FRICTIONAL SYSTEMS SUBJECTED TO OSCILLATING LOADS

J.R.Barber

Department of Mechanical Engineering
University of Michigan
Ann Arbor

This talk draws on work performed in collaboration with: Youngju Ahn, Enrico Bertocchi, Hanbum Cho, Michele Ciavarella, David Hills, Yong Hoon Jang, Anders Klarbring, Mehdi Movassaghi, and Andrea Spagnioli.

Marseille, October 2010

Engineering systems typically comprise numerous components and the interfaces between these components are often frictional.

- Damping due to microslip in nominally static bolted joints.
- Fretting fatigue (e.g. in turbine blade roots).
- Settlement of soils or masonry structures.
- Tectonic plate movement.
- Frictional ‘wedging’ during automatic assembly processes.
- Frictional slip in belt drives....

Many of these applications involve periodic (cyclic) loading, due to vibration, or to repetitive operations.

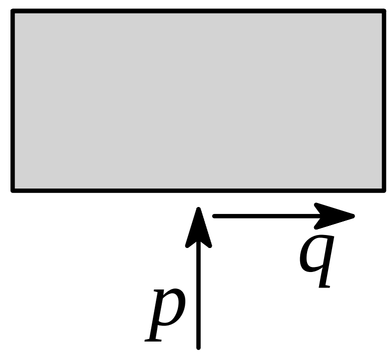
Friction is a major factor in fretting fatigue failures (aircraft engines) and it also determines the amount of hysteresis loss (effective damping) in systems that are nominally monolithic.

It has been estimated that frictional hysteresis in assembled structures accounts for more energy dissipation that internal material damping.

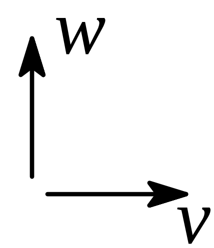
The Coulomb friction law

For a two-dimensional system, each point in the nominal interface must be in one of four states:-

Stick	$w = 0\ ;\ \dot{v} = 0\ ;\ p \geq 0\ ;\ \lvert q \rvert \leq fp$
Separation	$w > 0\ ;\ p = 0\ ;\ q = 0$
Forward slip	$w = 0\ ;\ \dot{v} > 0\ ;\ p \geq 0\ ;\ q = -fp$
Backward slip	$w = 0\ ;\ \dot{v} < 0\ ;\ p \geq 0\ ;\ q = fp\ ,$



contact tractions



displacements

where f is the coefficient of friction.

The Coulomb law has been criticized extensively by tribologists, but it is sufficiently close to the experimentally observed behaviour of macroscopic systems to justify use in engineering design, because of its simplicity.

This simplicity is a bit of an illusion. The governing equations are linear (in two-dimensions), but the associated inequalities lead to quite complex behaviour [non-uniqueness, non-existence, jumps in quasi-static response, wedging,...], when the coefficient of friction is ‘sufficiently large’.

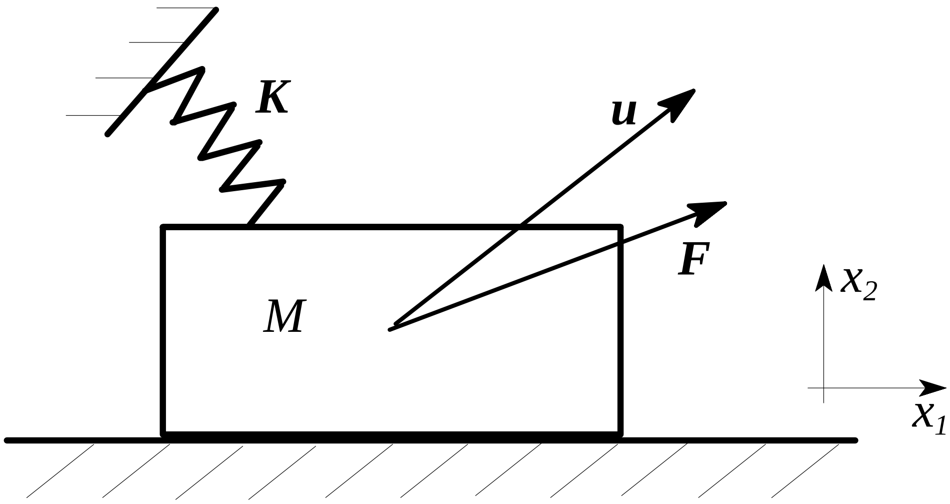
In this talk, I focus on cases where we are below the critical coefficient of friction, so the incremental problem is well-posed and has a unique solution.

A critical feature of the frictional contact problem is that it is load-history dependent.

This arises because the slip condition includes the sign of the slip velocity $\dot{v}$ which is a time derivative.

We recall

Stick	$w=0\,;\;\; \dot{v}=0\,;\;\; p\geq 0\,;\;\; \lvert q\rvert \leq fp$
Separation	$w>0\,;\;\; p=0\,;\;\; q=0$
Forward slip	$w=0\,;\;\; \dot{v}>0\,;\;\; p\geq 0\,;\;\; q=-fp$
Backward slip	$w=0\,;\;\; \dot{v}<0\,;\;\; p\geq 0\,;\;\; q=fp\,,$



For values of the applied force $\boldsymbol{F}$(t) that involve contact, the position $\boldsymbol{u}$(t) of the mass M cannot be determined from the instantaneous value of $\boldsymbol{F}$(t) alone.

We need an initial condition and a loading history.

Multinode system

Suppose there are now N contact nodes and we use the notation

$$\left\{ \begin{array}{c} q_i \\ p_i \end{array} \right\} = \left\{ \begin{array}{c} q_i^w \\ p_i^w \end{array} \right\} + \left[\begin{array}{cc} A_{ij} & B_{ij} \\ B_{ji} & C_{ij} \end{array} \right] \left\{ \begin{array}{c} v_j \\ w_j \end{array} \right\} ,$$

where p_i^w, q_i^w are the contact tractions that would be produced if the nodes were all welded in contact at $v_i = w_i = 0$.

If all the nodes are in contact ($w_j = 0$ for all j),

$$q_i = q_i^w + A_{ij} v_j$$
$$p_i = p_i^w + B_{ji} v_j$$

History dependence implies that the system in some sense posesses **‘memory’**.

In frictional systems, it is clear that the system memory must reside in the slip displacements in regions that are not slipping (at the present instant), and hence that lie strictly between the two boundaries defined by the frictional inequalities.

These are contained in the vector v_i.

The trajectory in v_i-space is a good way of tracking the evolution of the system.

If the applied loads are periodic in time, we anticipate eventually reaching a steady state.

We might get to this after a finite time, or we might approach it asymptotically.

Because of the history dependence, the steady state may depend on the initial condition, or on an initial loading path leading to the periodic state.

What does this dependence on initial conditions look like?

Suppose that the loading is such that no nodes leave contact ($w_i = 0$ for all i).

The instantaneous state of the system $\{v_1, v_2, v_3 ... v_N\}$ defines a point P in v_i-space.

P must satisfy the frictional constraints

$$(A_{ji} - fB_{ji})v_i(t) \leq fp_j^w(t) - q_j^w(t)$$
$$(A_{ji} + fB_{ji})v_i(t) \geq -fp_j^w(t) - q_j^w(t)$$

These define a set of $2N$ directional hyperplanes in v_i-space if there are N contact nodes.

These hyperplanes advance and recede as the load varies periodically, but retain the same slope (which is determined by $\boldsymbol{A}, \boldsymbol{B}$ and f.

During periods of slip, one or more of the advancing hyperplanes 'push' the point P around the space.

