Journal of the Mechanics and Physics of Solids 56 (2008) 3433–3440

Contents lists available at ScienceDirect

Journal of the Mechanics and Physics of Solids

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

Shakedown of coupled two-dimensional discrete frictional systems

Young Ju Ahn [a], Enrico Bertocchi [b], J.R. Barber [a,*]

[a] Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109-2125, USA
[b] Dipartimento di Ingegneria Meccanica e Civile, Università degli Studi di Modena e Reggio Emilia Via Vignolese 905, Modena, Italy

ARTICLE INFO

Article history:
Received 4 April 2008
Received in revised form
21 July 2008
Accepted 9 September 2008

Keywords:
Coulomb friction
Shakedown
Cyclic loading
Melan's theorem
History dependence

ABSTRACT

Recent results have established that Melan's theorem can be applied to discrete elastic systems governed by the Coulomb friction law only when the normal contact reactions are uncoupled from the tangential (slip) displacements. For coupled systems, periodic loading scenarios can be devised which lead to either shakedown or cyclic slip depending on the initial condition. Here we explore this issue in the simplest coupled system involving two contact nodes. The evolution of the system 'memory' is conveniently represented graphically by tracking the instantaneous condition in slip-displacement space. The frictional inequalities define directional straight line constraints in this space that tend to 'sweep' the operating point towards the safe shakedown condition if one exists. However, if the safe shakedown region is defined by a triangle in which two adjacent sides correspond to the extremal positions of the two frictional constraints for the same node, initial conditions can be found leading to cyclic slip. The critical value of a loading parameter at which this occurs can be determined by requiring that three of the four constraint lines intersect in a point. Below this value, shakedown occurs for all initial conditions. Similar concepts can be extended to multi-node systems.

1. Introduction

The classical Coulomb friction law states that for a given normal reaction p, no relative tangential motion (slip) can occur as long as the tangential reaction $|q|<fp$, where f is a constant, after which unlimited slip can occur at constant q. This represents a significant idealization, but its combination of simplicity and reasonable approximation to experimentally observed behaviour has led to its widespread use in the analysis of contacting systems.

Systems comprising elastic bodies in frictional contact share many features with bodies governed by an elastic–plastic constitutive law. Frictional slip and plastic deformation both involve energy dissipation and generally cause the instantaneous stress state to depend on loading history. Also, the idealization of elastic-behaviour in which unlimited plastic strain is allowed to occur at constant yield stress clearly has close parallels with the Coulomb friction law. These parallels and experience with the solution of specific example problems have led some authors (Churchman and Hills, 2006) to speculate that Melan's (1936) theorem might apply to frictional systems governed by the Coulomb friction law—in other words, that if a state of residual stress associated with frictional slip can be identified that would be sufficient to inhibit slip throughout a periodic loading cycle, then the system will in fact 'shake down', meaning that no further slip will occur after some initial transient. Churchman et al. (2006) explored this question in more detail in the context of

* Corresponding author. Tel.: +1734 936 0406; fax: +1734 615 6647.
E-mail address: jbarber@umich.edu (J.R. Barber).

doi:10.1016/j.jmps.2008.09.003

a simple two-node discrete elastic system. They identified loading regimes in which the long term behaviour involves shakedown, cyclic slip and frictional 'ratchetting', respectively, and presented these in the form of the Bree diagram. In this model, the normal force remains constant and hence the slip displacements retain associativity with respect to the time-varying reactions, thus meeting the conditions for Melan's theorem to apply. Similar considerations apply to the 'Tresca' friction law used by Antoni et al. (2007) in their study of ratchetting of a bushing in a connecting rod end.

By contrast, the Coulomb friction law is non-associative and there is therefore no reason to expect Melan's theorem to apply. This question was investigated in two previous papers in the context of discrete (Klarbring et al., 2007) and continuous (Barber et al., 2008) systems, respectively. A necessary condition for shakedown to occur in a discrete system is that there exist at least one vector of nodal slip displacements (the 'safe shakedown vector') such that the resulting time-varying nodal reactions satisfy the condition $|q_i| < fp_i$ at all nodes i throughout the loading cycle. The frictional Melan's theorem, if true, would then imply that this was also a sufficient condition for shakedown.

