Available online at www.sciencedirect.com
ScienceDirect
International Journal of Solids and Structures 44 (2007) 8355–8365
INTERNATIONAL JOURNAL OF SOLIDS AND STRUCTURES
www.elsevier.com/locate/ijsolstr

Shakedown in elastic contact problems with Coulomb friction

A. Klarbring [a], M. Ciavarella [b], J.R. Barber [c,*]

[a] *Division of Mechanics, Institute of Technology, Linköping University, SE-581 83 Linköping, Sweden*
[b] *CEMEC-PoliBA – Centre of Excellence in Computational Mechanics, V.le Japigia 182, Politecnico di Bari, 70125 Bari, Italy*
[c] *Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109-2125, USA*

Received 19 February 2007; received in revised form 13 June 2007; accepted 15 June 2007
Available online 21 June 2007

Abstract

Elastic systems with frictional interfaces subjected to periodic loading are sometimes predicted to 'shake down' in the sense that frictional slip ceases after the first few loading cycles. The similarities in behaviour between such systems and monolithic bodies with elastic–plastic constitutive behaviour have prompted various authors to speculate that Melan's theorem might apply to them – i.e., that the existence of a state of residual stress sufficient to prevent further slip is a sufficient condition for the system to shake down.

In this paper, we prove this result for 'complete' contact problems in the discrete formulation (i) for systems with no coupling between relative tangential displacements at the interface and the corresponding normal contact tractions and (ii) for certain two-dimensional problems in which the friction coefficient at each node is less than a certain critical value. We also present counter-examples for all systems that do not fall into these categories, thus giving a definitive statement of the conditions under which Melan's theorem can be used to predict whether such a system will shake down.

Keywords: Contact problems; Fretting fatigue; Shakedown; Melan's theorem; Coulomb friction

1. Introduction

Nominally static frictional contacts between engineering components are extremely prevalent, including for example bolted joints (Berczynski and Gutowski, 2006; Law et al., 2006), blade root contacts in jet engines (Murthy et al., 2004) and shrink fit assemblies (Booker et al., 2004). Such contacts are often subjected to a substantial mean load and a superposed oscillatory load due to mechanical vibrations. Under these conditions, failure can occur due to fretting fatigue, associated with periodic reversed microslip in parts of the contact region (Lovrich and Neu, 2006; Nowell et al., 2006). However, when the contact is 'complete' – i.e., when the contact area is known *a priori* and does not change during the loading cycle – analysis of the resulting contact problem sometimes predicts that the system shakes down (Banerjee and Hills, 2006; Churchman

* Corresponding author.Tel.: +1 7349360406; fax: +1 7346156647.
E-mail address: jbarber@umich.edu (J.R. Barber).

0020-7683/$ - see front matter
doi:10.1016/j.ijsolstr.2007.06.013

and Hills, 2006). In other words, microslip occurs at the interface only during the first few cycles, after which 'stick' conditions apply throughout the contact area for all subsequent times.

There are clear parallels between frictional slip and plastic deformation in elastic–plastic solids. Both are dissipative mechanisms that depend on the history but not the rate of loading. Analogies can also be found between the phenomena of cyclic plasticity and ratchetting and corresponding phenomena in frictional contact. For example, a rigid flat punch loaded by a constant tangential force and a periodic normal force is found to 'walk' across an elastic half plane by a constant increment in each loading cycle, even though the global friction limit is not exceeded at any time (Mugadu et al., 2004). The gradual monotonic rotation of bushings in a conrod small end can probably be attributed to a similar mechanism (Antoni et al., 2007). These similarities have caused tribologists to speculate (e.g., Churchman et al., 2006) as to whether it might be possible to prove a frictional equivalent of the well-known theorem due to Melan regarding the conditions for shakedown of an elastic–plastic body subjected to oscillatory loading (Melan, 1936).

Such a theorem might 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."*

Drucker (1954) showed that although some simple one-degree-of-freedom frictional systems show analogies with elastic–plastic behaviour, the non-associative nature of the flow rule for friction prevents the limit theorems from being applied directly to the friction problem. Fredriksson and Rydholm (1981) argued that if the normal traction at the interface is constant and prescribed, the flow rule essentially becomes associative in the sense that its dependence of flow on pressure merely alters the effective yield criterion and the permitted slip within the plane does satisfy the condition that it aligns with the direction of maximum tangential traction. This requires both that there be no coupling between tangential displacements and normal tractions and that the periodic component of the applied load does not generate a corresponding variation in normal traction. These conditions are met in the simple example treated by Churchman et al. (2006). Note also that the theorems and counter-examples of the present paper concern systems with Coulomb friction in contrast to problems involving the rather artificial so-called Tresca friction. For such systems one easily proves shakedown theorems (Antoni et al., 2007), but no information is then provided as to when such theorems are also valid for Coulomb frictional systems.

Necessary and sufficient conditions for shakedown in non-associated plasticity, which in principle can be generalized to friction problems, where given by Maier (1969). However, Björkman and Klarbring (1987, 1988) pointed out that Maier's sufficient condition, using the concept of a 'reduced elastic domain', is not useful in case of friction since this domain then becomes just the half line of positive normal contact forces with zero tangential forces, and the calculated lower bound for the load will usually be zero. The necessary condition of Maier, on the other hand, is the obvious statement that shakedown will never occur unless there exists a residual stress state that prevents further slip. However, Björkman and Klarbring made numerical calculations that showed that the upper bound for the shakedown load that one gets by using this necessary condition often gives considerably higher values than those obtained from direct quasi-static solutions.

In the present paper, we shall establish definitive rules for the conditions that must be met by a discrete (e.g., finite element) formulation of an elastic frictional contact problem in order that Melan's theorem should apply, subject only to the restriction that the contact be 'complete' – i.e., that the contact area should remain constant throughout the loading cycle.