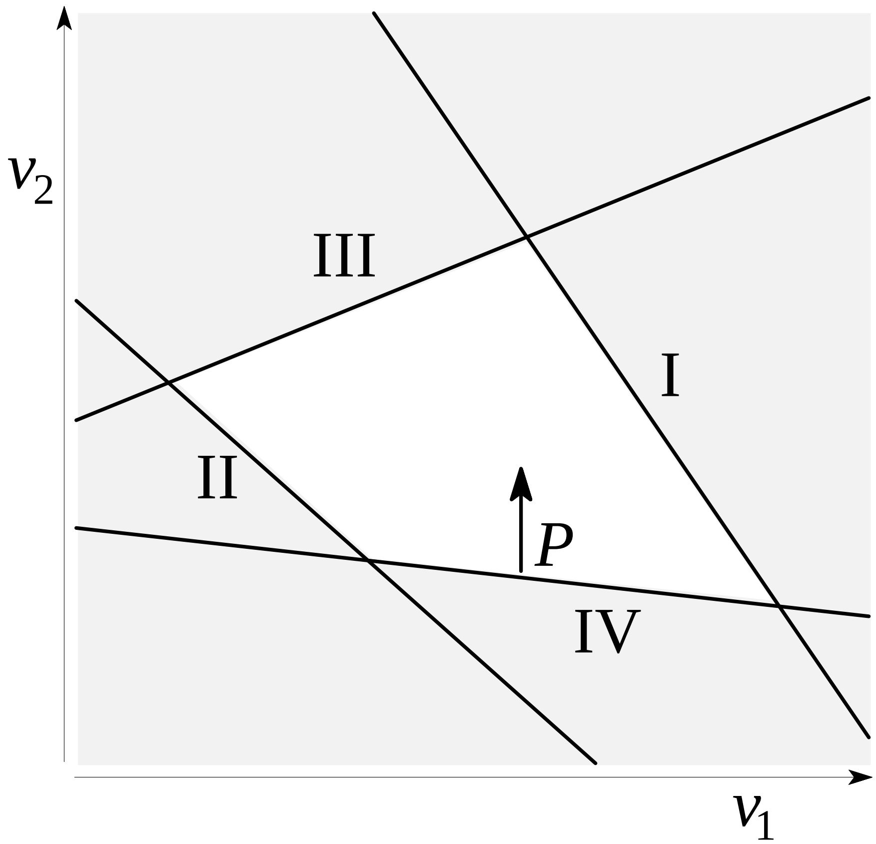
Two-node system

For $N = 2$, there are just four such constraints

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

These control the motions

I: $\dot{v}_1 < 0$; II: $\dot{v}_1 > 0$; III: $\dot{v}_2 < 0$; IV: $\dot{v}_2 > 0$



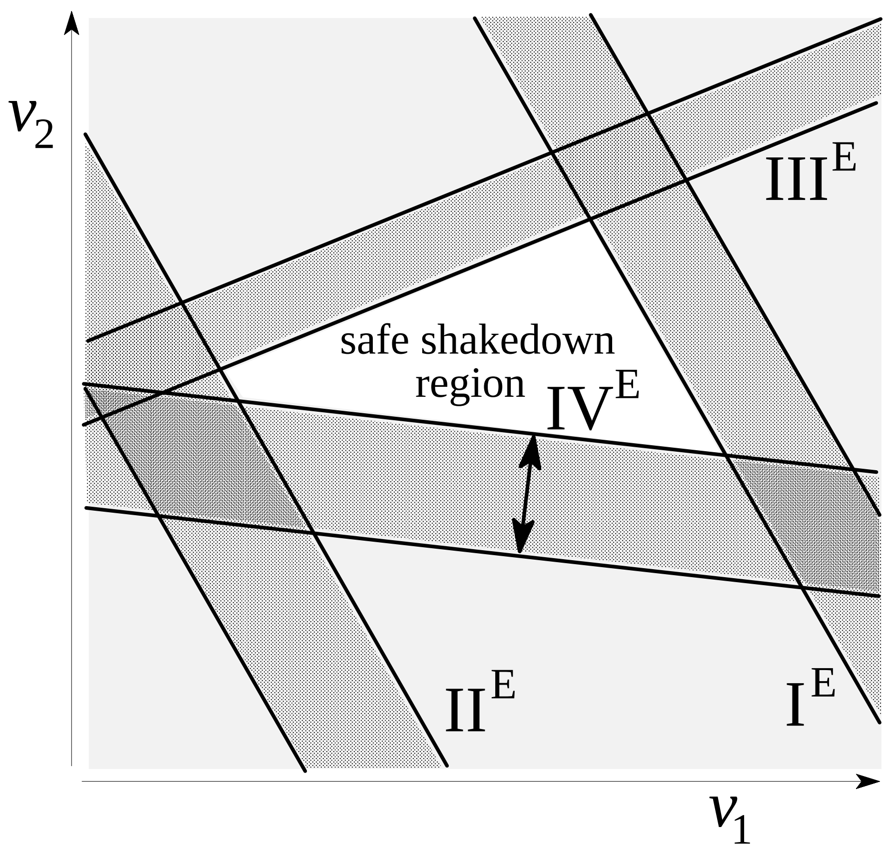
If IV advances, it will push P in the direction $\dot{v}_2 > 0$ (forward slip at node 2).

A special case here concerns ‘shakedown’ which means that in the steady state, there is no frictional slip.

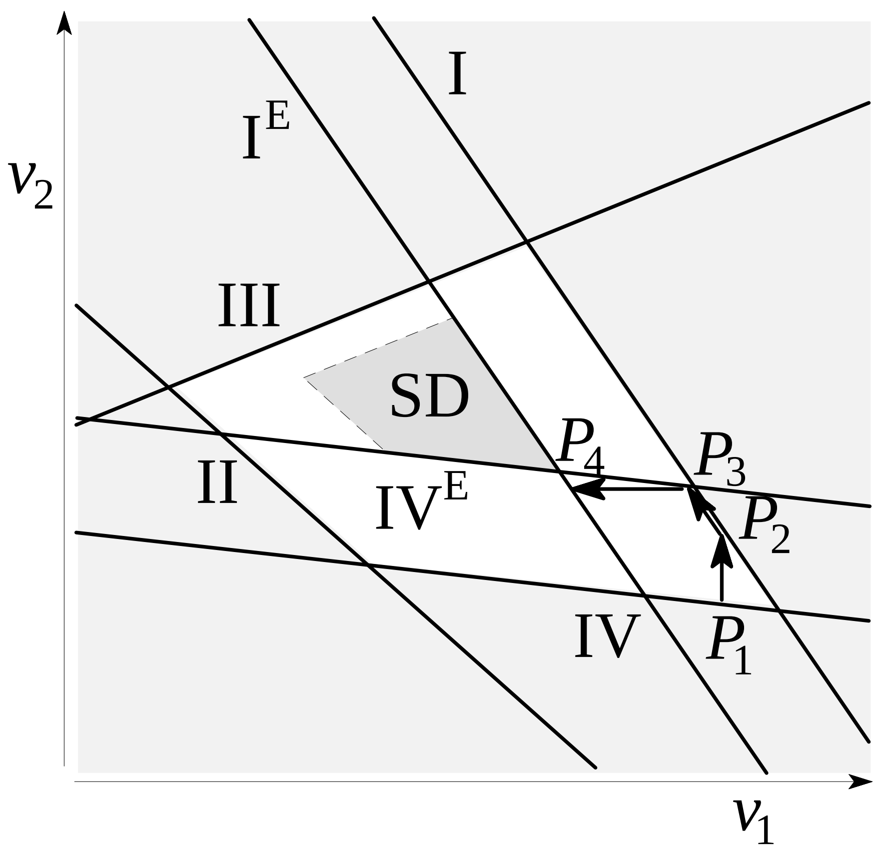
In other words, slip that has occurred early in the loading process is sufficient to prevent further slip in the steady state.

The lines I^E, II^E etc. define the extreme positions (maximum advance) of the constraints.

