8

FRICTION LAWS

The tangential tractions discussed in the previous chapter must be associated with some mechanism such as friction, unless the bodies are bonded [welded] together, in which case we are not strictly in the domain of contact mechanics. The investigation of the nature of such frictional tractions is more properly a subject for the related discipline of Tribology, though we shall touch on the subject in §8.6, and also in Chapter 16 where we investigate the contact of rough surfaces. In most of this book, we shall make use of the simple theory of friction associated with the names of Amontons and Coulomb, according to which the maximum tangential force $|\boldsymbol{Q}|$ that can be transmitted between two bodies at a plane interface is linearly proportional to the compressive normal force P and independent of the area of contact.

8.1 Amontons' law

We distinguish between states of stick and slip, such that during slip the magnitude of the tangential force is

$$|\boldsymbol{Q}| = fP \ , \tag{8.1}$$

where f is the *coefficient of friction*, and the direction of the force must be such as to oppose the slip velocity $\boldsymbol{V}_s$ as shown in Figure 8.1.

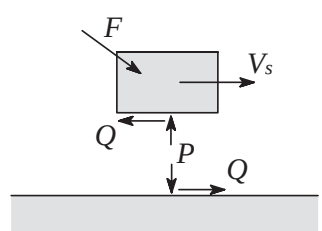


Figure 8.1: Frictional reactions induced during slip.

More generally, we have

$$\frac{\boldsymbol{Q}}{|\boldsymbol{Q}|} = \frac{\boldsymbol{V}_s}{|\boldsymbol{V}_s|} \,, \tag{8.2}$$

where the vector $\boldsymbol{Q}$ defines the tangential reaction on the *lower* body. During periods of stick,

$$\boldsymbol{V}_s = \boldsymbol{0} \quad \text{and} \quad |\boldsymbol{Q}| \leq fP \,. \tag{8.3}$$

We shall generally assume that the same coefficient of friction f applies in both equations (8.3, 8.1), in other words that the so-called static and dynamic coefficients of friction are equal. Some of the implications of differing coefficients are discussed in §8.6.3.

Numerous experimental investigations have documented significant deviations from Amontons' law, particularly at very small length scales, but it remains a remarkably good approximation to frictional behaviour at the macroscale and its mathematical simplicity recommends it for use in many engineering applications, if only as an approximation.

8.1.1 Continuum problems

The statement of Amontons' law in equations (8.1–8.3) is appropriate for the contact of rigid bodies, and it can clearly be extended to systems of rigid bodies and springs, and hence by extension to finite element models of contact, where the friction conditions will be imposed at a discrete set of contact nodes, some of which may stick and some slip. However, in problems for the continuum, we must restate the law in terms of contact tractions [forces per unit area] obtaining

$$|\boldsymbol{q}(x,y)| = fp(x,y) \quad \text{and} \quad \frac{\boldsymbol{q}(x,y)}{|\boldsymbol{q}(x,y)|} = \frac{\dot{\boldsymbol{h}}(x,y)}{|\dot{\boldsymbol{h}}(x,y)|} \tag{8.4}$$

in slip regions, and

$$|\boldsymbol{q}(x,y)| \leq fp(x,y) \quad \text{and} \quad \dot{\boldsymbol{h}}(x,y) = 0 \tag{8.5}$$

in stick regions, where $\dot{\boldsymbol{h}}$ is defined in equation (7.2) and implies the sign convention of Figure 7.1, and $\boldsymbol{q}(x,y)$ is the tangential traction acting on the *lower* body 1.

8.1.2 Two-dimensional problems

In three-dimensional problems, the slip velocity $\dot{\boldsymbol{h}}(x,y)$ and the frictional traction $\boldsymbol{q}(x,y)$ are vectors within the plane, but in two-dimensional problems, we can only distinguish between *forward slip* [$\dot{h}(x) > 0$] and *backward slip* [$\dot{h}(x) < 0$]. The friction law for the continuum can then be stated as $p(x) \geq 0$ for all x and