2. The three-dimensional discrete system

We consider the general three-dimensional problem in which an elastic body is loaded by time-varying external forces and makes frictional contact with one or more rigid obstacles. The body is discretized by the finite element method in such a way that there are m nodes of which n ($n < m$) are contact nodes and $m - n$ are interior nodes. The deformation in the discrete solution is completely defined by the set of nodal displacement vectors $\boldsymbol{u}_j$, $j = 1, m$. Reaction forces $\boldsymbol{r}_i$, $i = 1, n$ will be generated at the contact nodes in response to the external nodal forces $\boldsymbol{F}_j(t)$. From these reaction forces, we form the vector

$$\boldsymbol{r} = [\boldsymbol{r}_1^{\mathrm{T}}, \ldots, \boldsymbol{r}_n^{\mathrm{T}}]^{\mathrm{T}} \in R^{3n}. \tag{1}$$

Also, from the displacement vectors $\boldsymbol{u}_i$, $i = 1, n$ associated with the contact nodes only, we form a global contact displacement vector

$$\boldsymbol{u} = [\boldsymbol{u}_1^{\mathrm{T}}, \ldots, \boldsymbol{u}_n^{\mathrm{T}}]^{\mathrm{T}} \in R^{3n}. \tag{2}$$

By a standard static condensation procedure, we can eliminate the displacements associated with the internal nodes and write the reaction forces in the form

$$\boldsymbol{r} = \boldsymbol{r}^w + \boldsymbol{\kappa}\boldsymbol{u}, \tag{3}$$

where $\boldsymbol{\kappa}$ is a contact stiffness matrix that is symmetric and positive semi-definite, but not necessarily non-singular (i.e., the structure can perform rigid-body displacements). The term $\boldsymbol{r}^w = \boldsymbol{r}^w(t)$ in (3) represents the time-varying reaction forces that would be generated at the contact nodes by the external forces if displacement at these nodes were constrained to be zero – i.e., if the contact nodes were welded to the obstacle(s). The remaining term $\boldsymbol{\kappa}\boldsymbol{u}$ represents the modification to these reactions resulting from the displacements of the contact nodes.

2.1. The Coulomb friction law

To state the mathematical definition of the Coulomb friction law, we first define a unit normal vector $\boldsymbol{n}_i$ for each contact node pointing from the obstacle into the body. Nodal contact displacements and reaction forces can then be decomposed into tangential and normal vectors

$$\boldsymbol{u}_i = \boldsymbol{v}_i + w_i\boldsymbol{n}_i; \quad \boldsymbol{v}_i \cdot \boldsymbol{n}_i = 0, \tag{4}$$

$$\boldsymbol{r}_i = \boldsymbol{q}_i + p_i\boldsymbol{n}_i; \quad \boldsymbol{q}_i \cdot \boldsymbol{n}_i = 0. \tag{5}$$

With this decomposition, vectors $\boldsymbol{v}_i$, $\boldsymbol{q}_i$ are restricted to the local interfacial plane. Henceforth, these will therefore be considered as vectors belonging to R^2, rather than R^3.

Clearly $p_i = \boldsymbol{r}_i \cdot \boldsymbol{n}_i$ and $w_i = \boldsymbol{u}_i \cdot \boldsymbol{n}_i$ (no sum). In this paper, we shall restrict attention to 'complete' contact problems in which the contact area is known *a priori* and no separation occurs as a result of the external forces. We therefore require that

$$w_i = 0; \quad p_i \geqslant 0. \tag{6}$$

The Coulomb friction law for node i can then be stated as

$$|\boldsymbol{q}_i| \leqslant f_i p_i, \tag{7}$$

$$|\boldsymbol{q}_i| < f_i p_i \Rightarrow \dot{\boldsymbol{v}}_i = \boldsymbol{0}, \tag{8}$$

$$0 < |\boldsymbol{q}_i| = f_i p_i \Rightarrow \dot{\boldsymbol{v}}_i = -\lambda_i \boldsymbol{q}_i, \quad \lambda_i \geqslant 0, \tag{9}$$

where f_i is the coefficient of friction and a superposed dot denotes the time derivative. Condition (7) states that the reaction forces should belong to an admissible set known as the *Coulomb friction cone* and (8) states that as long as $\boldsymbol{r}_i$ is in the interior of this cone, there can be no slip. Condition (9) is the frictional analogue of the flow rule in plasticity and states that during slip, the friction force $\boldsymbol{q}_i$ must oppose the direction of slip $\dot{\boldsymbol{v}}_i$. Notice that the inequality in (6) is implied by (7).

The quasi-static frictional contact problem consists of finding the time evolution of $\boldsymbol{u} = \boldsymbol{u}(t)$ and $\boldsymbol{r} = \boldsymbol{r}(t)$ for a given history of external loading, such that (3) and (6)–(9) are satisfied at all times. Notice that the effect of the external nodal forces $\boldsymbol{F}_j(t)$ on the problem appears only through the function $\boldsymbol{r}^w(t)$. The problem is meaningful only if initial conditions at time $t = 0$ are such that the contact forces $\boldsymbol{r}(0)$ belong to the friction cone – i.e.,

$$p_i(0) \geqslant 0; \quad |\boldsymbol{q}_i(0)| \leqslant f_i p_i(0). \tag{10}$$

It is assumed in this paper that conditions guaranteeing existence of solution of the quasi-static contact problem are satisfied (Andersson, 1999; Andersson and Klarbring, 2000). Moreover, the conditions of our two theorems are such that they do, in fact, also guarantee uniqueness of solutions.

