Chapter 1
Discrete Coulomb Frictional Systems Subjected to Periodic Loading

Jim R. Barber and Young Ju Ahn

Abstract. If elastic systems with frictional interfaces are subjected to periodic loading, the system may *shake down*, meaning that frictional slip is restricted to the first few cycles, or it may settle into a steady periodic state involving cyclic slip. Furthermore, if the system posesses a rigid-body mode, the slip may also cause an increment of rigid-body motion to occur during each cycle — a phenomenon known as *ratcheting*.

Here we investigate this behaviour for discrete systems such as finite element models, for which the contact state can be described in terms of a finite set of nodal displacements and forces. If the system is '*uncoupled* — i.e. if the stiffness matrix is such that the tangential nodal displacements are uninfluenced by the normal nodal forces, a frictional Melan's theorem can be proved showing that shakedown will occur for all initial conditions if there exists a safe shakedown state for the periodic loading in question. For coupled systems, we develop an algorithm for determining the range of periodic load amplitudes within which the long-time state might be cyclic slip or shakedown, depending on the initial condition. The problem is investigated using a geometric representation of the motion of the frictional inequality constraints in slip displacement space. Similar techniques are used to explore ratcheting behaviour in a low-order system.

1.1 Introduction

Many engineering systems comprise one or more contacting elastic bodies in nominally static contact. Examples include bolted joints between machine components and the centrifugally loaded contact between aero engine

Jim R. Barber · Young Ju Ahn
Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI48109-2125, USA
e-mail: {jbarber,yjahn}@umich.edu

G.E. Stavroulakis (Ed.): Recent Advances in Contact Mechanics, LNACM 56, pp. 1–11.
springerlink.com

turbine blades and the blade disk. These systems are typically subjected to mechanical vibrations, which can cause the contact tractions to exceed the limiting friction condition at part of the interface, leading to a state of cyclic microslip. This in turn results in energy dissipation which affects the dynamics of the system and may also lead to the initiation of fretting fatigue cracks emanating from the microslip region.

The Coulomb friction law is still arguably the best simple approximation to the observed behaviour of unlubricated contacts and it introduces a history dependence into the problem. In particular, the steady cyclic state will generally differ from that during the first cycle of loading. We would like to be able to solve for this steady state directly, and hence determine the location and magnitude of damage due to fretting fatigue and/or estimate the energy loss so as to define an equivalent (frequency-dependent) damping element. However, the steady cyclic state is often inherently non-unique, with the state achieved depending on the initial condition or the initial transient period of loading.

1.2 Shakedown and Melan's Theorem

If the time-independent component in the compressive normal tractions is sufficiently large, the system may *shake down*, meaning that the steady state is one in which all points on the interface remain in a state of stick after an initial transient that may involve microslip.

Shakedown is a well known phenomenon in the analogous process of elastic/plastic deformation, where it can be predicted using Melan's theorem [9] which broadly speaking states that if the system can shake down, it will do so regardless of initial conditions. For frictional systems, an equivalent theorem might be stated 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."* Transient studies of cyclic frictional systems seem to confirm the validity of this theorem [7], but the proof of Melan's theorem depends on the associativity of the plastic flow rule, whereas the Coulomb friction law is non-associative. The theorem has recently been proved in both discrete [8] and continuum [4] formulations, but only for the restricted class of systems in which there is no coupling between normal tractions and tangential displacements. This class includes the much studied case of the contact of two similar elastic half planes, and more generally, any system that is symmetric about the contact plane.

The discrete theorem is established by defining a non-negative norm

$$A = \frac{1}{2}\left(\tilde{\mathbf{v}} - \mathbf{v}\right)^T \boldsymbol{\kappa}\left(\tilde{\mathbf{v}} - \mathbf{v}\right) \ , \tag{1.1}$$

where $\mathbf{v}$ is a vector of instantaneous nodal slip displacements, $\tilde{\mathbf{v}}$ is a 'safe' shakedown vector and $\boldsymbol{\kappa}$ is the reduced stiffness matrix. The norm A is a measure of the deviation of the instantaneous deviation of the system from the shakedown state and the theorem is established by demonstrating that for all permissible slip motions, the time derivative $\dot{A} < 0$ and hence the shakedown state is approached monotonically.

