaces of elastic spheres Proc 1st US Nat Cong Appl Mech 1951 203–208
- [22] Mindlin R D and Deresiewicz H 1953 Elastic spheres in contact under varying oblique forces *ASME Journal of Applied Mechanics* **75** 327–344
- [23] Johnson K L 1955 Surface interaction between elastically loaded bodies under tangential forces *Proceedings of the Royal Society of London* **A230** 531–548

- [24] Ciavarella M 1998a The generalized Cattaneo partial slip plane contact problem I-Theory II-Examples *International Journal of Solids and Structures* **35** 2349–78
- [25] Ciavarella M 1998b Tangential loading of general three-dimensional contacts *ASME Journal of Applied Mechanics* **65** 998–1003
- [26] Jäger J 1998 A new principle in contact mechanics *ASME Journal of Tribology* **120** 677–684
- [27] Munisamy R L Hills D A and Nowell D 1994 Static axisymmetrical Hertzian contacts subject to shearing forces *ASME Journal of Applied Mechanics* ASME JAppl Mech **61** 278–283
- [28] Barber J R Davies M and Hills D A 2011 Frictional elastic contact with periodic loading *International Journal of Solids and Structures* **48** 2041–47
- [29] Davies M Barber J R and Hills D A 2012 Energy dissipation in a frictional incomplete contact with varying normal load *International Journal of Mechanical Sciences* in press
- [30] Sevostianov I and Kachanov M 2008 Normal and tangential compliances of interface of rough surfaces with contacts of elliptic shape *International Journal of Solids and Structures* **45** 2723–2736
- [31] Campañá C Persson B N J and Müser M H 2011 Transverse and normal interfacial stiffness of solids with randomly rough surfaces *Journal of Physics: Condensed Matter* **23** 085001 [DOI: 10.1088/0953-8984/23/8/085001].
- [32] Greenwood J A and Williamson J B P 1966 The contact of nominally flat surfaces *Proceedings of the Royal Society of London* **A295** 300–319
- [33] Persson B N J 2001 Theory of rubber friction and contact mechanics *Journal of Chemical Physics* **115** 3840–61
- [34] Akarapu S Sharp T and Robbins M O 2011 Stiffness of contacts between rough surfaces *Physics Review Letters* **106** Art 204301
- [35] Barber J R 2003 Bounds on the electrical resistance between contacting elastic rough bodies *Proceedings of the Royal Society (London)* **459** 53–66
- [36] Putignano C Ciavarella M and Barber J R 2011 Frictional energy dissipation in contact of nominally flat rough surfaces under harmonically varying loads *Journal of the Mechanics and Physics of Solids* **59** 2442–54