2.2. Coupling coefficients

We denote by $\mathcal{D}$ the set of displacement vectors $\boldsymbol{u}$ such that $w_i = 0$, $i = 1, n$. For each $\boldsymbol{u} \in \mathcal{D}$, we calculate the corresponding contact reactions through the equation $\boldsymbol{r} = \boldsymbol{\kappa u}$, thereby defining a set of n functions

$$\mathcal{D} \ni \boldsymbol{u} \mapsto \boldsymbol{r}_i(\boldsymbol{u}) \in R^3. \tag{11}$$

We can then define a set of n real-valued functions $g_i(\boldsymbol{u})$ through

$$g_i(\boldsymbol{u}) = \begin{cases} \frac{\boldsymbol{r}_i(\boldsymbol{u})}{|\boldsymbol{r}_i(\boldsymbol{u})|} \cdot \boldsymbol{n}_i & \text{if } \boldsymbol{r}_i(\boldsymbol{u}) \neq \boldsymbol{0}, \\ 0 & \text{if } \boldsymbol{r}_i(\boldsymbol{u}) = \boldsymbol{0}. \end{cases} \tag{12}$$

Finally, we can define a *coupling coefficient* c_i for each contact node as

$$c_i = \max_{\boldsymbol{u} \in \mathcal{D}} g_i(\boldsymbol{u}). \tag{13}$$

Clearly $c_i \leqslant 1$ and since $-\boldsymbol{u} \in \mathcal{D}$ whenever $\boldsymbol{u} \in \mathcal{D}$, the linearity of $\boldsymbol{r} = \boldsymbol{\kappa u}$ implies also that $c_i \geqslant 0$. The limiting case where $c_i = 0$ for all contact nodes arises when tangential displacements at the contact area (slip) have no effect on the normal component of the contact reactions – i.e., there is no normal–tangential coupling in the contact problem. This occurs for example in continuum problems for the contact of two elastic half spaces if Dundurs' constant $\beta = 0$ (Barber, 2002).

We shall find that the coupling constants c_i play a major role in determining when a frictional Melan's theorem can be applied in contact problems. Notice that they are related to, but distinct from the constants introduced by Andersson (1999) in his investigation of existence and uniqueness of the discrete frictional contact problem.

2.2.1. The one-node two-dimensional system

The simplest case is that of a two-dimensional system with only one contact node. This system was studied extensively by Klarbring (1990) and Cho and Barber (1998) with particular reference to the effect of coefficient of friction on uniqueness and stability of solution. The displacement $\boldsymbol{u}$ then has only two degrees of freedom u_1, u_2 corresponding to tangential and normal displacement at the contact node, respectively, and $\boldsymbol{\kappa}$ is a 2×2 matrix. If the node remains in contact, $u_2 = 0$. It follows from (12, 13) that the sole coupling coefficient c_1 is given by

$$c_1 = \frac{|\kappa_{21}|}{\sqrt{\kappa_{11}^2 + \kappa_{21}^2}}. \tag{14}$$

Notice that since $\boldsymbol{\kappa}$ is positive semi-definite, it follows that for a single node system, $0 \leqslant c_1 < 1$.

2.2.2. Multi-node systems

By contrast, for almost all systems with normal–tangential coupling and more than one node, we shall find that all the coupling coefficients are unity. In this section, we shall investigate what values of c_i are to be expected for the three-dimensional case, but the changes needed for covering the plane case are obvious.

Introducing the vectors

$$\boldsymbol{v} = [\boldsymbol{v}_1^{\mathrm{T}}, \ldots, \boldsymbol{v}_n^{\mathrm{T}}]^{\mathrm{T}} \in R^{2n}, \quad \boldsymbol{w} = [w_1, \ldots, w_n] \in R^n,$$
$$\boldsymbol{q} = [\boldsymbol{q}_1^{\mathrm{T}}, \ldots, \boldsymbol{q}_n^{\mathrm{T}}]^{\mathrm{T}} \in R^{2n}, \quad \boldsymbol{p} = [p_1, \ldots, p_n] \in R^n,$$

we may decompose $\boldsymbol{\kappa}$ and write $\boldsymbol{r} = \boldsymbol{\kappa u}$ as

$$\boldsymbol{q} = \boldsymbol{A v} + \boldsymbol{B}^{\mathrm{T}} \boldsymbol{w}, \quad \boldsymbol{p} = \boldsymbol{B v} + \boldsymbol{C w}, \tag{15}$$

where $\boldsymbol{A}$ and $\boldsymbol{C}$ are symmetric and possibly non-singular matrices.

We restrict attention to sets of displacements $\boldsymbol{u}$ for which all the nodes remain in contact and hence $\boldsymbol{w} = \boldsymbol{0}$. A function $\boldsymbol{r}_i(\boldsymbol{u})$ satisfying this condition, as in (11), can be represented by two rows from $\boldsymbol{A}$, forming a $2 \times 2n$

matrix $\boldsymbol{A}_i$, and one row from $\boldsymbol{B}$, forming a $1 \times 2n$ matrix $\boldsymbol{B}_i$. We may stack these submatrices on top of each other to form a $3 \times 2n$ matrix $\boldsymbol{\kappa}_i$ such that

$$\boldsymbol{q}_i = \boldsymbol{A}_i \boldsymbol{v},\ p_i = \boldsymbol{B}_i \boldsymbol{v} \Longleftrightarrow \boldsymbol{r}_i = \boldsymbol{\kappa}_i \boldsymbol{v}. \tag{16}$$

If $\boldsymbol{\kappa}_i$ has the full rank three (the two rows of $\boldsymbol{A}_i$ and the one row of $\boldsymbol{B}_i$ are linearly independent), then any $\boldsymbol{r}_i$ is a result of some $\boldsymbol{v}$ and, in particular, we can always find a vector $\boldsymbol{v}$ such that $\boldsymbol{r}_i$ has the same direction as $\boldsymbol{n}_i$, resulting in $c_i = 1$. Thus, $c_i \neq 1$ requires $\boldsymbol{\kappa}_i$ to be of rank less than three. When it is of rank two or one, the range space $\mathcal{R}(\boldsymbol{\kappa}_i)$ of $\boldsymbol{\kappa}_i$ can be represented by a surface or a line, respectively, in R^3.