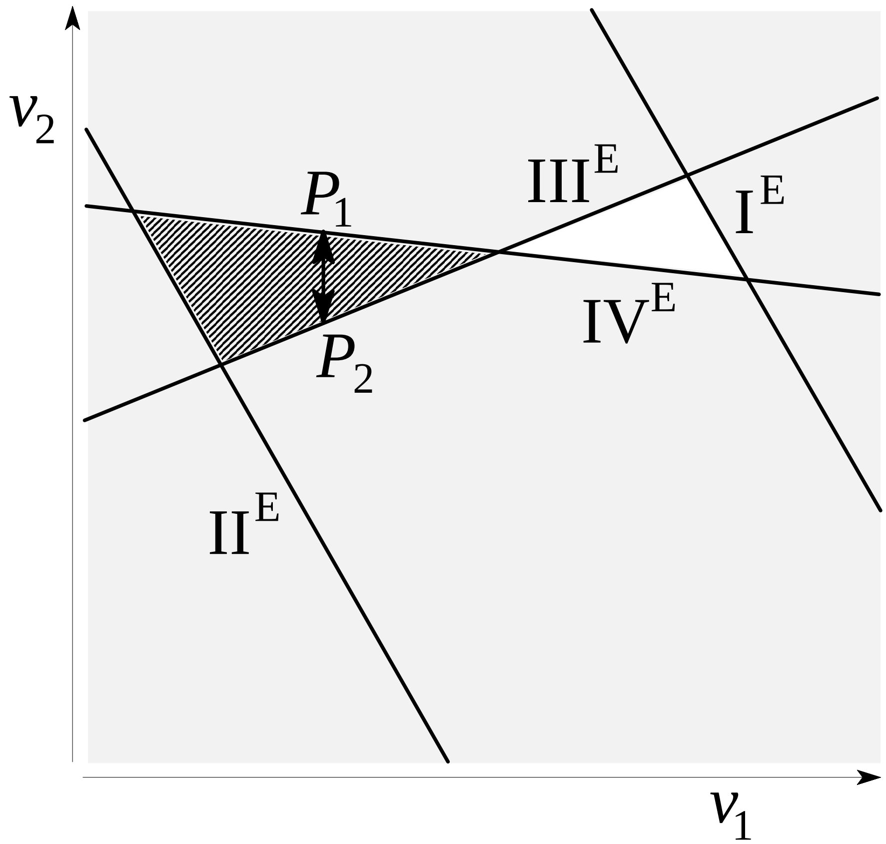
For shakedown to be possible, there must be a *safe shakedown region* that is never excluded by any of the constraints.



Advance of IV ($\dot{v}_2 > 0$) to IV$^{\mathrm{E}}$, followed by advance of I ($\dot{v}_1 < 0$) to I$^{\mathrm{E}}$, pushes P towards the safe shakedown region 'SD'.



P always reaches the safe shakedown region if this is a quadrilateral, but not always if it is triangular.



With an appropriate initial condition, it can end up oscillating between P_1 and P_2.

This criterion can be generalized to an N-node system.

In the analogous process of elastic-plastic deformation there is a theorem (Melan's theorem) which states essentially that

The system will shake down if it can.

— i.e. if a suitable state of residual stress exists.

Clearly this is not always true for frictional systems.

[We just found a counter-example — the triangular safe region.]

We have proved that the frictional Melan's theorem applies *if and only if* there is no coupling between normal and tangential loading.

In other words, in

$$\left\{\begin{array}{c} q_i \\ p_i \end{array}\right\} = \left\{\begin{array}{c} q_i^w \\ p_i^w \end{array}\right\} + \left[\begin{array}{cc} A_{ij} & B_{ij} \\ B_{ji} & C_{ij} \end{array}\right] \left\{\begin{array}{c} v_j \\ w_j \end{array}\right\} ,$$

the matrix $\boldsymbol{B} = \boldsymbol{0}$.

In mathematical terms, the theorem states:-

"If we can find a time-independent *safe shakedown vector* $\tilde{\boldsymbol{v}}$ such that the corresponding reaction vector

$$\tilde{\boldsymbol{r}}(t) = \boldsymbol{r}^w(t) + \boldsymbol{K}\tilde{\boldsymbol{v}}$$

lies strictly within the friction cone at all nodes and all times t, then an uncoupled system will shake down, though not necessarily to the state defined by $\tilde{\boldsymbol{v}}$."

$$\boldsymbol{r}_i = \begin{Bmatrix} q_i \\ p_i \end{Bmatrix}$$

Actually, the proof applies to both two and three-dimensional systems, so the tangential reaction q_i can itself be a vector in the contact plane.

Proof

We define the norm

$$A = \frac{1}{2}\left(\tilde{\boldsymbol{v}} - \boldsymbol{v}\right)^T \boldsymbol{K}\left(\tilde{\boldsymbol{v}} - \boldsymbol{v}\right) \;,$$

which is a scalar measure of the difference between the instantaneous slip-displacement vector and the safe shakedown vector. Since $\boldsymbol{K}$ is positive definite, $A \geq 0$.

The time-derivative of A is

$$\dot{A} = -\left(\tilde{\boldsymbol{v}} - \boldsymbol{v}\right)^T \boldsymbol{K}\dot{\boldsymbol{v}} = -\sum_{i-1}^{n}\left(\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i\right)\cdot\dot{\boldsymbol{v}}_i$$

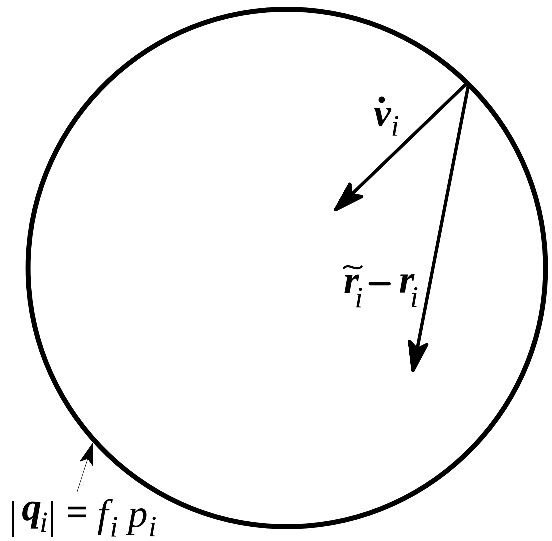
The only contributions to

$$\dot{A} = -\sum_{i-1}^{n} (\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i) \cdot \dot{\boldsymbol{v}}_i$$

come from nodes where $\dot{v}_i \neq 0$ — i.e. nodes that are slipping.

If $\boldsymbol{B} = 0$, the normal component of $\boldsymbol{r}_i$ is not affected by $\boldsymbol{v}$.

It follows that $\boldsymbol{r}_i$ and $\tilde{\boldsymbol{r}}_i$ have the same normal component p_i and therefore lie on a plane in $\boldsymbol{r}$-space that cuts the friction cone for node i in a circle.



If node i is slipping, $\boldsymbol{r}_i$ must be on the boundary of the circle, the safe shakedown reaction $\tilde{\boldsymbol{r}}_i$ must be strictly inside the circle, and the Coulomb friction flow rule requires that the slip velocity $\dot{\boldsymbol{v}}_i$ be directed perpendicular into the circle at $\tilde{\boldsymbol{r}}_i$.

Thus, the dot product $(\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_i) \cdot \dot{\boldsymbol{v}}_i$ is always positive and $\dot{A} < 0$.

Any slip that occurs reduces A and hence makes the system approach the safe shakedown state monotonically.

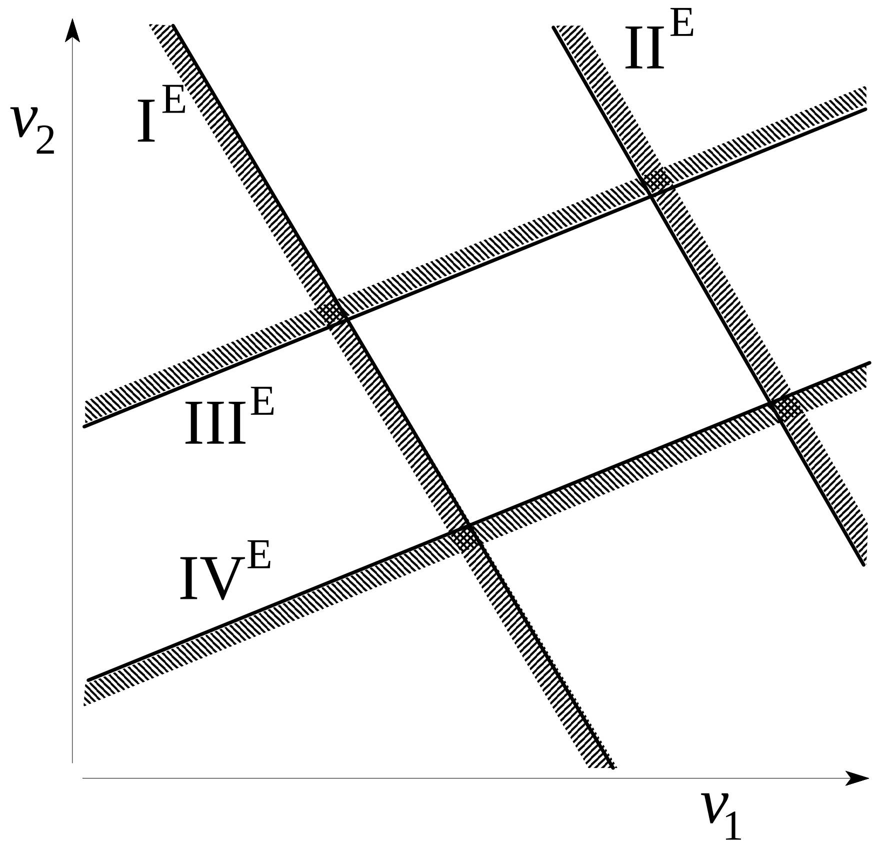
Whenever there is coupling $(\boldsymbol{B} \neq \mathbf{0})$, counter examples to the theorem can be found — the occurrence of shakedown depends on initial conditions.