1.3 Coupled Systems

That the normal and tangential elastic problems be uncoupled is both a necessary and sufficient condition for Melan's theorem to apply, except for certain very special low order discrete systems [8]. For coupled systems, it is always possible to construct counter examples to the theorem — i.e. periodic loading scenarios for which the long term state of the system may be either shakedown or cyclic slip depending only on the initial conditions.

To explore this phenomenon, we consider the behaviour of a two-dimensional N-node discrete system subjected to external loading of the form

$$\mathbf{F}(t) = \mathbf{F}_0 + \lambda\mathbf{F}_1(t) \ , \tag{1.2}$$

where $\mathbf{F}_0$ is a time-invariant mean load, $\mathbf{F}_1(t)$ is a periodic load, t is time and λ is a scalar loading factor.

The discrete description of the elastic system can be condensed so as to include only the contact degrees of freedom, giving a system of linear equations

$$\begin{aligned} q_j &= q_j^w + A_{ji}v_i + B_{ij}w_i \\ p_j &= p_j^w + B_{ji}v_i + C_{ji}w_i \ , \end{aligned} \tag{1.3}$$

where v_i, w_i are respectively the tangential and normal nodal displacements, q_i, p_i are the tangential and normal (compressive) nodal forces, q_j^w, p_j^w are the nodal reactions that would be generated by the external forces $\mathbf{F}$ if all the nodal displacements were constrained to be zero and $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are partitions of the reduced stiffness matrix $\boldsymbol{\kappa}$. We note that with this terminology, the coupling between tangential displacements and normal reactions is defined by the matrix $\mathbf{B}$ and hence the condition for Melan's theorem to apply is $\mathbf{B} = \mathbf{0}$.

We define the Coulomb friction law for node i by the relations

$$w_i \ \geq \ 0 \ ; \quad p_i \geq 0 \tag{1.4}$$

$$w_i > 0 \Rightarrow p_i = q_i = 0 \tag{1.5}$$

$$p_i > 0 \Rightarrow w_i = 0 \tag{1.6}$$

$$-fp_i \leq q_i \leq fp_i \tag{1.7}$$

$$|q_i| < fp_i \Rightarrow \dot{v}_i = 0 \tag{1.8}$$

$$0 < |q_i| = fp_i \Rightarrow \text{sgn}(\dot{v}_i) = -\text{sgn}(q_i) \; , \tag{1.9}$$

where f is the coefficient of friction.

For shakedown to be possible, it is necessary that there exists at least one shakedown vector $\mathbf{v} = \tilde{\mathbf{v}}$, for which the contact tractions satisfy the Coulomb friction inequalities (1.7) at all nodes $i \in (1, N)$ and at all times during the loading cycle. Assuming that all nodes remain in contact in this state, so that $w_i = 0$ for all i, and substituting (1.3) into the Coulomb friction inequalities (1.7), we obtain

$$\begin{aligned}(A_{ji} - fB_{ji})v_i &< fp_j^w - q_j^w \\ (A_{ji} + fB_{ji})v_i &> -fp_j^w - q_j^w \; ,\end{aligned} \tag{1.10}$$

Each of these $2N$ inequalities can be represented as a directional hyperplane in the N-dimensional space of coordinates v_i, such that points on one side only of each hyperplane are admissible. During the loading cycle, the hyperplanes move, whilst retaining the same orientation, and if they impinge on the instantaneous operating point P defined by the coordinates v_i, they will cause it to move in the coordinate direction associated with slip at the node in question in the direction defined by (1.9).

Figure 1.1 illustrates this process in v_i-space for a two-node system. The lines I, II, III, IV define the frictional constraints associated with incipient nodal slip in the directions $\dot{v}_1 < 0, \dot{v}_1 > 0, \dot{v}_2 < 0, \dot{v}_2 > 0$ respectively and the regions excluded by the frictional constraints are shown shaded. Thus, at the instant illustrated, the operating point $P(v_1, v_2)$ can exist only in the 'safe' unshaded region between the four lines. If changes in the applied loads cause the active constraint IV to advance (in the sense of excluding more of the space), slip will occur in the direction $\dot{v}_2 > 0$ and P will be 'pushed' upwards by the constraint. For a fairly general periodic loading scenario, the constraints will advance and recede whilst retaining the same slope (which is determined only by the stiffness matrix and the coefficient of friction). Notice incidentally that the direction of slip is generally not orthogonal to the constraint line, which in a heuristic sense is an indication of the non-associative nature of the friction law.