Klarbring et al. (2007) were able to establish such a theorem for two- and three-dimensional discrete systems that are uncoupled, meaning that changes in nodal slip displacements do not influence the normal contact tractions p_i. In this proof, shakedown was defined such that any slip that occurred during the initial cycles of loading caused the instantaneous slip displacements to approach the safe shakedown state monotonically in the sense of a certain norm. A similar result was established for continuous systems by Barber et al. (2008). Klarbring et al. (2007) also identified a class of two-dimensional systems with a degeneracy in the stiffness matrix, for which the theorem could be established if the friction coefficient f was less than a certain critical value. For all other discrete coupled systems, they showed that counter-examples could be identified—i.e. particular loading scenarios such that, depending on the initial conditions, the system may experience either shakedown or cyclic slip.

These counter-examples often require rather contrived loading. For example, the present authors have explored the problem of a rectangular elastic block in contact with a rigid plane surface and found that failure to reach the optimal safe shakedown state occurred only when the mean load induced large normal reactions at nodes near both edges of the block. There is therefore reason to hope that a reduced form of Melan's theorem might still apply to such systems under suitable restrictions on the loading history.

In this paper, we shall examine this question in the context of a simple two-dimensional coupled discrete system, comprising two contact nodes. In particular, we shall demonstrate that by considering the range of permissible slip displacements at the two nodes, it is possible to determine a lower bound on the amplitude of the cyclic load below which the system will always shake down, regardless of the initial transient, and an upper bound above which it cannot shake down. The methodology introduced is also capable of extension to more general multi-node systems.

2. The Coulomb friction law

Consider a two-dimensional elastic system that is supported against rigid-body motion and that makes contact with a rigid obstacle at a set of discrete nodes. Time-varying external loads are applied to the system, resulting in contact reactions

$$\boldsymbol{r}_i = [q_i, p_i]^{\mathrm{T}}, \tag{1}$$

where q_i, p_i are, respectively, the tangential and normal reactions at node i and we adopt the convention that compressive normal reactions are positive. The corresponding nodal displacements are

$$\boldsymbol{u}_i = [v_i, w_i]^{\mathrm{T}}, \tag{2}$$

where a positive value of normal displacement w_i corresponds to a gap between the elastic body and the obstacle.

The Coulomb friction law for node i can now be stated such that

$$w_i \geqslant 0, \quad p_i \geqslant 0, \tag{3}$$

$$w_i > 0 \Rightarrow p_i = q_i = 0, \tag{4}$$

$$p_i > 0 \Rightarrow w_i = 0, \tag{5}$$

$$|q_i| \leqslant fp_i, \tag{6}$$

$$|q_i| < fp_i \Rightarrow \dot{v}_i = 0, \tag{7}$$

$$0 < |q_i| = fp_i \Rightarrow \mathrm{sgn}(\dot{v}_i) = -\mathrm{sgn}(q_i), \tag{8}$$

where $f > 0$ is the coefficient of friction and a superposed dot denotes the time derivative.

As in Klarbring et al. (2007) we can employ a standard static condensation procedure to eliminate displacements at interior nodes of the discretized elastic body and hence write the reactions in the form

$$[\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots] = [\boldsymbol{r}_1^w, \boldsymbol{r}_2^w, \ldots] + \boldsymbol{\kappa}[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots], \tag{9}$$

where $\boldsymbol{r}_i^w$ are the reactions that would be generated by the external forces if all the nodal displacements were constrained to be zero—i.e. if the contact nodes were welded to the obstacle, and $\boldsymbol{\kappa}$ is the contact stiffness matrix which we take to be symmetric and positive definite. Using Eqs. (1) and (2), we can partition $\boldsymbol{\kappa}$ into three sub-matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, such that

$$q_j = q_j^w + A_{ji}v_i + B_{ij}w_i,$$
$$p_j = p_j^w + B_{ji}v_i + C_{ji}w_i, \tag{10}$$

where $\boldsymbol{A}$, $\boldsymbol{C}$ are symmetric and positive definite and $\boldsymbol{B}$ represents the coupling between tangential displacements and normal reactions which is not subject to these restrictions.