We have seen an example of this in the two-node case.

If $\boldsymbol{B} = \boldsymbol{0}$, the constraints reduce to

$$\begin{aligned} A_{ji} v_i(t) &\leq f p_j^w(t) - q_j^w(t) \\ A_{ji} v_i(t) &\geq -f p_j^w(t) - q_j^w(t) \end{aligned}$$

and hence the two hyperplanes corresponding to a given node j are parallel.



Clearly the safe shakedown region, if it is not null, will be a parallelogram, and hence *a fortiori* a quadrilateral, as required for Melan's theorem to apply.

A conjecture

- For uncoupled systems in the steady-state , the tractions at the slipping nodes, the time-varying terms in the remaining tractions and displacements, and hence the energy dissipation per cycle are independent of initial conditions.

  [The frictional Melan's theorem would be a special case, where the dissipation is zero].

- For coupled systems, the steady-state and the energy dissipation per cycle will sometimes depend of initial conditions.

There is anecdotal evidence for this conjecture:-

(i) Fretting fatigue tests [where frictional slip is causing damage that eventually leads to a fatigue failure] are very consistent for smooth ‘Hertzian-like’ contact geometries.

In this case, the system is reasonably approximated by two half planes, which involves no coupling.

Similar tests for a flat indenter pressed against a plane surface, [where there is significant coupling] give very erratic results.

(ii) Experimental tests on the effective damping in bolted joints shows that the results are very erratic.

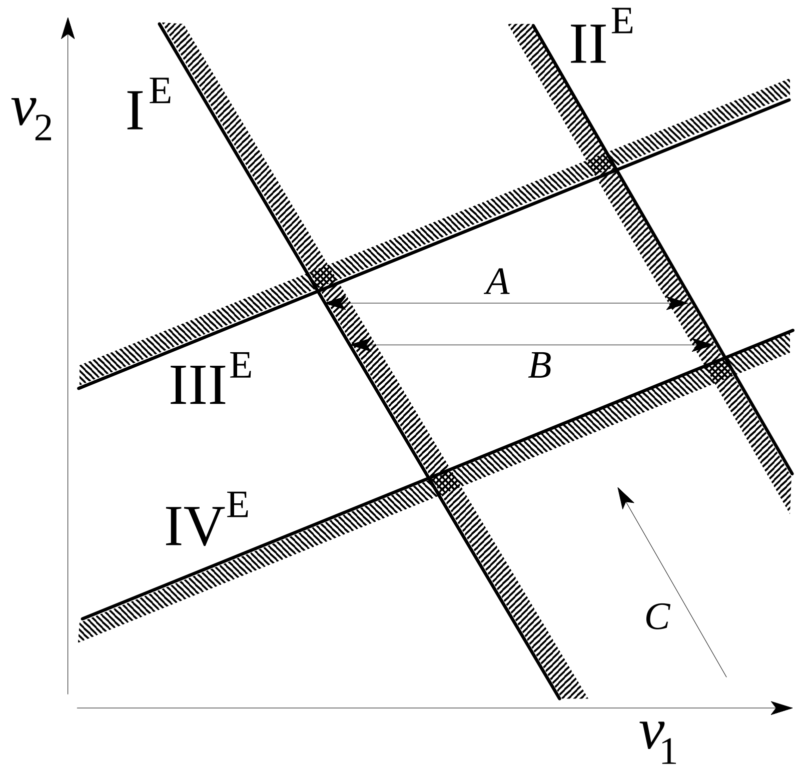
Apparently identical systems give different results and even the same system, if disassembled and then reassembled, can give very different results.

[This is like changing the initial conditions].

For uncoupled systems, the *normal nodal forces are unique* (since they are equal to p_j^w and are defined by the periodic loading).

It follows that *during sliding* (in a given direction) the tangential tractions are also unique.

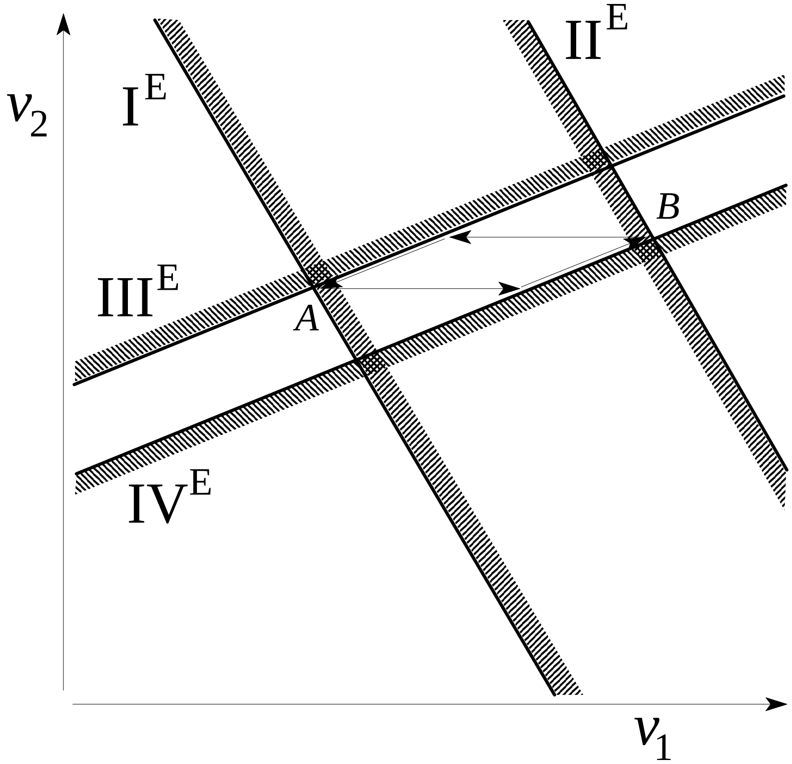
Consider some scenarios for the two-node system:-



The extreme positions of $\mathrm{III}^{\mathrm{E}}$, IV^{E} leave a ‘safe’ space, but I^{E}, II^{E} overlap, forcing cyclic slip at node 1.

A and *B* are two possible steady states with different ‘locked-in’ displacements v_2.

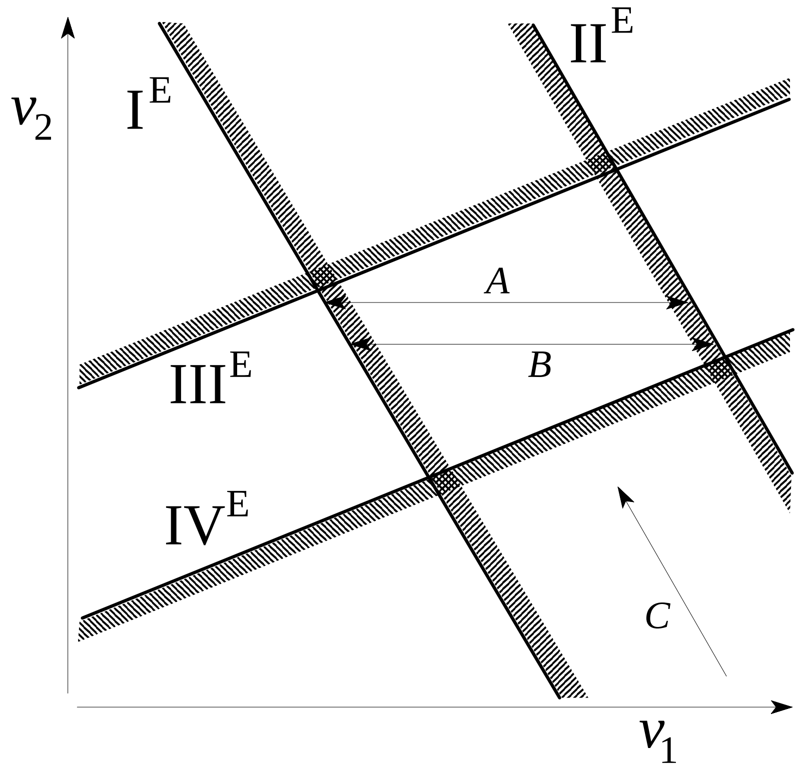
The safe space between III$^{\mathrm{E}}$ and IV$^{\mathrm{E}}$ is not sufficient to ensure that node 2 is always stuck.