$$\begin{array}{llll} \dot{h}(x) = 0 & g(x) = 0 & |q_x(x)| \leq fp(x) & \text{stick} \\ q_x(x) = fp(x) & g(x) = 0 & \dot{h}(x) > 0 & \text{forward slip} \\ q_x(x) = -fp(x) & g(x) = 0 & \dot{h}(x) < 0 & \text{backward slip} \\ q_x(x) = 0 & p(x) = 0 & g(x) > 0 & \text{separation,} \end{array} \tag{8.6}$$

where we have supplemented the frictional conditions by the normal contact conditions (1.3). Notice that every point on the interfacial plane must be in one of four states, separation, stick, forward slip or backward slip, and if the state is known for all such points, we then have two linear equations for each x for the two unknowns $p(x),\ q_x(x)$. We also note that the equations are supplemented by inequalities which, as in frictionless contact, serve to determine the partition of the plane into regions of each state. By contrast, the three-dimensional condition (8.4) is pointwise non-linear, even if the partition of the interfacial plane into zones in different states is known.

8.1.3 Existence and uniqueness theorems

Equations (8.6) show that frictional contact problems have some similarities with the simpler frictionless problem of equations (1.3) and numerous investigators have endeavoured to establish theorems of existence and uniqueness. In this context, it is important to distinguish between the discrete problem, where there exists a denumerable set of discrete contact points, and the continuum problem discussed above, where contact may occur at a non-denumerable set of real points. In the engineering literature, finite element methods are often used to reduce the continuum problem to a discrete problem, where the contact and frictional conditions are imposed at a finite set of contact nodes. Underlying such investigations lies the assumption that with a sufficiently fine mesh, the solution will in some sense approach the continuum solution asymptotically. However, from a mathematical perspective, the two problems are quite different and indeed considerably more progress has been made in the investigation of the discrete problem.

The problem defined by (8.4) or (8.6) is history dependent, because it contains the time derivative $\dot{\boldsymbol{h}}$. In the mathematical literature, this is described as the *rate problem* and is contrasted with the *static problem* in which conditions are imposed on the instantaneous values of $\boldsymbol{h}$ rather than $\dot{\boldsymbol{h}}$. From an engineering perspective, the static problem formulation is meaningful if and only if the loading is proportional and monotonic and all contact points were already in contact at zero force, in which case $\dot{\boldsymbol{h}}$ can legitimately be integrated in time.

Difficulties are encountered with both existence and uniqueness proofs for the general quasi-static elastic contact problem with Coulomb friction[1]. Some of these difficulties can be avoided by 'regularizing' the friction law [i.e. smoothing the discontinuities associated with the inequalities] or using a non-local friction law, but uniqueness can still generally only be established subject to the proviso that the coefficient of friction f should be 'sufficiently small' (Oden and Pires, 1983). In other words, there may exist a system-specific critical coefficient f_0 such that for $f > f_0$ multiple solutions or non-existence of solution may occur. This was a serious embarrassment to the developers of finite element codes, since these are required to be 'robust' and capable of use by practitioners lacking a detailed knowledge of the underlying algorithms.

[1] See for example Kikuchi and Oden, (1988).

8.2 The Klarbring model

Insight into the behaviour of frictional systems was considerably advanced by Klarbring's investigation of the simple two-dimensional single node system (Klarbring, 1990) shown in Figure 8.2 (a).

A point mass M makes frictional contact with a rigid plane and is also supported by a generalized linear spring $\boldsymbol{K}$. The stiffness matrix $\boldsymbol{K}$ must be positive definite, so $k_{11}, k_{22} > 0$, and we can choose the direction of x_1 so as to ensure that $k_{12} = k_{21} > 0$ without loss of generality. A time-varying force $\boldsymbol{F}(t) = \{F_1, F_2\}$ is applied to the mass and the resulting motion is described by a displacement vector $\boldsymbol{u}(t) = \{v, w\}$.

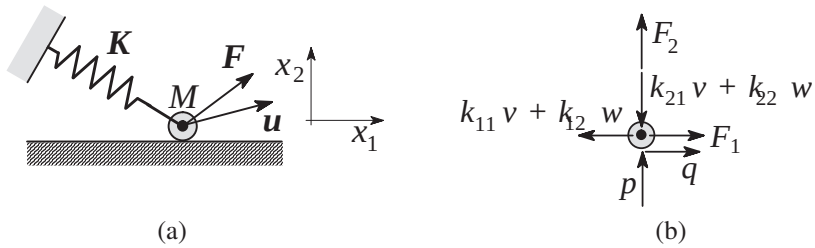


Figure 8.2: (a) The Klarbring model, (b) Free-body diagram.