2.1. Critical coefficients of friction

The case where there is only one contact node was first introduced by Klarbring (1990) to illustrate and elucidate the anomalous behaviour of frictional systems when the coefficient of friction is large. In the present notation, if

$$f > f_c \equiv \frac{A_{11}}{|B_{11}|}, \tag{11}$$

conditions (3)–(8) may fail to define a unique quasi-static evolution for the single-node system under a given loading history. Uniqueness can be restored by using an elastodynamic formulation (Cho and Barber, 1998), but dynamic instabilities are then predicted (Martins et al., 1995; Adams, 1996), resulting in rapid transitions from one state to another at certain points in the evolution. In the limit of low mass, where we would expect to recover the quasi-static solution, this translates into 'displacement jumps'—i.e. a loss of continuity of nodal displacements under continuously varying loads (Cho and Barber, 1998; Martins et al., 1994).

The coefficient f_c plays a pivotal role in determining the behaviour of the single-node system. It is the sole eigenvalue in Hild's eigenvalue problem (Hassani et al., 2003) and it was shown in Klarbring et al. (2007) that Melan's theorem applies to this system if and only if $f < f_c$. Also, the single-node system is capable of becoming 'wedged' if and only if $f > f_c$, meaning that it can sustain a deformed configuration in the absence of external loads (Barber and Hild, 2006). By contrast, if $f < f_c$, it always returns to the undeformed configuration when the external loads are removed.

For multi-node systems, the behaviour is more complex (Andersson and Klarbring, 2001). Hassani et al. (2003) establish some relations between the uniqueness of solution and the eigenvalue problem. Also, the solution is always non-unique if the system is capable of wedging (Andersson, 2008). In this case, we can always generate scenarios in which discontinuous displacement jumps occur. For example, suppose the system is wedged and tangential external loads are then gradually applied until all the wedged nodes just exceed the limiting friction condition (6) in the direction tending to reduce the wedging displacements. Relaxation due to incremental slip now causes a larger change in the normal force than is required to stay on the slip constraint and the system will accelerate (actually towards a state involving separation) with no further change in external forces. In effect, slip motion in the opposite direction to that required to establish the wedged state is dynamically unstable, as demonstrated by Cho and Barber (1998) for the single-node system. However, for multi-node systems we shall show that wedging is a sufficient but not a necessary condition for the transient evolution to be discontinuous. Notice also that multi-node systems can exhibit many different wedging modes.

For the most part, we shall restrict attention in the present paper to coefficients of friction that are low enough for wedging to be impossible and for the quasi-static evolution to be continuous and unique.

3. The two-node system

Consider a system comprising just two contact nodes, $i = 1, 2$ and suppose that at some point in the loading cycle both nodes are in contact, so that $w_1 = w_2 = 0$ and

$$q_j = q_j^w + A_{ji}v_i, \quad p_j = p_j^w + B_{ji}v_i, \tag{12}$$

from Eq. (10). For this state to be physically admissible for a given loading vector $\boldsymbol{r}^w$, the Coulomb friction law demands that we satisfy inequality (6) at each of the two nodes $j = 1, 2$. Using Eq. (12), this implies that

$$-fp_j^w - fB_{ji}v_i \leqslant q_j^w + A_{ji}v_i \leqslant fp_j^w + fB_{ji}v_i \tag{13}$$

and hence

$$(A_{11} - fB_{11})v_1 + (A_{12} - fB_{12})v_2 \leqslant fp_1^w - q_1^w \quad \text{I}$$
$$(A_{11} + fB_{11})v_1 + (A_{12} + fB_{12})v_2 \geqslant -fp_1^w - q_1^w \quad \text{II}$$
$$(A_{21} - fB_{21})v_1 + (A_{22} - fB_{22})v_2 \leqslant fp_2^w - q_2^w \quad \text{III}$$
$$(A_{21} + fB_{21})v_1 + (A_{22} + fB_{22})v_2 \geqslant -fp_2^w - q_2^w \quad \text{IV} \tag{14}$$

3.1. Admissible regions in v_1, v_2 space

In frictional problems, the 'memory' of the system resides in the values of tangential displacements at nodes that are instantaneously stuck. Once a node slips or separates, its condition is determined by an equation and its contribution to the system memory is erased. On the other hand, memory is 'created' at nodes which transition from separation or slip to stick (Dundurs and Comninou, 1983). The evolution of the system memory can therefore conveniently be represented graphically in v_1, v_2 space.