Fig. 1 illustrates the case where $\boldsymbol{\kappa}_i$ is of rank two and $\mathcal{R}(\boldsymbol{\kappa}_i)$ is a plane which contains the direction of all possible vectors $\boldsymbol{r}_i$ that can be obtained by substituting a tangential displacement vector $\boldsymbol{v}$ into Eq. (16). The definition of c_i (12, 13) then shows that

$$c_i = \cos \alpha_i, \tag{17}$$

where α_i is the smallest angle between a line in the plane $\mathcal{R}(\boldsymbol{\kappa}_i)$ and $\boldsymbol{n}_i$. The special case $\alpha_i = 90°$ (and hence $c_i = 0$) results when $\boldsymbol{B}_i = \boldsymbol{0}$, which is the case where there is no normal–tangential coupling. Alternatively, if $\mathcal{R}(\boldsymbol{\kappa}_i)$ has rank one, corresponding to a line, then all possible vectors $\boldsymbol{r}_i$ are parallel with this line and the definition of c_i then shows that $c_i = \cos\alpha_i$, where α_i is the angle between $\mathcal{R}(\boldsymbol{\kappa}_i)$ and $\boldsymbol{n}_i$.

A classical result of linear algebra is that the null space of $\boldsymbol{\kappa}_i^{\mathrm{T}}$, denoted $\mathcal{N}(\boldsymbol{\kappa}_i^{\mathrm{T}})$, is the orthogonal complement of $\mathcal{R}(\boldsymbol{\kappa}_i)$. For the case illustrated in Fig. 1, $\mathcal{N}(\boldsymbol{\kappa}_i^{\mathrm{T}})$ is the normal to the plane $\mathcal{R}(\boldsymbol{\kappa}_i)$ as shown. The symmetry of $\boldsymbol{A}$ then implies that a nodal displacement vector $\boldsymbol{u}_i^*$ that belongs to $\mathcal{N}(\boldsymbol{\kappa}_i^{\mathrm{T}})$ is such that, if all other nodal displacements are set to zero, the resulting full vector of tangential forces $\boldsymbol{q}$ is zero. In physical terms, it follows that we shall obtain a value of $c_i < 1$ only if there exist one or more directions for the nodal displacement $\boldsymbol{u}_i$ such that no tangential reactions are generated at any node.

To illustrate this, consider the two-dimensional two-node system of Fig. 2 in which tangential displacement $\boldsymbol{v}_1$ of node 1 results in forces from the 'self-influence' spring K_A and from the 'coupling' spring K_B that have the same direction. It then follows that (i) any tangential motion of node 1 will generate a force aligned with these springs and equivalently (in view of the symmetry of the stiffness matrix $\boldsymbol{\kappa}$) (ii) displacement of node 1 in direction $\boldsymbol{e}$ perpendicular to the line of the springs generates no tangential component of reaction. Notice that the springs at each node normal to the interface contribute only to matrix $\boldsymbol{C}$ and hence have no effect on c_i. It then follows that $c_1 = \cos\alpha_1$, which is of course a special case of (17).

However, we should emphasise that for $n > 1$, systems of this kind have to be very carefully tailored to give values of c_i that are not either zero or unity and are unlikely to result from, e.g., finite element discretization of realistic elastic contact problems. We explore these special cases here, simply so as to ensure that the following theorems have completely general applicability.

3. Shakedown theorems

A system is said to have reached a state of shakedown at time t_0 if for all future times $t > t_0$, $\dot{\boldsymbol{u}} = \boldsymbol{0}$ – i.e., no further frictional slip occurs and all nodes remain in a state of stick. An obvious necessary condition for this to occur is that there must exist a time-independent vector $\tilde{\boldsymbol{u}}$ such that the normal component $\tilde{\boldsymbol{w}} = \boldsymbol{0}$ and

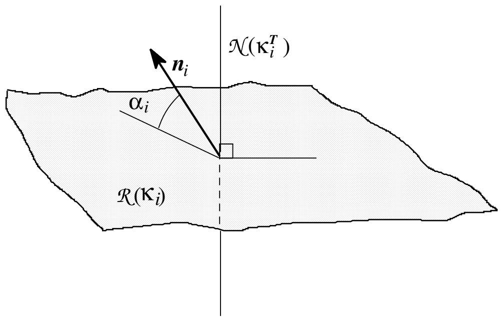

Fig. 1. The range space $\mathcal{R}(\boldsymbol{\kappa}_i)$ for the case where $\boldsymbol{\kappa}_i$ is of rank two.

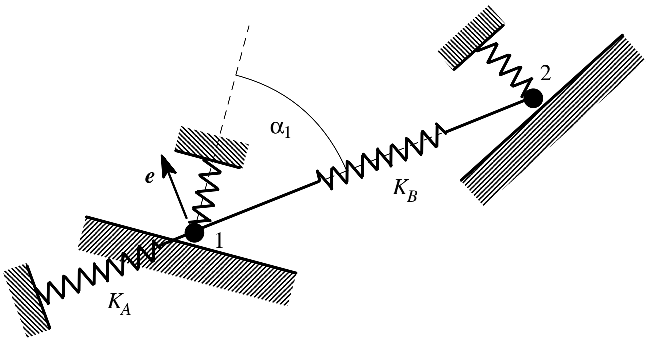

Fig. 2. A two-dimensional two-node system for which $0 < c_1, c_2 < 1$.

$$|\tilde{\boldsymbol{q}}_i| \leqslant f_i \tilde{p}_i, \quad i = 1, n, \tag{18}$$

where

$$\tilde{\boldsymbol{r}} = \boldsymbol{r}^w + \boldsymbol{\kappa}\tilde{\boldsymbol{u}}. \tag{19}$$

We shall call the $\tilde{\boldsymbol{u}}$ a *shakedown displacement vector*. We shall also refer to $\tilde{\boldsymbol{u}}$ as a *safe* shakedown displacement vector if the strict inequality is enforced in (18).