For shakedown to be possible, there must exist some region that is safe at *all* times during the loading cycle. This can be established by identifying the extreme positions of each constraint (i.e. the position at which the constraint excludes the maximum region of space) and plotting a diagram similar to Figure 1.1, but using these extreme positions (which generally will occur at different times during the cycle). It can be shown [1] that the two-node system will shake down for all possible initial conditions if the safe shakedown

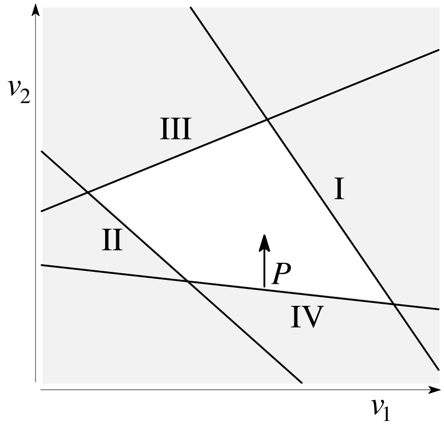

Fig. 1.1 Motion of the instantaneous operating point P due to the advance of constraint IV

region defined by these extreme constraints is a quadrilateral, but that if it is triangular, the steady state may be either shakedown or cyclic slip, depending on the initial conditions.

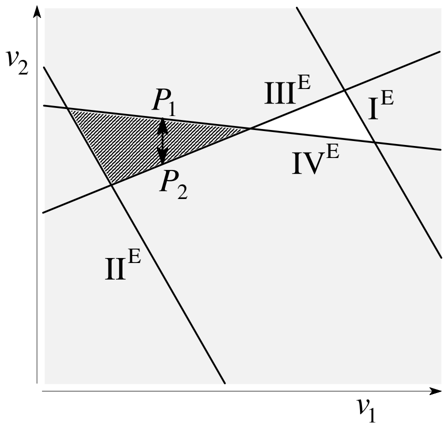

Fig. 1.2 Cyclic slip limit cycle in the case where the safe shakedown region is triangular

The latter case is illustrated in Figure 1.2, where the lines I^{E}, II^{E}, $\mathrm{III}^{\mathrm{E}}$, IV^{E}, represent the extreme positions of I, II, III, IV respectively and the safe shakedown region is the unshaded triangle. If an initial condition is chosen that lies within this triangle, no slip will ever occur, so the system shakes down *a fortiori*. However, if the initial condition lies in the dark shaded triangle, cyclic slip will occur as illustrated in the figure.

Consider now the effect of increasing λ in equation (1.2). We assume that the mean load $\mathbf{F}_0$ is such that the only possible states of the system for $\lambda = 0$ are those in which both nodes are in contact, in which case the safe shakedown region must be quadrilateral. As λ is increased, the extreme positions of the constraints advance, the safe shakedown region is decreased and at some critical value λ_L it becomes triangular. Further increase in λ reduces the size of this triangle until at a higher critical value λ_U it becomes null. We conclude that for $\lambda < \lambda_L$ the system always shakes down, for $\lambda_L < \lambda < \lambda_U$ we may get shakedown or cyclic slip depending on the initial conditions, and for $\lambda > \lambda_U$ shakedown is impossible for all initial conditions. Both critical values correspond to conditions where three of the four constraint lines intersect in a point. Thus, they can easily be found by solving all possible combinations of three linear equations and selecting those for which the resulting configuration satisfies certain inequalities [1]. Notice that an alternative statement of Melan's theorem in this context would be that $\lambda_L = \lambda_U$.

This strategy can be extended to the N-node discrete case, though the number of linear operations needed to establish the value of λ_L, λ_U increases combinatorially with N. An alternative method of establishing λ_U is to configure it as a constrained linear optimization problem [5].

When the system is uncoupled, the two constraints at any given node represent parallel hyperplanes. The topology of the safe shakedown region is then independent of λ and all the constraints remain active until this region becomes null, confirming that Melan's theorem applies when the system is uncoupled ($\mathbf{B}$=$\mathbf{0}$).