Each of the four inequalities I–IV in Eq. (14) defines a straight line boundary in this space and excludes the region on one side of the line. The admissible values of v_1, v_2 are defined by the intersection of the admissible regions for each inequality. A typical case is illustrated in Fig. 1, where only the central white region is admissible. Notice that since the pairs I, II and III, IV correspond to slip in opposite directions at the same node, they will not intersect except at the critical point where separation occurs at that node. Thus, when a quadrilateral is defined as in Fig. 1, I and II will generally represent opposite sides.

The location of the four constraint lines depends on the instantaneous values of q_i^w, p_i^w, but their slopes depend only on the matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and the coefficient of friction f.

The admissible region may take various shapes depending on $\boldsymbol{r}^w$ and the slopes of the lines. In particular, we may have

(i) A quadrilateral, as in Fig. 1.
(ii) A triangle, as in Fig. 2(a).
(iii) A region open to the point at infinity and bounded by two, three or four straight line segments (the case of four segments is illustrated in Fig. 2(b)).
(iv) A null space, which implies that our initial assumption of contact at both nodes is false, so one or both of the nodes must separate under the given loads.

Notice that the vertex A of the triangle in Fig. 2(a) differs from the other two vertices in that it is an intersection between the two constraints for the same node. It follows that this point corresponds to the case where node 2 is on the point of separating. By contrast, the other two vertices correspond to points where both nodes are slipping or on the point of slipping.

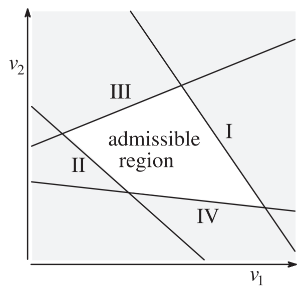

Fig. 1. Intersection of the admissible regions (values of v_1, v_2) that satisfy constraints I–IV.

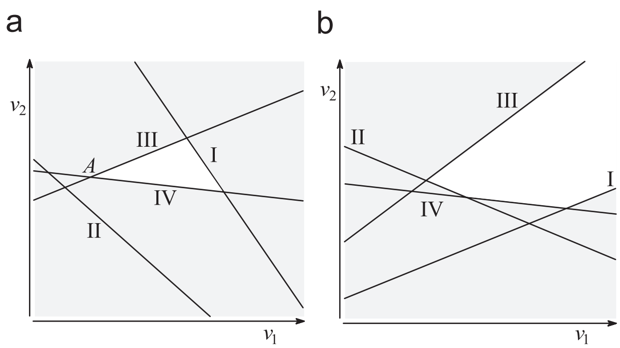

Fig. 2. Configurations of the constraints leading to an admissible region that is (a) a triangle or (b) a region open to infinity.

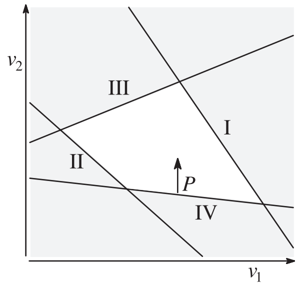

Fig. 3. Motion of the instantaneous operating point P due to the advance of constraint IV.

3.2. Wedging

In the special case where there is no external load ($\boldsymbol{r}_i^w = \boldsymbol{0}$), all four constraint lines pass through the origin but have the same slope as under any other value of load. If these slopes are such as to lead to a bounded admissible region, as in Figs. 1 and 2(a), this region will then shrink to a point, showing that the system relaxes to a unique position when the external loads are removed. However, under the same relaxation, case (iii) and Fig. 2(b) would reduce to a sector open to infinity, showing that the system is capable of becoming wedged, as defined in Section 2.1.