The fundamental question addressed in the present paper is to determine if and when the existence of such a vector provides also a *sufficient* condition for the system to shake down. For this purpose, we first define a modified definition of shakedown by introducing the norm

$$A = \frac{1}{2}(\tilde{\boldsymbol{u}} - \boldsymbol{u})^{\mathrm{T}}\boldsymbol{\kappa}(\tilde{\boldsymbol{u}} - \boldsymbol{u}) \geqslant 0,$$

which is a measure of the difference between the instantaneous displacement vector $\boldsymbol{u}$ and the shakedown displacement vector $\tilde{\boldsymbol{u}}$. We shall consider the system to shakedown if $\dot{A} < 0$ whenever $\dot{\boldsymbol{u}} \neq \boldsymbol{0}$ – in other words, any slip that occurs causes the system to approach $\tilde{\boldsymbol{u}}$ in the sense of the norm A. Clearly $\dot{A} = 0$ when $\dot{\boldsymbol{u}} = \boldsymbol{0}$ and hence A is then a non-increasing function of time. This condition recognizes that the final shakedown state of the system might differ from $\tilde{\boldsymbol{u}}$, giving $A > 0$ and $\dot{A} = 0$. It also places no restriction on the number of load cycles or the time required to achieve shakedown, so the possibility of monotonic asymptotic approach to a non-zero value of A is also open.

Theorem 1. *Assume there exists a safe time-independent shakedown displacement vector $\tilde{\boldsymbol{u}}$ such that $\tilde{\boldsymbol{w}} = \boldsymbol{0}$ and $|\tilde{\boldsymbol{q}}_i| < f_i\tilde{p}_i$, $i = 1,n$, where $\tilde{\boldsymbol{r}} = \boldsymbol{r}^w + \boldsymbol{\kappa}\tilde{\boldsymbol{u}}$. If $c_i = 0$ for all $i = 1,n$, the actual displacement $\boldsymbol{u}$ will approach $\tilde{\boldsymbol{u}}$ in the sense that $\dot{A} < 0$ whenever $\dot{\boldsymbol{u}} \neq \boldsymbol{0}$.*

Proof. The time derivative of the norm A is

$$\dot{A} = -(\tilde{\boldsymbol{u}} - \boldsymbol{u})^{\mathrm{T}}\boldsymbol{\kappa}\dot{\boldsymbol{u}} = -\sum_{i=1}^{n}(\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i)\cdot\dot{\boldsymbol{u}}_i.$$

Since $\tilde{\boldsymbol{u}} - \boldsymbol{u}$ belongs to the set $\mathcal{D}$, the condition $c_i = 0$ and the definition of c_i imply that $(\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i)\cdot\dot{\boldsymbol{n}}_i = 0$. This means that both $\tilde{\boldsymbol{r}}_i$ and $\boldsymbol{r}_i$ belong to the circle of radius $f_ip_i = f_i\tilde{p}_i$ formed by cutting the friction cone at the common normal force level $p_i = \tilde{p}_i$ and in view of the strict inequality for a safe shakedown state, $\tilde{\boldsymbol{r}}_i$ must lie strictly inside this circle.

If $\dot{\boldsymbol{u}}_i \neq \boldsymbol{0}$, the Coulomb friction law (7)–(9) implies that (i) $\boldsymbol{r}_i$ lies on the boundary of the circle and (ii) that the direction of $\dot{\boldsymbol{u}}_i$ coincides with the inward normal to the circle at the point $\boldsymbol{r}_i$, as shown in Fig. 3. It follows that $(\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i)\cdot\dot{\boldsymbol{u}}_i$ positive for all i. □

We conclude that Melan's theorem applies for any discrete elastic system involving complete contact and in which there is no coupling between tangential displacements and normal tractions.

Theorem 2. *Consider a coupled two-dimensional discrete elastic system and assume there exists a safe time-independent shakedown displacement vector* $\tilde{\boldsymbol{u}}$ *such that* $\tilde{\boldsymbol{w}} = \boldsymbol{0}$ *and* $\tilde{q}_i < f_i\tilde{p}_i$, $i = 1, n$, *where* $\tilde{\boldsymbol{r}} = \boldsymbol{r}^w + \boldsymbol{\kappa}\tilde{\boldsymbol{u}}$. *If*

$$f_i \leqslant \frac{\sqrt{1-c_i^2}}{c_i} \tag{20}$$

for all $i = 1, n$, *the actual displacement* $\boldsymbol{u}$ *will approach* $\tilde{\boldsymbol{u}}$ *in the sense that* $\dot{A} < 0$ *whenever* $\dot{\boldsymbol{u}} \neq \boldsymbol{0}$.

An alternative statement of the inequality (20) is

$$f_i \leqslant \tan \alpha_i, \tag{21}$$

where α_i *is defined in Fig. 1 and Eq. (17).*

Proof. Since f_i, c_i are both positive, (20) implies that

$$c_i \leqslant \frac{1}{\sqrt{1+f_i^2}}. \tag{22}$$

Since $\tilde{\boldsymbol{u}} - \boldsymbol{u}$ belongs to the set $\mathcal{D}$, we also have

$$c_i \geqslant \frac{\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i}{|\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i|} \cdot \boldsymbol{n}_i \tag{23}$$

from (12, 13) and hence

$$\frac{\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i}{|\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i|} \cdot \boldsymbol{n}_i \leqslant \frac{1}{\sqrt{1+f_i^2}}. \tag{24}$$

In Fig. 4, BAC defines the stick sector for the instantaneous value of p_i. During slip $\dot{\boldsymbol{u}} \neq \boldsymbol{0}$, $\boldsymbol{r}_i$ must lie on one of the lines AB, AC and in the figure is represented by the point D. The frictional flow rule (9) then demands that the slip vector $\dot{\boldsymbol{u}}$ be directed into the stick sector as shown.