1.3.1 Existence, Uniqueness and Wedging

It is well-known that frictional systems of this kind exhibit problems of existence and uniqueness of the quasi-static incremental solution if the coefficient of friction is sufficiently high. In the present formulation, the coefficient of friction changes the inclination of the constraint lines in Figure 1.1, which can cause two distinct kinds of anomolous evolutionary behaviour. The line IV in Figure 1.1 is associated with slip in the direction $v_2 > 0$ at node 2 and this is clearly possible if IV advances. However, if increasing the coefficient of friction causes IV to rotate clockwise past the vertical position, its advance will be inconsistent with the appropriate direction of slip and the quasi-static evolutionary algorithm fails to return a result. This behaviour is exactly analogous with that exhibited by Klarbring's one-node model [6] and results in an unstable motion to a new state involving separation at the node in question.

For multinode systems, a qualitatively different failure of the algorithm can occur in which the advance of two or more constraints each separately permit motion of the operating point P in the appropriate direction, but the several constraints conspire to eliminate all possible slip directions. This is illustrated for the 2-node system in Figure 1.3. Advance of either constraint allows P to move appropriately until it reaches the intersection of I and III when the quasi-static evolutionary algorithm fails. In this case, a more complex dynamic transition occurs to a state involving one or both of the nodes separating.

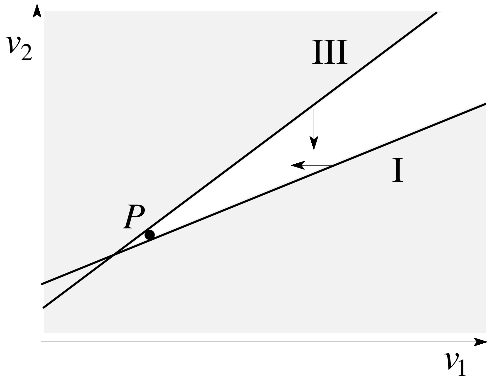

Fig. 1.3 Configuration in which advance of either constraint leads to failure of the evolutionary algorithm

Another phenomenon observed at high coefficients of friction is that of wedging [3], in which the system can exist in a state of stress even when all external loads are removed. In the present formalism, the removal of all external loads $p_i^w = q_i^w = 0$ implies that all the constraints (1.10) pass through the origin in v_i space. Recalling that the slopes of the constraint surfaces are independent of the applied loads, we conclude that wedging is possible if and only if the constraint surfaces moved to the origin leave a safe region that is open to infinity. In the two-node case this would be an infinite 'safe' sector. Changing the external loads will change the geometry of this region local to the origin, but cannot close the region at infinity.

1.3.2 Ratcheting

Qualitatively different behaviour can be obtained if the system exhibits a rigid-body mode. For example, Mugadu *et al.* [10] analyzed the motion of a flat rigid punch indenting an elastic half plane and subjected to varying normal and tangential loads. If the loads are such as to cause all points in the contact area to slip at some time during the cycle, but not all at the same time, it is possible for the punch to 'walk' over the half plane by a constant increment during each cycle.

A related problem concerns the frictional behaviour of an axisymmetric elastic bushing which is a force fit inside a connecting rod end, considered by

Antoni *et al.* [2]. In this case, a uniform mean load is generated by the force fit and superposed oscillating loads during engine operation may cause slip at the bushing/connecting rod interface. However, if every node is caused to slip circumferentially the same distance, the stress state of the system will be unaffected, so this constitutes a rigid-body mode for the system and shows that the contact stiffness matrix $\boldsymbol{\kappa}$ must be singular. In v_i-space, this implies that all the constraint surfaces will be parallel to the line $v_1 = v_2 = ... = v_N$ and in particular that all the constraint lines in Figure 1.1 would be inclined at 45 degrees. We illustrate this case in Figure 1.3, where we also show the directions of slip implied by each constraint. During periodic loading these constraints will advance and recede. If there is any region that remains safe throughout the cycle, the system will shake down. If not, the steady-state may consist of ratcheting (illustrated by the displacement steps at the top right of Figure 1.4) or of cyclic slip at only one node, depending on the loading sequence.