3.3. Transient evolution of the system

At each instant during the transient process (and assuming the no-separation condition is satisfied), we must have a figure similar to one of Figs. 1 and 2 and the instantaneous values of v_1, v_2 define an operating point within or on the boundary of the admissible region. As the external load $\boldsymbol{r}_i^w(t)$ changes in time, the four boundaries generally move. Slip will occur only when the operating point is on one or more boundaries and when the motion of these boundaries due to the change in load would otherwise cause the operating point to fall outside the new admissible region. Thus, the movement of the operating point can be viewed as a particle in the figure that is 'swept' over the plane by the four moving constraints. Notice, however, that the motion of the operating point is not generally orthogonal to the corresponding constraint line, but is directed along the appropriate axis. For example if the instantaneous operating point is at P in Fig. 3 and the constraint line IV advances so as to reduce the admissible region, the motion of P must be in the direction of the positive v_2-axis as shown ($\dot{v}_2 > 0$). In the same way, advance of constraint III will cause P to move downwards ($\dot{v}_2 < 0$) and that of constraints I, II cause slip at node 1 in the directions $\dot{v}_1 < 0$, $\dot{v}_1 > 0$, respectively.

It is this lack of orthogonality which constitutes the non-associative behaviour of the frictional system and we shall see later that it can permit a state of cyclic slip to occur even when a safe shakedown state exists for the given loading. Slip will cause P to remain on the corresponding constraint line until it reaches an intersection between two constraints or until the corresponding constraint 'recedes' leaving P strictly within the permissible region. In the former case, slip will also be initiated at the second node and P will remain in the corresponding corner until one or other of the intersecting constraints recedes.

This evolution mechanism places restrictions on the slopes of the four lines I–IV and hence on the magnitude of the coefficient of friction. For example, if constraint IV is rotated clockwise until it passes the vertical, it will no longer be able to push P upwards when it advances. If this condition holds and the other constraints are inactive, there is no admissible state involving both nodes in contact and we anticipate a discontinuous transition to separation at node 2. To avoid this possibility at each of the four constraints, the coefficient of friction must satisfy the condition

$$f < \min\left(\frac{A_{11}}{|B_{11}|}, \frac{A_{22}}{|B_{22}|}\right). \tag{15}$$

Comparison with Eq. (11) shows that this defines the critical coefficient of friction for a single-node system constructed by anchoring one of the two nodes. A related situation arises if the slope of the lines falls into the category (iii) of Section 3.1. In Fig. 2(b), all the constraints satisfy (15) and in particular, advance of I pushes the operating point P to the left $\dot{v}_1 < 0$ and III pushes it down ($\dot{v}_2 < 0$). However, if as a result of such motion P reaches an intersection between I and III, further advance of either of these constraints leads to an impossible scenario and once again we anticipate that there will be a discontinuous transition to a state involving separation. Of course, this case is also capable of wedging and we have already demonstrated in Section 2.1 that this implies the existence of loading scenarios involving displacement discontinuities.

3.4. Periodic loading

We suppose that the system is exposed to periodic external loads which can be expressed in the form

$$\boldsymbol{r}_i^w(t) = \boldsymbol{r}_i^0 + \lambda \boldsymbol{r}_i^1(t), \tag{16}$$

where $\boldsymbol{r}_i^0$ is a mean load that is independent of time t, $\boldsymbol{r}_i^1(t)$ is a periodic load with zero mean value and λ is a scalar load factor. The four components of $\boldsymbol{r}_i^1(t)$ are independent functions of time, so no restrictions are imposed on the form of the loading cycle. In particular, it is not necessarily sinusoidal and the separate components are not assumed to be in phase. The four constraint lines will now advance and recede in a periodic manner and each will experience a time (generally different for each constraint) at which the region excluded is a maximum. These extreme positions can be defined as

$$\begin{aligned}
&(A_{11} - fB_{11})v_1 + (A_{12} - fB_{12})v_2 \leqslant fp_1^0 - q_1^0 + \lambda(fp_1^1 - q_1^1)_{\min} \quad \mathrm{I}^{\mathrm{E}} \\
&(A_{11} + fB_{11})v_1 + (A_{12} + fB_{12})v_2 \geqslant -fp_1^0 - q_1^0 + \lambda(-fp_1^1 - q_1^1)_{\max} \quad \mathrm{II}^{\mathrm{E}} \\
&(A_{21} - fB_{21})v_1 + (A_{22} - fB_{22})v_2 \leqslant fp_2^0 - q_2^0 + \lambda(fp_2^1 - q_2^1)_{\min} \quad \mathrm{III}^{\mathrm{E}} \\
&(A_{21} + fB_{21})v_1 + (A_{22} + fB_{22})v_2 \geqslant -fp_2^0 - q_2^0 + \lambda(-fp_2^1 - q_2^1)_{\max} \quad \mathrm{IV}^{\mathrm{E}}
\end{aligned} \tag{17}$$

We can plot diagrams similar to Figs. 1 and 2 corresponding to these *maximum* excluded regions and any remaining admissible region will now be admissible at all times t during the loading cycle. If this region is not null, it implies that there exists a safe shakedown state for the system under the given cyclic loading. We shall demonstrate that shakedown *always* occurs if the safe shakedown region is defined by a quadrilateral, but that if the shakedown region is a triangle, initial conditions can be found in which the long term state is cyclic slip.