Here both nodes slip and the steady state is unique.

In general (including for coupled systems) we expect that the non-uniqueness in the steady state arises from the locked-in displacements at the ‘permanently stuck nodes’.

A and B differ only by a constant in v_1, v_2 actually by a translation of the trajectory in direction C.



The time-derivatives of the displacements (and hence of all the tractions) are the same for all scenarios.

The steady state will appear unique if we look in direction C (or project the motion on a line perpendicular to C).

To develop a similar projection that works for the N-node uncoupled system, we need to change the basis of the slip-displacement N-vector $\boldsymbol{v}$.

We define a new N-vector $\boldsymbol{f}$ through the linear operation

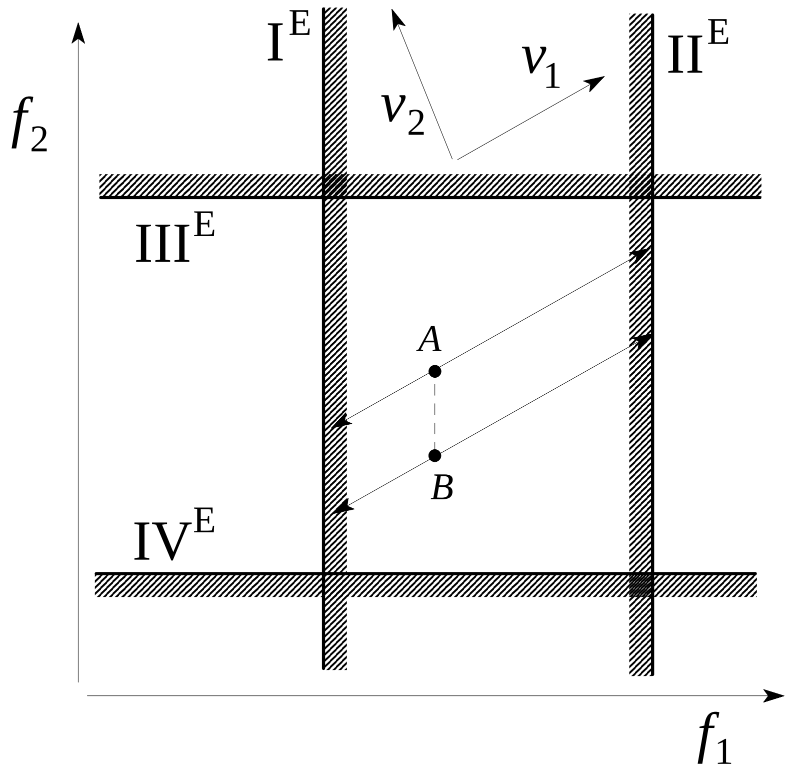
$$\boldsymbol{f} \equiv \boldsymbol{A}\boldsymbol{v} \quad \text{or} \quad f_j = A_{ji} v_i \ .$$

The constraints at node j now take the form

$$\begin{aligned} f_j &\leq f p_j^w - q_j^w \\ f_j &\geq -f p_j^w - q_j^w \ . \end{aligned}$$

Physically, f_j is the *change* in the tangential reaction q_j due to the slip displacements v_i.

The two-node trajectory in f-space looks like this:-



The constraints are now perpendicular to the axes, but slip at one node corresponds to an inclined trajectory.

Projection on a line perpendicular to f_2 generates a unique one-dimensional trajectory in f-space.

For N nodes, suppose that in the steady state, M nodes slip at least once during each cycle and $N - M$ nodes never slip (they comprise the set $\mathscr{S}_0$ of *permanently stuck nodes*).

Actually, if a node slips once during the steady state, it must slip at least twice if the process is to be cyclic.

We shall obtain a unique trajectory if we project the actual trajectory on the M-dimensional f-space orthogonal to $f_j, j \in \mathscr{S}_0$.

The tractions at the slipping nodes are

$$q_j = q_j^w + f_j$$

and hence are unique.

In this reduced space, each constraint will be active at least for part of the cycle.

The conjectured uniqueness of the steady state for uncoupled system could be proved if we could establish

(i) That for all possible steady states, the set $\mathscr{S}_0$ of permanently stuck nodes is unique. For example, that the assumption of two distinct states with different sets $\mathscr{S}_0$ leads to a contradiction of some kind.

(ii) That for a system in which $\mathscr{S}_0$ is null (i.e. all nodes slip at least twice in each cycle), the system must approach a unique steady state, even if only asymptotically.

If $\mathscr{S}_0$ is unique, we can consider the reduced M-node system comprising only the slipping nodes.

The tangential displacements at these nodes are equivalent to a time-invariant set of superposed nodal forces at the slipping nodes, which could be relaxed by corresponding time-invariant slips.

The reduced system has no permanently stuck nodes, so (ii) would then be sufficient to establish uniqueness.

We repeat (ii)
That for a system in which $\mathscr{S}_0$ *is null (i.e. all nodes slip at least twice in each cycle), the system must approach a unique steady state, even if only asymptotically.*

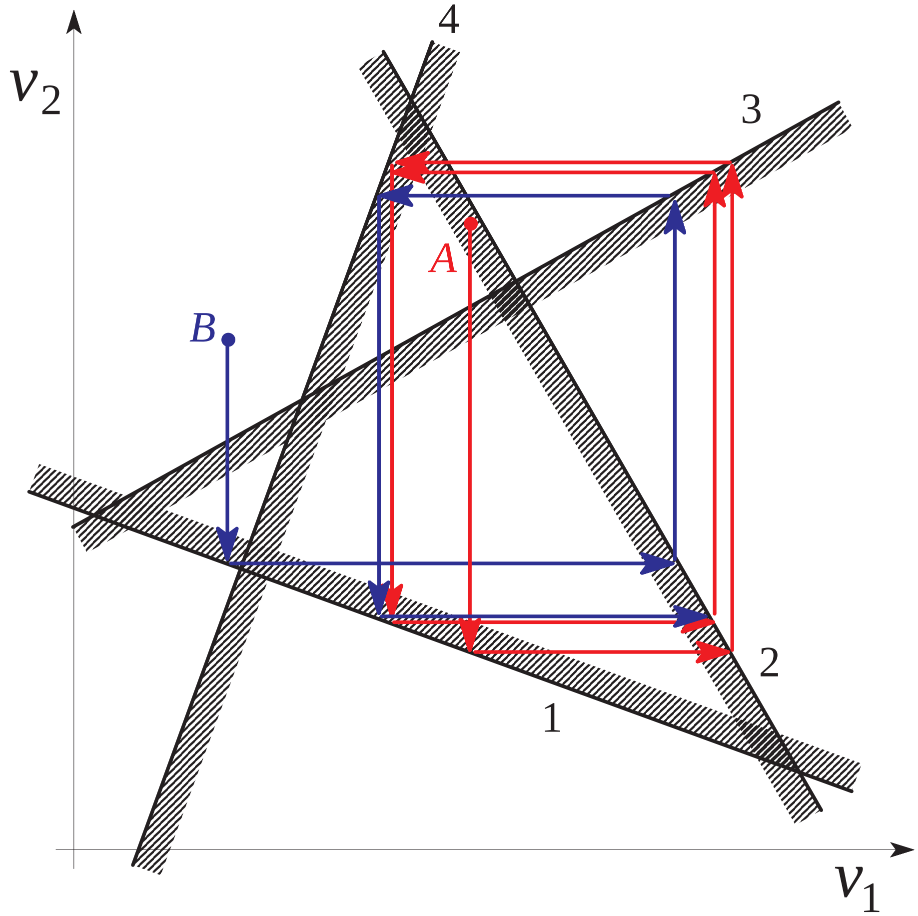
I believe this to be true even for coupled systems.

A heuristic argument is that the system memory is contained in the slip displacements at nodes that are instantaneously stuck.

No node is stuck all the time, so memory of the initial condition must be exchanged between nodes during each cycle.

This exchange process is likely to involve some loss of memory. [As educators, we should be very aware of this!]