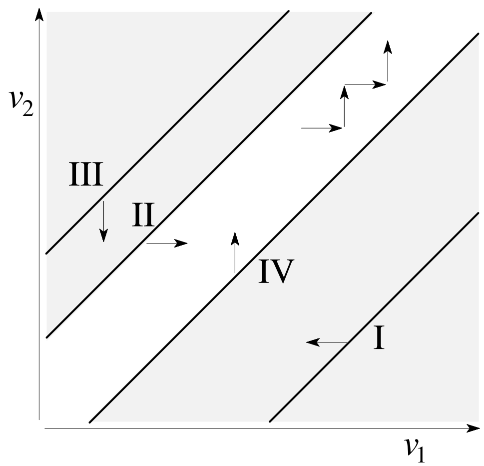

Fig. 1.4 Constraint space for a two-node system with a rigid-body mode

Suppose we look along the rigid-body line, or (equivalently) project the motion of the constraints onto the line $v_1 + v_2 = 0$ orthogonal to the rigid-body line. We shall then simply see each of the four constraints advancing and retreating along a line as time progresses. We could plot the position of each constraint and the region excluded as a function of time. Figure 1.5 shows such a plot. The unshaded region is safe at any particular time. Shakedown is impossible for the case illustrated, since there is no region that is unshaded at all times.

The operating point P at any given time will be moved up or down if forced to do so by an advancing constraint. For the particular case illustrated, the motion of P is shown by the dashed line. Initially P is pushed up by I ($\dot{v}_1 < 0$)

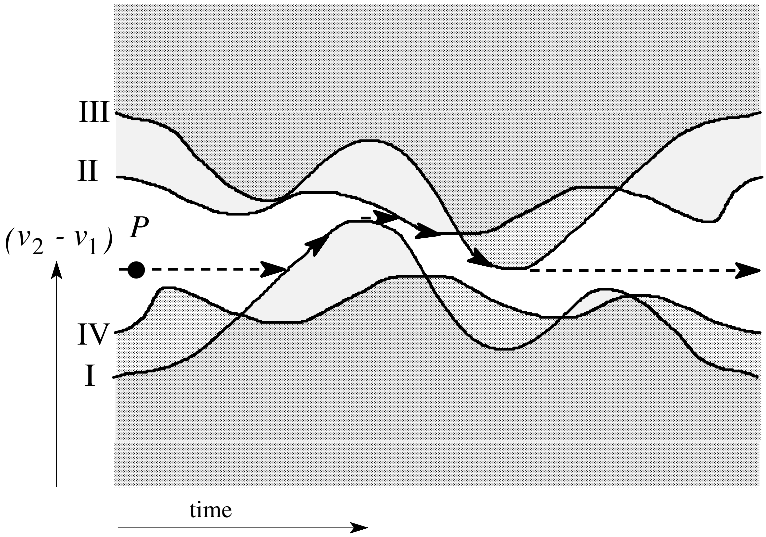

Fig. 1.5 Dominant constraints as a function of time

until it reaches its maximum. It is later pushed down by II ($\dot{v}_1 > 0$) and later still (also down) by IV ($\dot{v}_2 < 0$). Since the only motion at node 2 is that driven by constraint IV, the system will ratchet in the direction $\dot{v}_1, \dot{v}_2 < 0$, though the slip at node 1 is non-monotonic.

The problem of determining this scenario is equivalent to that of tracking the motion of a ball falling through the space between the extreme lines — i.e. the set of points that are safe with respect to all four constraints as a function of time. Although four constraints combine to form this safe region, it can then be characterized by only two surfaces — the envelope of I,III and the envelope of II,IV.

Since the point P has only one degree of freedom (up and down), its motion is completely determined once it strikes either one of these surfaces. The only effect of the initial condition is to determine which surface is struck first, but since *ex hyp.* we assume no shakedown, it must alternate between the two surfaces and hence the steady state must be unique. In fact, if we make even one pass through the system, the end point will be independent of the initial point. Thus, we reach a unique steady state after one cycle of loading.

This unique steady state could comprise ratcheting or cyclic slip, depending on which of the four constraints are active at some time during the cycle. Cyclic slip will occur if only those constraints associated with a single node (i.e. (I,II) or (III,IV)) are active during the steady-state cycle. If at least one constraint from each pair is active, ratcheting will almost always occur.

This procedure can be extended to multi-node systems with a rigid-body mode. For example, for a system of three nodes, we could project the instantaneous constraints onto a plane orthogonal to the direction $v_1 = v_2 = v_3$, as shown in Figure 1.6, which also shows the motion of the point P associated with the motion of each constraint. The motion of P could then be tracked by an algorithm similar to that illustrated in Figure 1.1 and would be equivalent to the path of a ball dropped through a tube whose axis is time and whose cross section has the form of the instantaneous safe region in Figure 1.6. The uniqueness of the steady state for this and higher-order systems remains an open question.