We consider the latter case first. We suppose that the instantaneous admissible region at all times during the loading cycle is a quadrilateral, so that separation is not possible, but that the safe shakedown region defined by the extremal constraints I^{E}, $\mathrm{III}^{\mathrm{E}}$, IV^{E} is triangular as shown in Fig. 4. This implies that III and IV reach their extreme positions at different times during the cycle. The remaining extremal constraint II^{E} is inactive in regard to determining the possibility of shakedown, but it also forms a triangle with constraints $\mathrm{III}^{\mathrm{E}}$, IV^{E} which we shall call the 'complementary' triangle.

Since the instantaneous admissible region is always a quadrilateral *ex hyp.*, there can be no time at which the instantaneous positions of III, IV intersect to the right of II^{E} and there must therefore be some instant at which an operating point P_1 on the IV^{E} boundary of the complementary triangle is admissible. Suppose we start from this initial condition. At some later time, constraint III advances to its extreme position $\mathrm{III}^{\mathrm{E}}$, pushing P down to the point P_2, as shown in Fig. 4. Constraint III now recedes, but IV then advances pushing P back up to P_1. This scenario permits indefinite cyclic slip to occur, since the only slip motion permitted on these boundaries involves the displacement v_2 and this does not enable P to make any progress towards the safe shakedown region. Notice that if instead the friction law were associative, the motion caused by each constraint advancing would be normal to the constraint boundary, resulting in an accumulation of motion of the operating point to the right which would tend asymptotically to the apex of the safe triangle.

This cyclic slip mechanism depends critically on the existence of the complementary triangle, which is a region on the 'inadmissible' side of the two constraints corresponding to a single node, but within the region that is admissible *throughout the loading cycle* in regard to the other node, which therefore remains stuck during the limit cycle. By contrast, consider the motion of the operating point due to the advance and recession of two constraints at different nodes. If constraint IV advances in Fig. 5, the operating point P_1 will move up to P_2 at which point slip will also be initiated at node 1. Further advance of IV will cause P to remain in the intersection of the two constraints (slipping at both nodes), following the path to P_3. If IV now recedes and I advances, slip will occur at node 1 only and will move P to the left to P_4 as shown. On the next loading cycle, P will move up and then to the left due to the advance of IV and I, respectively. This two step pattern will then be repeated, leading to an asymptotic and monotonic approach to the safe shakedown region (labelled SD in Fig. 5). Thus, when the two active constraints are associated with different nodes, the same sequence of advancing and receding constraints leads P towards the safe shakedown region. Notice that although approach to the safe shakedown

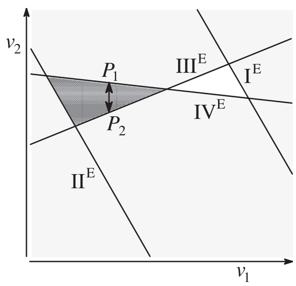

Fig. 4. Cyclic slip limit cycle in the case where the safe shakedown region is triangular.

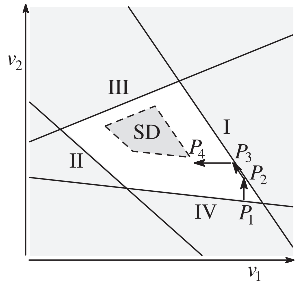

Fig. 5. Progression of the operating point towards the safe shakedown region (SD) due to motion of constraints for two different nodes.

region is monotonic, it can require a theoretically infinite number of loading cycles with geometrically decreasing slips. This falls within the definition of shakedown adopted by Klarbring et al. (2007).