Figure 8.2 (b) shows a free-body diagram of the mass and in particular identifies the normal and tangential reactions p, q respectively at the contact interface. Notice that the positive direction of the tangential reaction q is opposite to that shown in Figure 8.1, since here the focus is on the equilibrium of the upper body, and it is convenient to use the same coordinate system for $\{q, p\}$ and $\{v, w\}$.

We assume that the force is applied sufficiently slowly for the response to be *quasi-static*. In other words, the system passes through a sequence of equilibrium states, so at all times t,

$$F_1 + q - k_{11}v - k_{12}w = 0\,; \quad F_2 + p - k_{21}v - k_{22}w = 0\,. \tag{8.7}$$

For this single node system, the discrete equivalents of conditions (8.6) are $p \geq 0$ and

$$\begin{array}{llll} \dot{v} = 0 & w = 0 & |q| \leq fp & \text{stick} \\ q = -fp & w = 0 & \dot{v} \;\; > 0 & \text{forward slip} \\ q = fp & w = 0 & \dot{v} \;\; < 0 & \text{backward slip} \\ q = 0 & p = 0 & w \;\; > 0 & \text{separation.} \end{array} \tag{8.8}$$

Consider the case where the system is initially unloaded with $v = w = 0$ and the external force $\boldsymbol{F}(t) = \boldsymbol{C}t$, where $\boldsymbol{C} = \{C_1, C_2\}$ is a constant vector. In other words, the two force components F_1, F_2 are increased monotonically and in proportion. Using (8.7, 8.8) to determine p, q, v, w, we then find that the *inequalities* governing the four states are

$$
\begin{array}{lll}
fC_2 < C_1 < -fC_2 & & \text{stick} \\
\dfrac{C_1 + fC_2}{k_{11} + fk_{21}} > 0\,; & \dfrac{k_{11}C_2 - k_{21}C_1}{k_{11} + fk_{21}} < 0 & \text{forward slip} \\
\dfrac{C_1 - fC_2}{k_{11} - fk_{21}} < 0\,; & \dfrac{k_{11}C_2 - k_{21}C_1}{k_{11} - fk_{21}} < 0 & \text{backward slip} \\
k_{21}C_1 < k_{11}C_2 & & \text{separation.}
\end{array}
\tag{8.9}
$$

The assumed loading scenario corresponds to a straight line starting from the origin in $\{F_1, F_2\}$-space and Figure 8.3 shows the states predicted by (8.9) for a case where $k_{21} > 0$ and $f < k_{11}/k_{21}$. In particular, we see that one and only one state is possible for each set of values $\{C_1, C_2\}$ and that the same state will persist as long as the loading trajectory is defined by the straight line $\boldsymbol{F}(t) = \boldsymbol{C}t$.

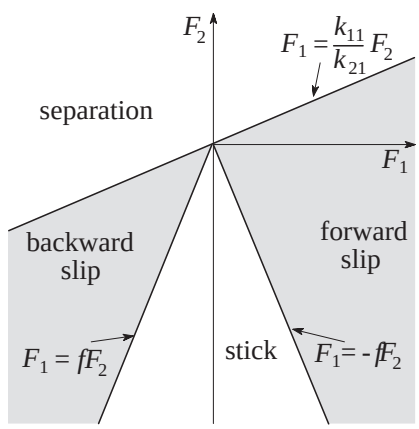


Figure 8.3: Domains of stick, slip and separation for $f < k_{11}/k_{21}$ and $\boldsymbol{F}(t) = \boldsymbol{C}t$.

8.2.1 General loading scenarios

Figure 8.3 can be modified to allow general loading scenarios. The separation segment remains unchanged, but the stick cone is moved by slip displacement v such that its vertex always lies on the separation line $F_1 = (k_{11}/k_{21})F_2$.