In view of the strict inequality for a safe shakedown state, the shakedown vector $\tilde{\boldsymbol{r}}_i$ must be strictly inside the sector BAC, but the inequality (24) also excludes it from the shaded sector BDE. We conclude that $\tilde{\boldsymbol{r}}_i$ must lie in the region $EDAC$, from which it is clear that $(\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i) \cdot \dot{\boldsymbol{u}}_i$ is positive for all i. □

Notice that in the special case where there is only one node ($n = 1$), (20, 14) gives

$$f_i \leqslant \left|\frac{\kappa_{11}}{\kappa_{21}}\right|, \tag{25}$$

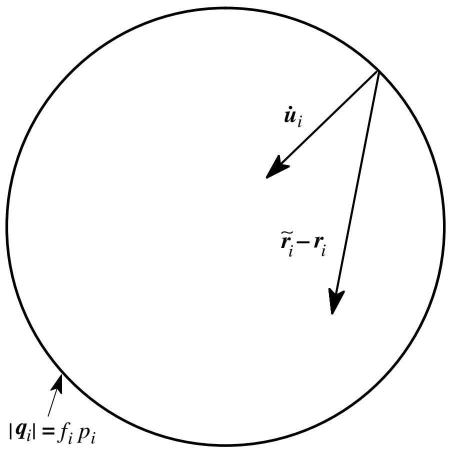

Fig. 3. The vector $(\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i)$ must be directed into the circle from the point $\boldsymbol{r}_i$ on the boundary.

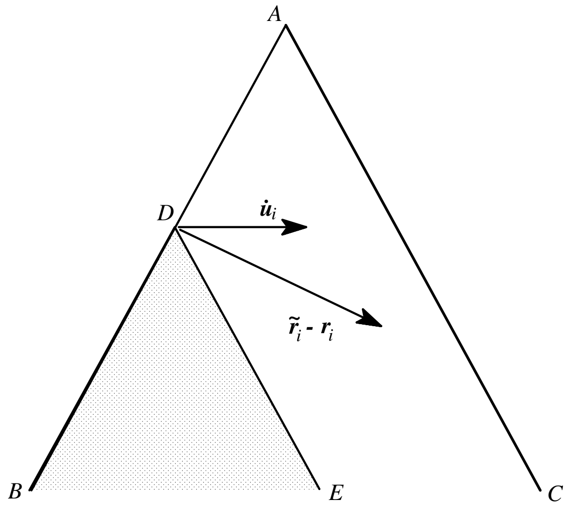

Fig. 4. The inequality (24) excludes $\tilde{\boldsymbol{r}}_i$ from the shaded region BDE.

which also defines the range of friction coefficients in which this system has a unique quasi-static solution (Klarbring, 1990; Cho and Barber, 1998).

3.1. Three-dimensional coupled systems

The reader might reasonably ask why the proof of Theorem 2 cannot be extended to three-dimensional coupled systems. For this case, the sectors in Fig. 3 are replaced by friction cones. We shall denote the stick cone replacing BAC by $\mathcal{S}$ and the conical region excluded by the inequality (24) by $\mathcal{R}$. The vector $\dot{\boldsymbol{u}}_i$ is directed into the circle obtained by intersecting $\mathcal{S}$ with a horizontal plane passing through $\boldsymbol{r}_i$. The vector $\tilde{\boldsymbol{r}}_i$ is now restricted to the region $\mathcal{S} - \mathcal{R}$ that is inside $\mathcal{S}$ but outside $\mathcal{R}$. This region contains points (out of the plane in Fig. 4) that are to the left of a vertical line through D (the vertex of $\mathcal{R}$) and for which $(\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i) \cdot \dot{\boldsymbol{u}}_i$ is therefore negative. Thus Theorem 2 cannot be proved for three-dimensional coupled systems. We shall show in the next section that counter-examples to Melan's theorem can be established in this case.

4. Counter-examples

In Section 3, we proved two theorems defining conditions under which the frictional Melan's theorem applies to elastic systems. In this section, we shall demonstrate by counter-example that these are the *only* conditions under which the theorem holds. In other words, if a discrete elastic system does not fall under the terms of one of our two theorems, there will always exist some loading scenarios under which the system does not shake down, even though a safe shakedown vector can be identified. To establish this result, it is sufficient to identify at least one loading scenario for any given system for which the predictions of Melan's theorem can be demonstrated to be false.

In preparation for these counter-examples, we consider the general n-node system subjected to a set of slip displacements $\boldsymbol{v}_i$. The $n-1$ nodes $i \neq j$ are then subjected to time-independent external normal forces that are sufficiently large to prevent further slip at these nodes during the subsequent loading scenario, so that the system is effectively reduced to a one-node system (node j).

4.1. Cases where $c_j = 1$

Consider the case where $\boldsymbol{v}_i$ is chosen to be the set that maximizes c_j as in Eq. (13). We denote this set of slip displacements by $\boldsymbol{v}_i^*$. We remarked in Section 2.2.2 that for almost all multi-node systems, $c_j = 1$, in which case with this choice of $\boldsymbol{v}_i$, the reaction at node j will be a purely normal force p_j. By imposing the slip displacements $\lambda \boldsymbol{v}_i^*$, where λ is a scalar multiplier of appropriate sign, we can generate a normal contact reaction $p_j > 0$ of any given value.

Now suppose that we choose $\lambda = 0$, so that the initial value of $p_j = 0$, and we impose a periodic force at node j only, such that some periodic slip occurs at node j (recall that the other nodes are being prevented from slipping by sufficiently large time-independent external normal forces). In other words, the system does not shake down. However, with the same periodic loading, a safe shakedown state can always be found by choosing a sufficiently large value of λ and hence of p_j. Thus, Melan's theorem fails for this case and hence for any system in which $c_j = 1$ at at least one node.

4.2. Cases where $0 < c_j < 1$

If $0 < c_j < 1$, we can still impose the tangential displacements $\lambda \boldsymbol{v}_i^*$, but the resulting reaction at node j will be inclined to the normal at the angle α_j defined in Fig. 1 and Eq. (17). Suppose at this stage we were to relax the tangential component $\boldsymbol{q}_j$ of this reaction, whilst keeping the remaining nodes $i \neq j$ fixed. This could be done, for example, by the thought example of temporarily making the contact at node j frictionless. This relaxation process must also cause the normal reaction p_j to go to zero, since if this were not the case, we would have defined a new slip state $\boldsymbol{v}_i$ at which $c_j = 1$, *contra hyp*. It also follows that a reaction in the optimal direction α_j can be generated by displacement (in an appropriate direction) of node j only (since it is unloading in this direction that constituted the relaxation). In other words, a possible choice for $\boldsymbol{v}_i^*$ comprises slip in this direction at node j only, with $\boldsymbol{v}_i = \boldsymbol{0}$, $i \neq j$.