3.5. Limiting values of the loading parameter λ

If the loading parameter λ in Eq. (16) is increased, the regions excluded by the extremal constraints I^{E}–IV^{E} is increased, reducing the size of the safe shakedown region. This region must be a quadrilateral for $\lambda = 0$, since we assume that the mean load does not permit separation at either node. Thus, there will exist some critical value $\lambda = \lambda_1$ at which the quadrilateral defined by the extremal constraints degenerates to a triangle. At this value, the edge of the quadrilateral that is about to disappear corresponds to a constraint line that passes through the apex of the newly generated triangle, and hence three of the extremal constraint lines pass through the same point.

Values of λ at which this condition is satisfied can be found by enforcing the equality in any three of I^{E}–IV^{E} and solving the resulting linear algebraic equations for v_1, v_2, λ. Since any one of the four extremal constraints may be inactive, four such values can be obtained which we label such that $\lambda_1 < \lambda_2 < \lambda_3 < \lambda_4$. They must all be positive, since the definition ensures that increasing λ reduces the admissible region. Clearly the critical value as defined above must be the lowest of the four, λ_1.

Increase in λ above λ_1 will reduce the size of the triangular safe shakedown region and this will become null for the case illustrated in Fig. 4 when the extremal constraints I^{E}, $\mathrm{III}^{\mathrm{E}}$, IV^{E} intersect in a point, corresponding to the value λ_2. The remaining values λ_3, λ_4 correspond to intersections of three constraints within a region that is inadmissible with regard to the fourth inactive constraint and hence have no physical significance.

Thus, we can define the shakedown behaviour of the system by solving the constraints as equalities in four groups of three, and sorting the resulting values of λ in ascending order. For $0 < \lambda < \lambda_1$, the system will always shake down. For $\lambda_1 < \lambda < \lambda_2$ it may shake down or may reach a limit cycle, depending on the initial conditions, and for $\lambda > \lambda_2$ it will never shakedown.

3.6. Separation

So far, we have restricted attention to systems in which the instantaneous admissible zone is always a quadrilateral, which implies that contact is maintained throughout the transient process. However, it is possible to remove this restriction if the coefficient of friction is below the value at which the transient evolution problem can exhibit displacement discontinuities. This excludes systems that are capable of wedging.

Suppose that constraint IV in Fig. 1 advances until the instantaneous feasible domain has the triangular form of Fig. 2(a). If the original operating point *P* lies to the left of point *A* in this figure, advance of IV will move it upwards until it lies at the instantaneous intersection of III and IV. Further advance of IV will then cause separation at node 2. We could still plot the location of $P(v_1, v_2)$ during the separation phase, but this would not be useful, since the assumption of separation at node 2 would change the equations for the reactions at node 1 and hence modify the slope and location of constraints I and II.

If the safe shakedown region is a quadrilateral, there must be some subsequent time at which constraint II lies to the right of the point *A*, at which point separation at node 2 is no longer possible. Under the assumption of continuity of displacements, the operating point must lie exactly at the intersection between III and IV at the instant that contact is re-established. Subsequent motion of *P* then follows the rules already established for both nodes being in contact and in particular *P* will be swept to the right of *A* by constraint II on its path to its extreme position, thus preventing separation in any subsequent cycle. Thus, it remains true that the system will shake down if the safe shakedown region is a quadrilateral, even if separation occurs at some point during the transient process.

3.7. Multi-node systems

It is clear that the same methodology could be extended to discrete systems with more than two nodes. As long as contact is maintained at all N nodes, the instantaneous slip displacements can be characterized as a point $P(v_1, v_2, \ldots, v_N)$ in an N-dimensional space and the admissible region at any instant in the load cycle will then be defined by a set of $2N$ constraints, each defined by the region on one side of a hyperplane. Shakedown is possible if and only if the region defined by the extremal positions of these constraints is non-null. If such a safe shakedown space exists, the motion of the constraints generally tends to push P towards it as in the two-dimensional case. However, this mechanism will fail a subset of initial conditions if any two adjacent facets of the safe space are defined by the two extremal constraints corresponding to a single node. If the mean load is such as to guarantee contact at all nodes and if λ is sufficiently small, all the constraints will be active in determining the safe space and the system will always shake down. Critical values of λ can be defined as in Section 3.5 by requiring that $N+1$ constraint planes intersect in a point—i.e. by solving a subset of $N+1$ constraints as equalities for the N values of v_i and λ. As λ is increased, the shape of the safe region will change at each of these critical values, and one of them will correspond to the upper shakedown limit—i.e. the case where the safe region just becomes null.