During periods of slip, as the instantaneous point $P(F_1, F_2)$ moves to Q in Figure 8.4 (a), it 'pushes' the vertex of the stick cone along this line. Separation, when it occurs, is always reached through the vertex of the cone and when a transition from separation to contact occurs, it always occurs through the vertex, as shown in Figure 8.4 (b). The state realized immediately after contact depends on the direction of the local force trajectory and hence on the ratio $\dot{F}_1/\dot{F}_2$. The case illustrated inFigure 8.4 (b) will transition from separation to forward slip.

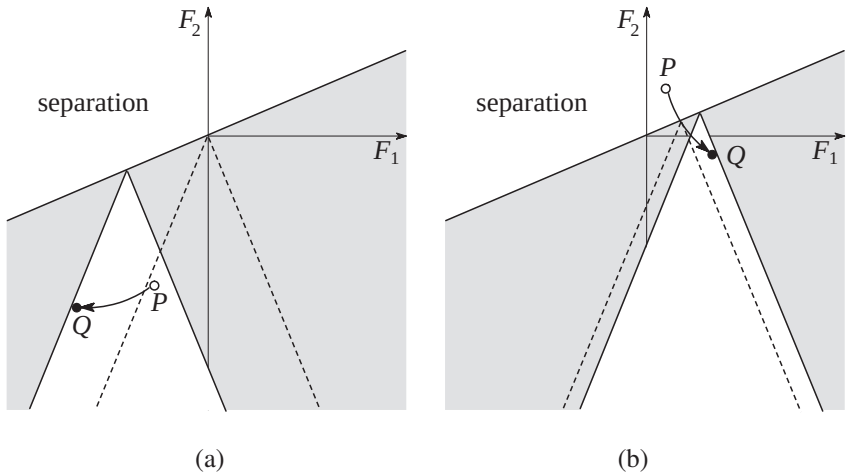


Figure 8.4: Movement of the stick cone.

8.2.2 The critical coefficient of friction

The arguments in the last section essentially establish that the rate problem for Klarbring's model has a unique solution for all possible loading scenarios. However, Figures 8.3, 8.4 are drawn for the case $f < k_{11}/k_{21}$. If instead the coefficient of friction $f > k_{11}/k_{21}$, the diagram has the form of Figure 8.5, where we see a domain in which stick, backward slip and separation are all possible. Thus we lose uniqueness of solution.

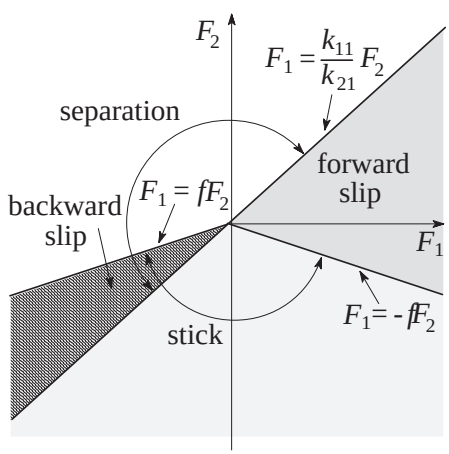


Figure 8.5: Domains of stick, slip and separation for $f > k_{11}/k_{21}$.

More seriously, if we devise a loading scenario that starts in the stick domain and then moves through this multi-state region, the stick inequalities are satisfied in the

strict sense so stick will persist until we reach the upper boundary beyond which only separation is possible. However, the separation state so reached involves a non-zero value of w, so the system seems to make a discontinuous jump, which violates the quasi-static assumption.

A scenario involving discontinuities in displacement was also predicted by Martins and co-workers[2] for the case where damping is included in the system and then allowed to approach zero asymptotically. These authors also note that existence and uniqueness theorems can be established with arbitrary coefficient of friction if the requirement of continuity of displacement is relaxed.

Cho and Barber (1998) showed that if the quasi-static assumption is relaxed — i.e. if the inertia term $M\ddot{\boldsymbol{u}}$ is included in equations (8.7), the discontinuous jump is resolved into a dynamic event on the time scale of the local dynamics, and the backward slip state in the multiple solution range is unstable, tending either to separation or stick depending on the exact initial conditions. However, in more complex discrete systems such as finite element models, the replacement of quasi-static by dynamic behaviour involves a major increase in computational effort. In simple systems such as the Klarbring model, one can devise modified quasi-static algorithms that capture the long time behaviour of the system without resolving the dynamics of the occasional jumps. However, to the authors' best knowledge, no such algorithm is yet available for complex systems above the critical friction coefficient.