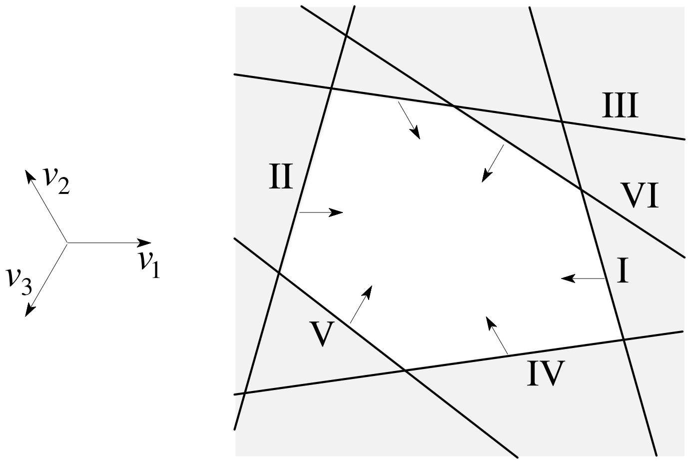

Fig. 1.6 Projected view of v_i space for a three-node system with a rigid-body mode

1.4 Conclusions

We have demonstrated that for two-dimensional discrete frictional systems subjected to periodic loading, upper and lower bounds can be placed on a scalar loading factor such that above the upper bound, shakedown is impossible, below the lower bound, shakedown occurs for all initial conditions, but between the bounds, either shakedown or cyclic slip may occur depending on the initial conditions. General procedures can be identified for determining these bounds. If there is no coupling between tangential nodal displacements and normal reactions, the two bounds coincide and Melan's theorem applies.

The motion of a point in slip-displacement space representing the instantaneous position of the system proves to be a fruitful tool for investigating the behaviour of discrete frictional systems. Application to a simple two-node system with a rigid-body mode shows that Melan's theorem applies and that above the shakedown limit, a unique steady state is achieved after one cycle.

Acknowledgements. We wish to thank Michele Ciavarella, Enrico Bertocchi and Yong Hoon Jang for their contributions to these results. Youngju Ahn also thanks the Electric Power National Scholarship Program of the Korean Ministry of Commerce, Industry and Energy (MOCIE) for financial support.

References

1. Ahn, Y., Bertocchi, E., Barber, J.: Shakedown of coupled two-dimensional discrete frictional systems. Journal of the Mechanics and Physics of Solids 56, 3433–3440 (2008)
2. Antoni, N., Nguyen, Q.S., Ligier, J.L., Saffré, P., Pastor, J.: On the cumulative microslip phenomenon. European Journal of Mechanics A/Solids 26, 626–646 (2007)
3. Barber, J., Hild, P.: On wedged configurations with coulomb friction. In: Wriggers, P., Nachenhorst, U. (eds.) Analysis and Simulation of Contact Problems, pp. 205–213. Spinger Verlag (2006)
4. Barber, J., Klarbring, A., Ciavarella, M.: Shakedown in frictional contact problems for the continuum. Comptes Rendus Mecanique 336, 34–41 (2008)
5. Björkman, G., Klarbring, A.: Shakedown and residual stresses in frictional systems. In: Gladwell, H.G.G.M.L., Kalousek, J. (eds.) Contact Mechanics and Wear of Rail/Wheel Systems II: Proceedings of the 2nd International Symposium, pp. 27–39. University of Waterloo Press (1987)
6. Cho, H., Barber, J.: Dynamic behavior and stability of simple frictional systems. Mathematical and Computer Modeling 28, 37–53 (1998)
7. Churchman, C.M., Korsunsky, A.M., Hills, D.: The application of plasticity principles to friction. Journal of Strain Analysis 41, 323–328 (2006)
8. Klarbring, A., Ciavarella, M., Barber, J.: Shakedown in elastic contact problems with coulomb friction. International Journal of Solids and Structures 44, 8355–8365 (2007)
9. Melan, E.: Theorie statisch unbestimmter Systeme aus ideal-plastichem Baustoff. Sitzungsber. d. Akad. d. Wiss., Wien 2A(145), 195–218 (1936)
10. Mugadu, A., Sackfield, A., Hills, D.: Analysis of a rocking and walking punch part i: Initial transient and steady state. ASME Journal of Applied Mechanics 71, 225–233 (2004)