It follows that if the nodes $i \neq j$ are locked as above by sufficiently large external normal forces and if $f_j > \tan \alpha_j$, it will be possible to 'wedge' node j (Barber and Hild, 2006). In other words, an appropriate tangential displacement at node j will generate a sufficient normal and hence frictional reaction to prevent it from relaxing back to the original position when the external force at node j is removed. Furthermore, the wedged displacement can be chosen to be arbitrarily large and hence sufficient to define a safe shakedown state for any conceivable periodic loading cycle at node j, thus providing a counter-example as in Section 4.1. It follows that, Melan's theorem fails for any system in which $f_j > \tan \alpha_j$ at at least one node. Notice that this counter-example also applies to systems with only one contact node.

4.3. Three-dimensional systems with $0 < c_j < 1$ and $f_j < \tan \alpha_j$

The preceding counter-examples cover all cases not covered by our two theorems *except* three-dimensional systems with $0 < c_j < 1$ and $f_j < \tan \alpha_j$. As in Section 4.2, these systems can be reduced to an equivalent one-node system by locking the nodes $i \neq j$ with sufficiently large external normal forces. It is therefore sufficient to demonstrate a counter-example for the single node three-dimensional system with $f_j < \tan \alpha_j$.

For the one-node system, coupling between normal and tangential displacements is defined by a 3×3 stiffness matrix $\boldsymbol{\kappa}$. Suppose that the loading vector is denoted by $\boldsymbol{r}^w = \{F_1, F_2, F_3\}^{\mathrm{T}}$, with F_2 being the component in the direction normal to the contact interface. The contact reaction is then given by Eq. (3), with $\boldsymbol{u} = \{u_1, u_2, u_3\}^{\mathrm{T}}$. If the node separates from the obstacle, there will be no contact reaction ($\boldsymbol{r} = 0$) and hence

$$\boldsymbol{\kappa u} = -\boldsymbol{r}^w \tag{26}$$

from (3). The gap between the node and the obstacle is u_2 and this is zero if

$$\boldsymbol{n} \cdot \boldsymbol{\kappa}^{-1} \boldsymbol{r}^w = 0, \tag{27}$$

which defines a plane in (F_1, F_2, F_3) space. Separation is possible only if $u_2 > 0$, which defines the region on one side of this plane. Cho and Barber (1999) analyzed this system and showed that the apex of the friction cone in (F_1, F_2, F_3) space must lie on the separation plane (27), but its position depends on the slip displacement $\boldsymbol{v} = \{u_1, u_3\}^{\mathrm{T}}$. They also identified a critical friction coefficient f_{cr} such that for $f > f_{\mathrm{cr}}$ the separation plane and the friction cone intersect, defining a range of loads $\boldsymbol{r}^w$ in which both stick and separation are possible. In the terminology of the present paper, $f_{\mathrm{cr}} = \tan \alpha$.

For $f < f_{\mathrm{cr}}$, the friction cone and the separation plane intersect only at the apex, the quasi-static solution is unique, and wedging is not possible. The direction of the in-plane coordinates x_1, x_3 can always be chosen so

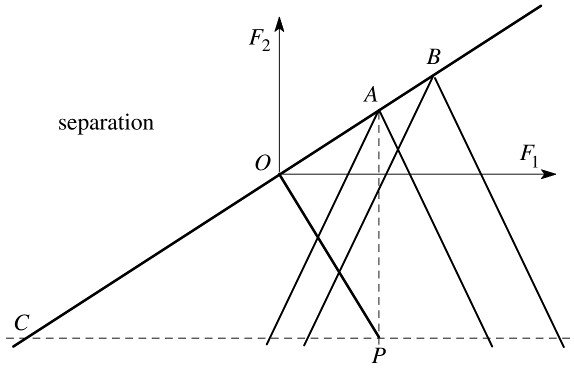

Fig. 5. The $F_1F_2F_3$ diagram for the three-dimensional one-node system.

as to ensure that there is no coupling between u_3 and the normal reaction p and a cross-section of the resulting diagram is shown in Fig. 5, with two possible locations for the apex of the friction cone being the points A, B.

Slip causes the apex of the stick cone to move about the separation plane. Suppose we load along the line OP, after which F_1, F_2 are held constant, whilst F_3 oscillates in time. If the amplitude of the periodic force is sufficiently large to cause slip, the resulting displacement will tend to move the stick cone to the 'optimum' point where the apex of the cone is at A, vertically above P. If the amplitude is larger than the diameter of the circle obtained by intersecting this cone with the horizontal plane CP, periodic slip will occur, with the cone moving back and forth in the F_3-direction.

Fig. 6 shows the cross-section of the diagram at the value of F_2 corresponding to P. It is clear that the amplitude of oscillation DE can be chosen such that it is larger than the diameter of the circle corresponding to the apex being at A in Fig. 5, but that it is still contained within the circle corresponding to a different location of the apex B. However, the direction of slip dictated by the flow rule does not permit the stick cone to be moved in this direction. We therefore conclude that for this case there exists a safe shakedown state which however cannot be reached by certain loading scenarios. Thus, the putative shakedown theorem is disproved for the one-node thee-dimensional system. Since the multi-node system can be reduced to an equivalent one-node system as explained above, it is also disproved for all multi-node three-dimensional coupled systems.