8.2.3 Wedging

Frictional systems are capable of becoming 'wedged' if the coefficient of friction is sufficiently high. We define a wedged state as one in which the system can support internal forces even when all external forces are set to zero. Wedging is important in automated assembly systems, where it might result in an incorrect assembly.

For the Klarbring model, if a wedged state is accessible, it implies that we can devise a loading scenario $\boldsymbol{F}(t)$ in $0<t<t_0$ such that $\boldsymbol{F}(0)=0$, $\boldsymbol{F}(t_0)=0$, $w(t_0)=0$ and $v(t_0)=v_0 \neq 0$. In other words, the system is left in a state of stick with non-zero tractions

$$q(t_0) = -k_{11}v_0 \; ; \quad p(t_0) = k_{21}v_0 \; , \tag{8.10}$$

from (8.7). The stick inequalities then require that $k_{21}v_0 > 0$ and

$$k_{11} < f|k_{21}| \quad \text{or} \quad f > \frac{k_{11}}{|k_{21}|} \; . \tag{8.11}$$

Notice that in this case, the critical coefficient of friction for wedging coincides with that at which the existence and uniqueness theorems for the rate problem break down. However, no general proof exists of this equivalency for more complex discrete systems, and indeed counter-examples can be devised.

[2] Martins and Oden (1987), Martins *et al.* (1992, 1994).

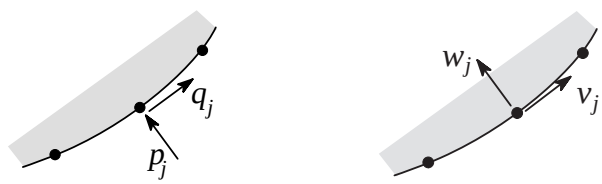
8.3 Multinode systems

These arguments can be generalized by considering a discrete two-dimensional system with N contact nodes $i \in (1, N)$. We can then define N-vectors with components p_i, q_i, w_i, v_i and if the system is linearly elastic, these quantities must be linearly related through the equations

$$\left\{ \begin{array}{c} q_i \\ p_i \end{array} \right\} = \left\{ \begin{array}{c} q_i^w \\ p_i^w \end{array} \right\} + \begin{bmatrix} A_{ij} & B_{ij} \\ B_{ji} & C_{ij} \end{bmatrix} \left\{ \begin{array}{c} v_j \\ w_j \end{array} \right\} , \tag{8.12}$$

where p_i^w, q_i^w are the contact tractions that would be produced by the time-varying external forces if the nodes were all welded in contact with $v_i = w_i = 0$.

It is convenient to think of such a system as a generalization to N nodes of Figure 8.2 (a), but a mathematically equivalent system can be generated by discretizing a continuum elastic contact problem by the finite element method, provided we restrict attention to small strains. In this case, nodal forces are non-zero only at the contact nodes and the externally loaded nodes, so the degrees of freedom at internal nodes can be eliminated, reducing the system to the form of equation (8.12). The sign convention for contact nodal forces and displacements is then as defined in Figure 8.6.



(a) (b)

Figure 8.6: Sign convention for (a) nodal forces and (b) nodal displacements in a finite element discretization, leading to equation (8.12).

Notice that

$$\boldsymbol{K} = \begin{bmatrix} A_{ij} & B_{ij} \\ B_{ji} & C_{ij} \end{bmatrix} \tag{8.13}$$

is the reduced stiffness matrix for the system and hence must be symmetric and positive definite. It follows that $\boldsymbol{A}$ and $\boldsymbol{C}$ are also symmetric and positive definite, but no such restriction applies to the matrix $\boldsymbol{B}$ which defines the coupling between normal and tangential effects. This coupling matrix is the discrete analogue of the coupling terms with multiplier β in the continuum problems of Chapter 7, and as in that case, it has a profound effect on the qualitative behaviour of the frictional system.