This two node system is coupled (the constraints are not parallel) and both nodes slip during each cycle:-



Extreme positions are reached in the sequence
$1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1 \ldots$

With this scenario, we can solve for the unique steady state.

Each corner of the rectangle must lie on one of the constraints and shares one nodal displacement with the previous corner.

There are therefore four constraint equations and four ‘shared-displacement’ equations for eight unknowns — the coordinates of the four corners.

These equations are linear and hence have only one solution.

Each segment of slip approaches monotonically to a unique final state from whichever side we start.

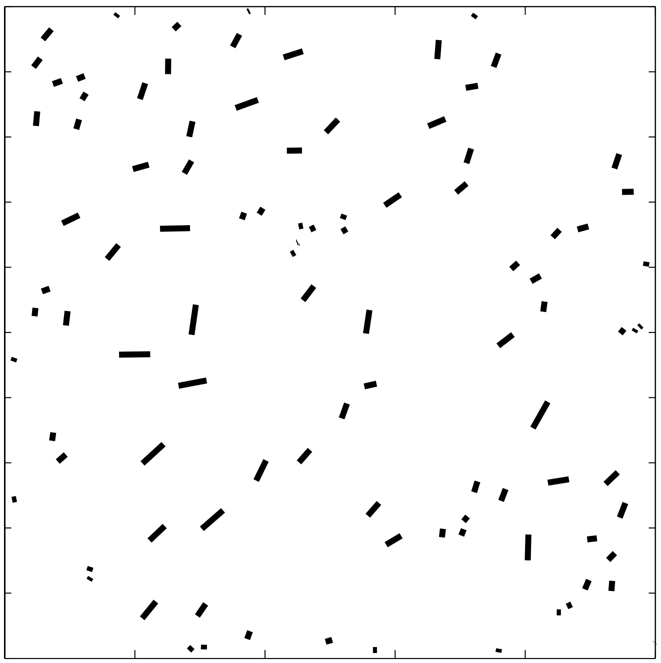
If there is any instant in the cycle when both nodes slip, the steady state is reached at that instant.

Can we prove this for general scenarios, or for the N-node system?

Can we define a norm representing somehow the difference between the instantaneous state and an unspecified steady state [at the same point in the loading cycle?]?

......

Application I: A random distribution of two-dimensional microcracks.



Suppose this body is subjected to nominally uniform cyclic in-plane stresses.

Under a uniform far-field stress, a single *isolated* plane crack must either:-

(i) open completely;

(ii) close completely and remain stuck everywhere;

(iii) close completely and slip everywhere.

There are no conditions under which partial closure or partial slip can occur.

A sufficiently sparse distribution of N plane cracks acts (mathematically) like a completely uncoupled discrete N-node frictional system.

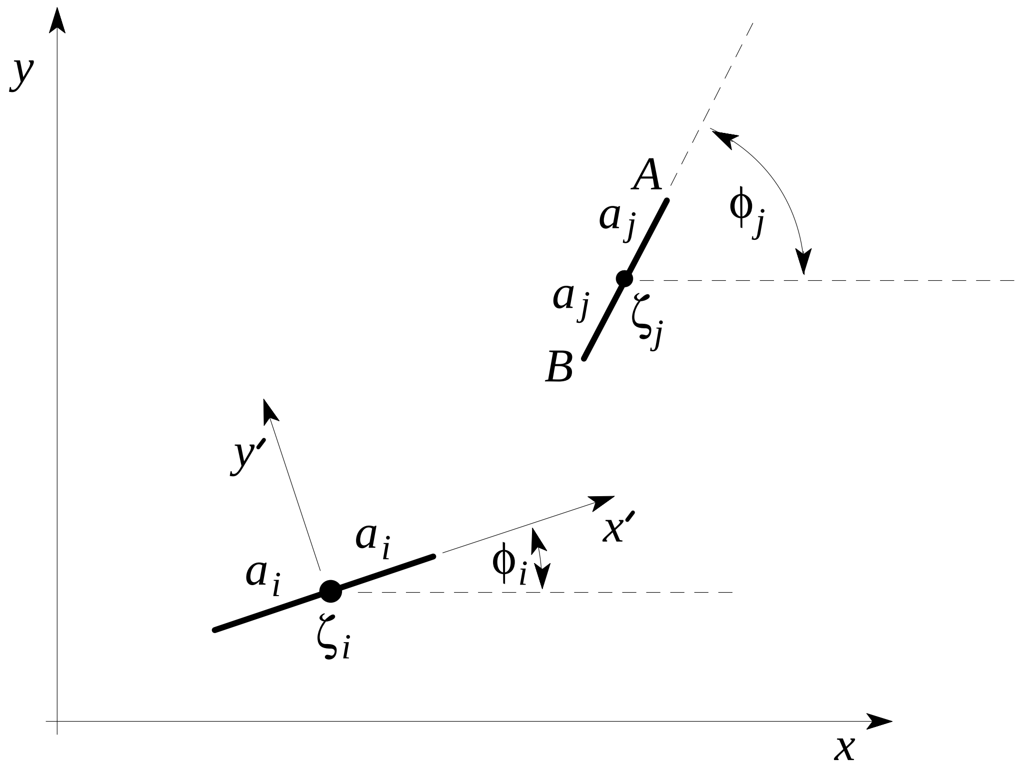
All the off-diagonal components of the stiffness matrix, including the whole matrix $\boldsymbol{B}$ are zero.

We can extend to moderate levels of crack interaction using Kachanov’s ‘simple’ theory.

The perturbation in the stress field at crack j due to opening or slip of crack i is approximated as a uniform local stress.

This uniform stress is taken as the average of the tractions on the line of crack j

The resulting discrete system is coupled ($\boldsymbol{B} \neq \mathbf{0}$).



This average traction is easily calculated using the complex-variable solution to the isolated crack problem (crack i) to determine the forces transmitted across the line of crack j.

We calculated the influence coefficients (stiffness matrix) for a random distribution of 100 cracks.

The far-field loading was assumed to be

$$\boldsymbol{S} = \boldsymbol{S}_0 + \lambda \boldsymbol{S}_1 \cos(\omega t) \; ,$$

where $\boldsymbol{S} \equiv \{\sigma_{xx}, \sigma_{xy}, \sigma_{yy}\}$ and

$$\boldsymbol{S}_0 = \{-1, 0, -1\} \; ; \quad \boldsymbol{S}_1 = \{0, 1, 0\} \; .$$

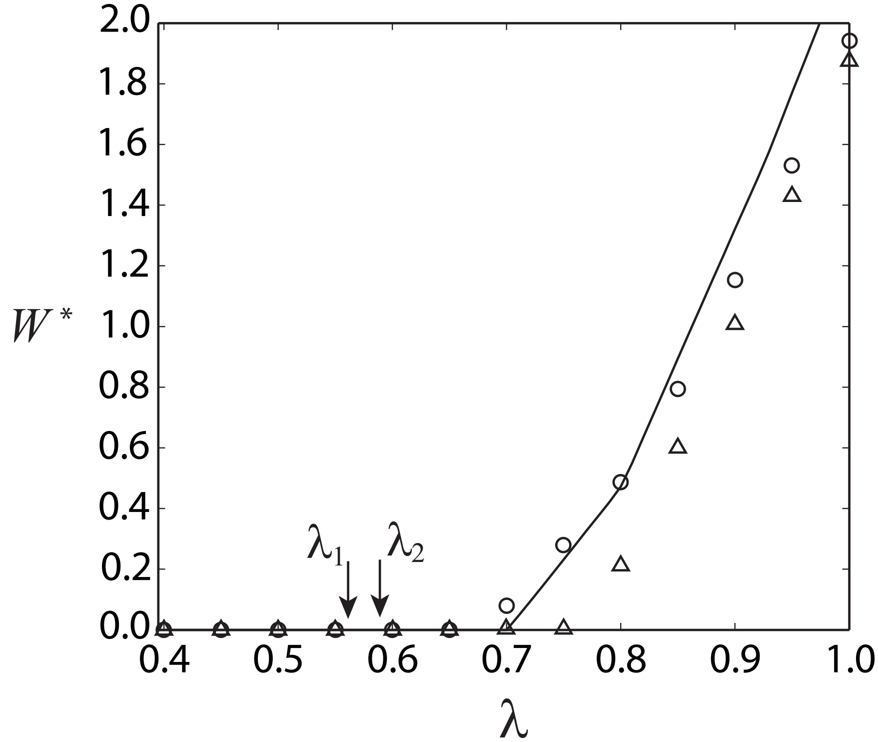
A constant hydrostatic compressive stress and a periodic shear stress.