4. Conclusions

The evolution of nodal slip v_i for the coupled two-node system is conveniently described by a process in which the frictional constraints 'sweep' the instantaneous operating point $P(v_1, v_2)$ about the $v_1 v_2$-plane. Shakedown is possible if and only if the extremal positions of these constraints define a non-null 'safe shakedown region'.

Critical values of the load factor λ on the periodic component of load can be determined by solving any three of the four frictional constraints as equalities for v_1, v_2, λ. Four such values are obtained and below the smallest one, the system shakes down from all initial conditions. Between the first and second root, shakedown depends on the initial condition, whereas above the second root, shakedown is impossible. A similar strategy can be devised for multi-node discrete frictional systems. These results apply for coefficients of friction below that at which the incremental problem involves displacement discontinuities and/or multiple solutions.

Acknowledgements

We wish to thank Professors Michele Ciavarella and Andrea Spagnoli for their contributions to these results. We also thank the Rotary Foundation Ambassadorial Scholarship Program and the Rotary Club of Bologna, Italy (Enrico Bertocchi) and the Electric Power National Scholarship Program of the Korean Ministry of Commerce, Industry and Energy (MOCIE) (Youngju Ahn) for financial support.

References

Adams, G.G., 1996. Self-excited oscillations in sliding with a constant friction coefficient—a simple model. ASME J. Tribol. 118, 819–823.
Andersson, L.-E., 2008. Private communication.
Andersson, L.-E., Klarbring, A., 2001. A review of the theory of static and quasi-static frictional contact problems in elasticity. Philos. Trans. R. Soc. London 359, 2519–2539.
Antoni, N., Nguyen, Q.-S., Ligier, J.-L., Saffré, P., Pastor, J., 2007. On the cumulative microslip phenomenon. Eur. J. Mech. A/Solids 26, 626–646.
Barber, J.R., Hild, P., 2006. On wedged configurations with Coulomb friction. In: Peter, W., Udo, N. (Eds.), Analysis and Simulation of Contact Problems. Springer, Berlin pp. 205–213.
Barber, J.R., Klarbring, A., Ciavarella, M., 2008. Shakedown in frictional contact problems for the continuum. C. R. Méc. 336, 34–41.
Cho, H., Barber, J.R., 1998. Dynamic behavior and stability of simple frictional systems. Math. Comput. Modeling 28, 37–53.
Churchman, C.M., Hills, D.A., 2006. General results for complete contacts subject to oscillatory shear. J. Mech. Phys. Solids 54, 1186–1205.
Churchman, C.M., Korsunsky, A.M., Hills, D.A., 2006. The application of plasticity principles to friction. J. Strain Anal. 41, 323–328.
Dundurs, J., Comninou, M., 1983. An educational elasticity problem with friction—Part III: general load paths. ASME J. Appl. Mech. 50, 77–84.
Hassani, R., Hild, P., Ionescu, I.R., Sakki, N.-D., 2003. A mixed finite element method and solution multiplicity for Coulomb frictional contact. Comput. Methods Appl. Mech. Eng. 192, 4517–4531.
Klarbring, A., 1990. Examples of non-uniqueness and non-existence of solutions to quasi-static contact problems with friction. Ing.-Arch. 60, 529–541.
Klarbring, A., Ciavarella, M., Barber, J.R., 2007. Shakedown in elastic contact problems with Coulomb friction. Int. J. Solids Struct. 44, 8355–8365.
Martins, J.A.C., Montiero Marques, M.D.P., Gastaldi, F., 1994. On an example of non-existence of solution to a quasistatic frictional contact problem. Eur. J. Mech. A/Solids 13, 113–133.
Martins, J.A.C., Guimaraes, J., Faria, L.O., 1995. Dynamic surface solutions in linear elasticity and viscoelasticity with frictional boundary conditions. J. Vib. Acoust. 117, 445–451.
Melan, E., 1936. Theorie statisch unbestimmter Systeme aus ideal-plastichem Baustoff. Sitzungsber. d. Akad. d. Wiss., Wien 2A(145), 195–218.