8.3.1 The evolution and rate problems

If the external forces $q_i^w(t)$, $p_i^w(t)$ are known functions of time, we wish to be able to determine the displacements $v_j(t)$, $w_j(t)$ such that equations (8.12) and conditions

(8.8) are satisfied at all nodes i and all times t. We shall refer to this as the *evolution problem*. The presence of time-derivatives in conditions (8.8) implies that the evolution problem can be solved if and only if a corresponding *rate problem* can be solved. In two dimensions the rate problem can be stated as:

Given a set of nodal forces and displacements that satisfy equations (8.12) and the conditions

$$|q_i| \leq f p_i \; ; \quad w_i \geq 0 \; ; \quad p_i \geq 0 \; ; \quad w_i p_i = 0 \tag{8.14}$$

at each node i, and given the loading rates $\dot{q}_i^w$, $\dot{p}_i^w$, determine the time derivatives $\dot{q}_i, \dot{p}_i, \dot{v}_i, \dot{w}_i$ such that each node satisfies the conditions for one and only one of the four states

$$\begin{array}{llll} \dot{v}_i = 0 & \dot{w}_i = 0 & |q_i| \leq f p_i & \text{stick} \\ \dot{q}_i = f\dot{p}_i & \dot{w}_i = 0 & \dot{v}_i \;\; > 0 & \text{forward slip} \\ \dot{q}_i = -f\dot{p}_i & \dot{w}_i = 0 & \dot{v}_i \;\; < 0 & \text{backward slip} \\ \dot{q}_i = 0 & \dot{p}_i = 0 & w_i \;\; > 0 & \text{separation.} \end{array} \tag{8.15}$$

Notice that at any node i where either of the strict inequalities $|q_i| < f p_i$ or $w_i > 0$ holds, the corresponding state (stick or separation respectively) must continue into the solution of the rate problem. However, nodes where $|q_i| = f p_i$ may evolve into either stick or slip [in the appropriate direction], and nodes where both $w_i = 0$ and $p_i = 0$ may evolve into any of the four states, depending on the external loading rate.

8.3.2 Algorithms for two dimensional problems with time-varying forces

A simple numerical algorithm for the evolution problem is to assume provisionally that the states holding at each node immediately prior to time t continue to hold through a finite time increment Δt. We then use (8.12) and the equations in (8.8) to determine the new nodal forces and displacements, and compare the resulting values and implied rates against the respective inequalities. If any inequality is violated, the nature of the violation is a clue to the correct state at the node in question. For example, if we assume node i is in forward slip and we detect that $\Delta v_i < 0$ during Δt, then the most likely alternative state is one of stick. Ahn and Barber (2008) developed an algorithm of this kind in which the nodes are solved for sequentially in a Gauss-Seidel sense.

An alternative approach is to recognize that as long as the states at all the nodes remain the same, the evolution of the discrete solution is defined by linear equations. If the external forces can also be described as piecewise linear functions of time, this opens up the possibility of a linear complementarity solution. Briefly, we assume that the states remain the same at all nodes in which case the equality conditions are linear and we can solve for the nodal forces and displacements as linear functions of time. We then substitute the resulting expressions into the inequalities and solve for the earliest time at which a violation occurs at any node. We then jump to this time and 'pivot' about it starting a new time step with an appropriate change of state at the pivoted node. This method was first applied to the friction problem by Klarbring & Björkman (1988) and an improved algorithm is given by Bertocchi (2009).

8.3.3 History-dependence and memory

The necessity of a rate formulation shows that the instantaneous state of a system involving friction can depend on the history of loading as well as on the instantaneous forces. Thus, frictional systems fall into the broader class of systems with 'memory'. It is clear from conditions (8.8) that the system memory is associated with the fact that at a 'stick' node, only the velocity $\dot{v}_i$ is determined from the instantaneous conditions, the actual position of the node v_i depending on the previous history. Thus it is reasonable to consider the memory of the system as residing in those nodes that are instantaneously stuck, and as being defined in some sense by the displacements v_i at those nodes.