We used Ahn's algorithm to determine the upper and lower bounds for the scalar load factor λ for shakedown:-

For $\lambda < \lambda_1$, shakedown occurs for all initial conditions.

For $\lambda > \lambda_2$ shakedown cannot occur [there is no safe shakedown state].

For $\lambda_1 < \lambda < \lambda_2$ shakedown depends on the intial conditions.



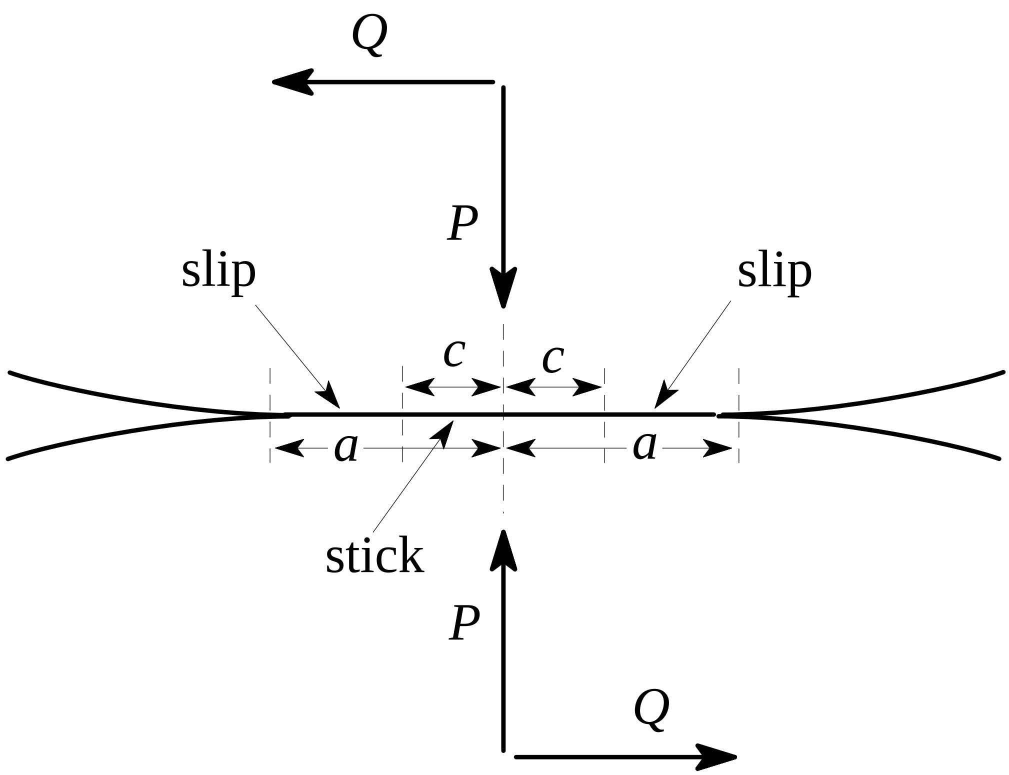
The solid line shows the (unique) frictional energy dissipation if coupling is neglected

The symbols △, ○ show the minimum and maximum dissipation respectively, depending on initial conditions.

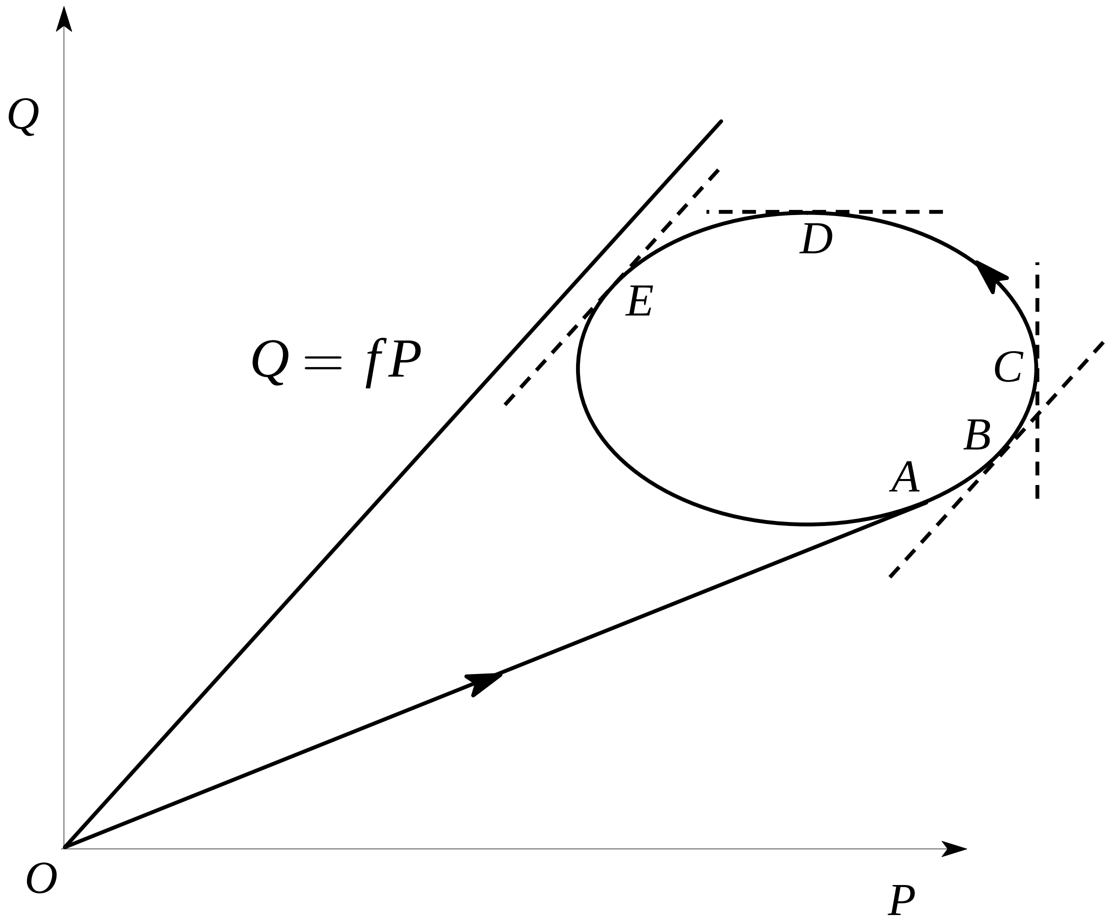
At sufficiently large λ, the dissipation becomes unique. Probably because all nodes (cracks) slip at some time during the cycle.

We confirmed this by recording the proportion of cracks that experience at least one period of slip and/or opening per cycle in the steady state.

Application II: Generalized-Hertzian contact: the effect of separation.



Two non-conforming bodies of similar materials are pressed together by a force P and sheared by a force Q, both of which can vary in an arbitrary but periodic way.



To avoid gross slip, we need $|Q| < fP$.

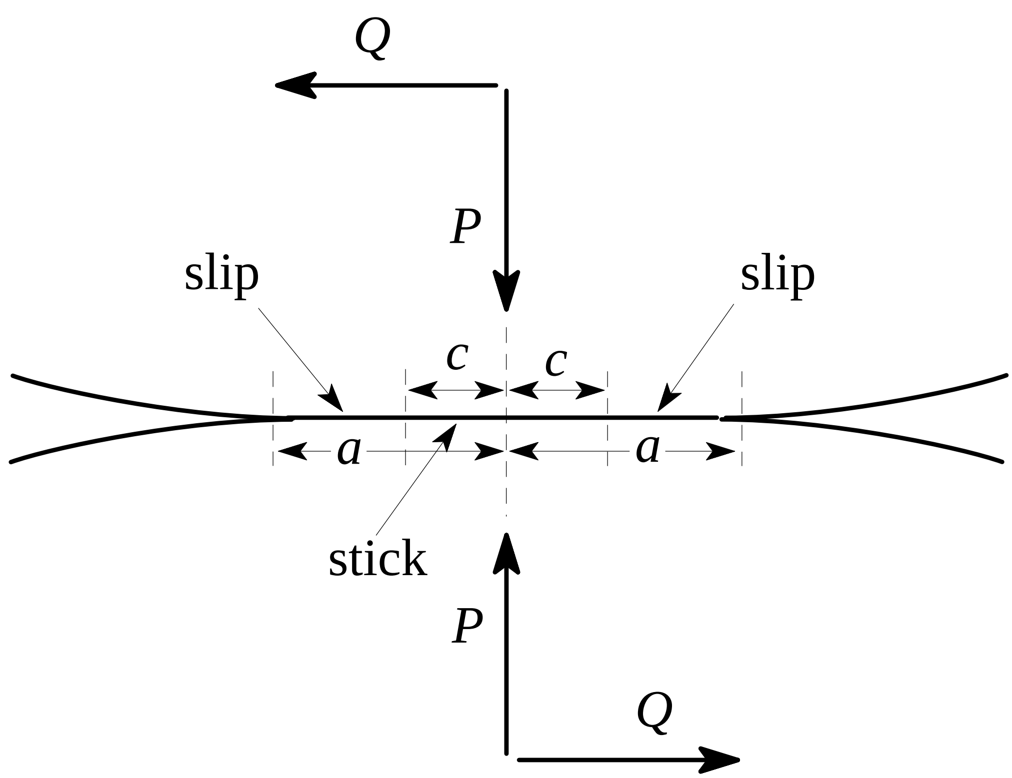
There will be an initial transient trajectory OA, the maximum P is reached at C and the maximum Q at D.