5. Conclusions

We conclude that Melan's theorem, in the sense defined in Section 3, applies to the discrete complete frictional contact problem:

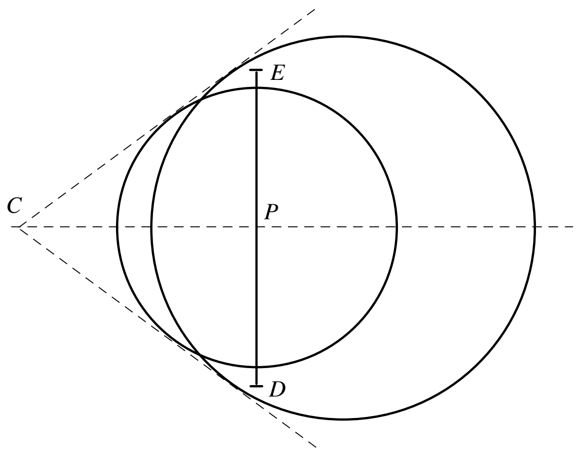

Fig. 6. Cross-section through the plane CP in Fig. 5.

(1) For any two or three-dimensional system in which there is no coupling between the tangential displacement vector $\boldsymbol{v}$ and the normal reaction vector p – i.e., the matrix $\boldsymbol{B}$ of Eq. (15) is null.
(2) For a two-dimensional coupled system for which the nodal coefficient of friction f_i satisfies the inequality (20) at each node. In this context, we note that the critical nodal coefficient for a multi-node system differs from zero only if there exists a direction of displacement $\boldsymbol{u}_i$ such that if all the other nodal displacements are set to zero, all the reactions $\boldsymbol{r}$ have zero tangential component $\boldsymbol{q} = \boldsymbol{0}$.

Counter-examples to the theorem can be established for all systems not falling into one of these categories.

Acknowledgements

The authors thank Dr. Luciano Afferante for his contribution to the discussions leading to the present paper. JRB and MC thank the National Science Foundation and Sandia National Laboratory for the organization and financial support of a workshop on the subject of frictional contacts subject to periodic loading.

References

Andersson, L.-E., 1999. Quasistatic frictional contact problems with finitely many degrees of freedom. LiTH-MAT-R-1999-22.
Andersson, L.-E., Klarbring, A., 2000. Quasi-static frictional contact of discrete mechanical structures. Eur. J. Mech. A/Solids 19, S61–S78.
Antoni, N., Nguyen, Q.-S., Ligier, J.-L., Saffré, P., Pastor, J., 2007. On the cumulative microslip phenomenon. Eur. J. Mech. A/Solids 26, 611–625.
Banerjee, N., Hills, D.A., 2006. Analysis of stick-slip and contact-edge behaviour in a simplified fretting fatigue test. J. Strain Analysis 41, 183–192.
Barber, J.R., 2002. Elasticity, second ed. Kluwer, Dordrecht, 410 pp.
Barber, J.R., Hild, P., 2006. On wedged configurations with Coulomb friction. In: Wriggers, Peter, Nackenhorst, Udo (Eds.), Analysis and Simulation of Contact Problems. Springer-Verlag, Berlin, pp. 205–213.
Berczynski, S., Gutowski, P., 2006. Identification of the dynamic models of machine tool supporting systems. Part I: an algorithm of the method. J. Vib. Control 12, 257–277.
Björkman, G., Klarbring, A., 1987. Shakedown and residual stresses in frictional systems. In: Gladwell, G.M.L., Ghonem, H., Kalousek, J. (Eds.), Contact Mechanics and Wear of Rail/Wheel Systems II: Proceedings of the 2nd International Symposium, University of Waterloo Press, pp. 27–39.
Booker, J.D., Truman, C.E., Wittig, S., Mohammed, Z., 2004. A comparison of shrink-fit holding torque using probabilistic, micromechanical and experimental approaches. Proc. Inst. Mech. Eng. B – J. Eng. Manufacture 218, 175–187.
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 Analysis 41, 323–328.
Drucker, D.C., 1954. Coulomb friction, plasticity, and limit loads. ASME J. Appl. Mech. 21, 71–74.
Fredriksson, B., Rydholm, G., 1981. Shakedown in frictional systems – some considerations. In: Rydholm, G. (Ed.), On Inequalities and Shakedown in Contact Problems, Diss. No. 61, Linköping Studies in Science and Technology, pp. F1–F21.
Cho, Hanbum, Barber, J.R., 1998. Dynamic behavior and stability of simple frictional systems. Mathematical and Computer Modeling 28, 37–53.
Cho, Hanbum, Barber, J.R., 1999. Stability of the three-dimensional Coulomb friction law. Proc. Roy. Soc. (Lond.) 455 A, 839–861.
Klarbring, A., 1990. Examples of non-uniqueness and non-existence of solutions to quasi-static contact problems with friction. Ingenieur-Archiv 60, 529–541.
Klarbring, A., Björkman, G., 1988. The treatment of problems in contact mechanics by mathematical programming. Journal de Mécanique Théorique et Appliquée 7 (suppl. 1), 83–96 (special issue).
Law, S.S., Wu, Z.M., Chan, S.L., 2006. Analytical model of a slotted bolted connection element and its behaviour under dynamic load. J. Sound Vib. 292, 777–787.
Lovrich, N.R., Neu, R.W., 2006. Effect of mean stress on fretting fatigue of Ti-6Al-4V on Ti-6Al-4V. Fatigue Fract. Eng. Matls. Struct. 29, 41–55.
Maier, G., 1969. Shakedown theory in perfect elastoplasticity with associated and nonassociated flow-laws: a finite element, linear programming approach. Meccanica 4, 250–260.
Melan, E., 1936. Theorie statisch unbestimmter Systeme aus ideal-plastichem Baustoff. Sitzungsber. d. Akad. d. Wiss., Wien 2A (145), 195–218.
Mugadu, A., Sackfield, A., Hills, D.A., 2004. Analysis of a rocking and walking punch –Part I: initial transient and steady state. ASME J. Appl. Mech. 71, 225–233.
Murthy, H., Harish, G., Farris, T.N., 2004. Efficient modeling of fretting of blade/disk contacts including load history effects. ASME J. Tribol. 126, 56–64.
Nowell, D., Dini, D., Hills, D.A., 2006. Recent developments in the understanding of fretting fatigue. Eng. Fract. Mech. 73, 207–222.