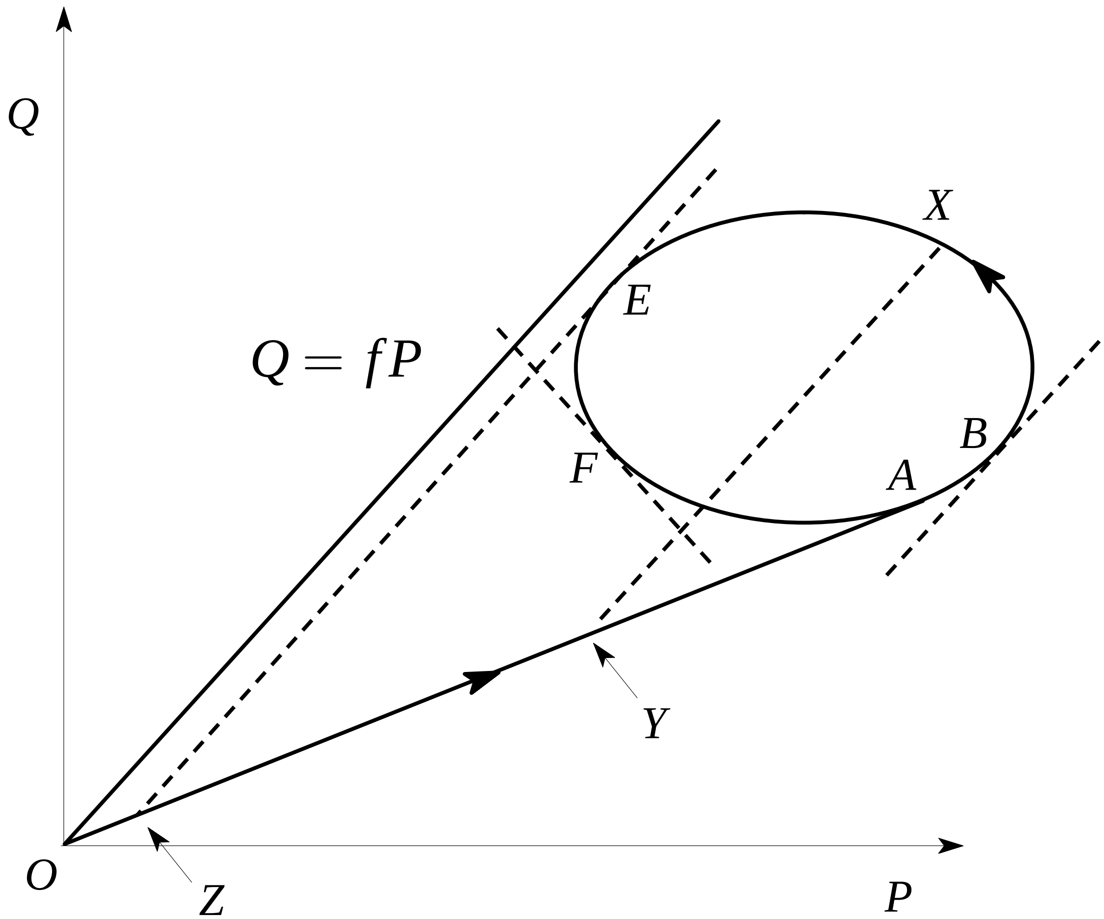
We can show that there will be no (incremental) slip as long as

$$\frac{dP}{dt} > 0 \qquad \text{and} \qquad \left|\frac{\partial Q}{\partial P}\right| < f \ ,$$

where f is the coefficient of friction.

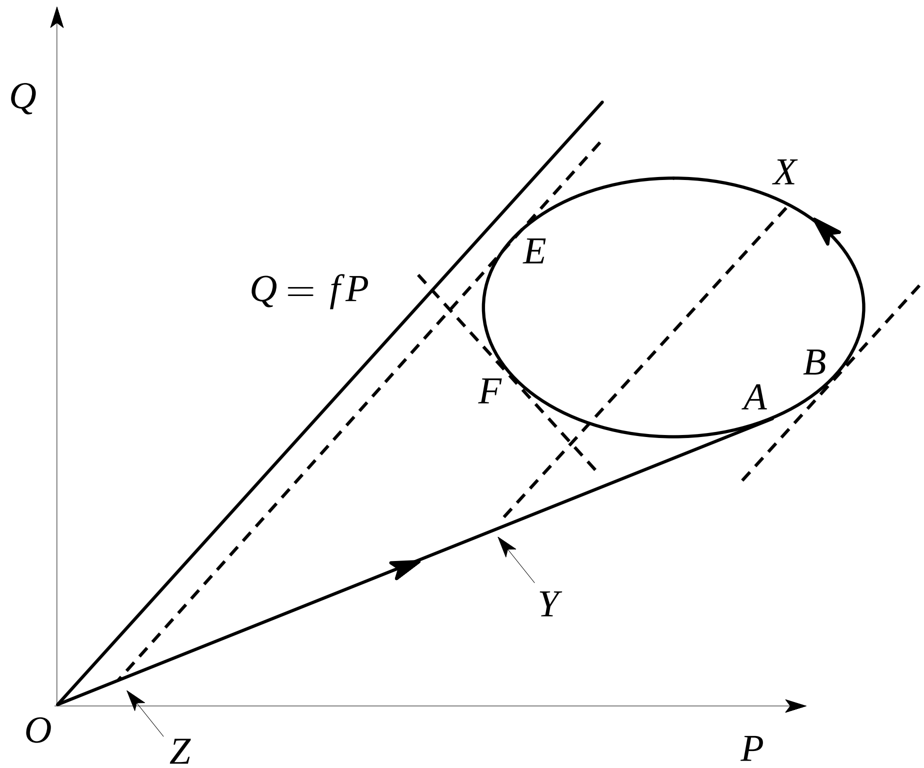
In all other cases there will be a central stick zone $-c < x < c$ and surrounding slip zones $c < |x| < a$.





The slip zone at X has the same extent as the whole contact region at Y [Ciavarella-Jäger theorem].

The memory of the transient phase between Z and A is erased by the time we reach E.



In the steady state, the contact area grows without slip between F and B.

Between B and E there is a growing slip zone.

AT E, this slip zone instantaneously sticks everywhere and a reverse slip zone grows between E and F, at which point again there is instantaneously full stick.

Conclusions

- Frictional systems exhibit seemingly endless complexities of behaviour despite the apparently simple constitutive description.

- Simple models are helpful to make sense of the resulting complex behaviour.

- Melan’s theorem applies only to systems with no coupling between slip displacements and normal reactions.

- The periodic steady state for coupled systems generally exhibits dependence on initial conditions.

- History-dependence in frictional problems depends on ‘system memory’ that resides in regions that are instantaneously stuck.

- Some memory is ‘lost’ when a node slips, so if all nodes slip at some time during each cycle, it is likely that all memory of the initial state will be lost, at least asymptotically, leading to a unique steady state, even when the system is coupled.

- If a subset of nodes remains stuck (the permanent stick zone), their locked-in slip displacements will influence the cyclic slip at other nodes *if and only if* the system is coupled.