Some insight into this question and into the frictional evolution problem in general can be gained by considering the motion of the point $P(v_1, v_2, ...v_N)$ in the N-dimensional hyperspace of slip displacements. Here we shall restrict attention to the simpler case where the external loading ensures that all the nodes remain in contact at all times [$\boldsymbol{w}(t) = \mathbf{0}$]. Equation (8.12) and the frictional inequality $|q_i| \leq f p_i$ at node i then imply that

$$-f p_i^w - f B_{ji} v_j \leq q_i^w + A_{ji} v_j \leq f p_i^w + f B_{ji} v_j \ , \tag{8.16}$$

or

$$(A_{ji} - f B_{ji}) v_j \leq f p_i^w - q_i^w \quad \text{and} \quad (A_{ji} + f B_{ji}) v_j \geq -f p_i^w - q_i^w \ , \tag{8.17}$$

where summation is implied over repeated indices. Each of these inequalities [two for each node] excludes the region on one side of a hyperplane in $\boldsymbol{v}$-space and when the node in question is slipping, P lies *on* this hyperplane [corresponding to an appropriate equality in (8.17)]. The external forces $\boldsymbol{p}^w(t)$, $\boldsymbol{q}^w(t)$ appear in the right-hand sides of these inequalities and hence each hyperplane moves in time whilst preserving the same inclination, which is determined by the elements of the matrices $\boldsymbol{A}$ and $\pm f \boldsymbol{B}$. Thus, the hyperplanes 'push' the point P about the space, the local motion being determined by the appropriate direction of slip corresponding to the active inequalities (Ahn *et al.*, 2008).

Figure 8.7 illustrates this process for a simple two-node system, where there are four inequality constraints I, II, III, IV governing forward and backward slip at each of the two nodes. For example, during backward slip at node 1 [$\dot{v}_1 < 0$], P must lie on I, whereas during forward slip at node 2 [$\dot{v}_2 > 0$], it must lie on IV. The region excluded by each hyperplane is shown shaded in the figure, so at any instant P must lie inside the unshaded region. However, the instantaneous position within this region depends on the history of loading. If [for example] the changes in $\boldsymbol{p}^w(t)$, $\boldsymbol{q}^w(t)$ cause IV to move upwards to the point P_3, after which I moves to the left, the point P would follow the trajectory shown.

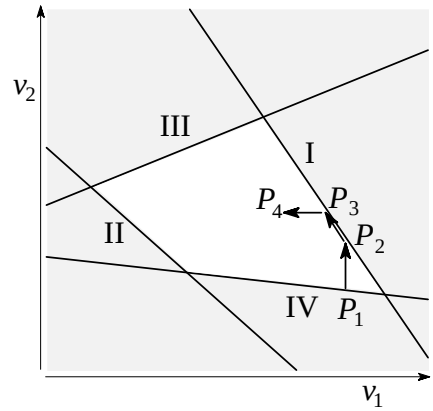


Figure 8.7: Motion of the instantaneous state $P(v_1, v_2)$ due to changes in external loading.

8.3.4 Klarbring's P-matrix criterion

Under some circumstances, the point P can become 'trapped' between two or more constraints, so that no further evolution satisfying the friction law can occur. Figure 8.8 shows such a case, where the constraints I and IV govern slip in directions $\dot{v}_1 < 0$ and $\dot{v}_2 > 0$ respectively. As the constraints move, both directions of motion are separately possible, but if P reaches the intersection between the constraints as shown, further reduction in the permissible region leads to a contradiction that cannot be resolved by a legitimate motion of P.

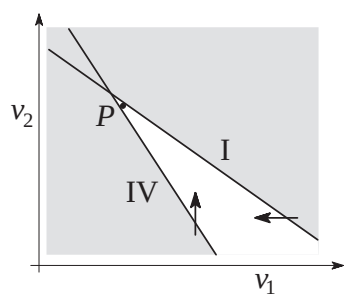


Figure 8.8: P becomes trapped between constraints I and IV.

Similar problems can arise at all edges and corners of the instantaneously admissible hypervolume in the N-node case. At a corner where M hyperplanes intersect, the condition for a situation analogous to that in Figure 8.8 to arise can be stated in terms of a subset of M of the $2N$ constraints (8.17), involving only M of the slip displacements v_i, and with appropriate signs for f at each node to reflect the direction of slip that is prevented.

The resulting criterion was enunciated by Klarbring (1999), who showed that for two-dimensional problems, the rate problem is well posed [i.e. situations analogous to that in Figure 8.8 cannot arise], if and only if all matrices of the form $(\boldsymbol{A}+\boldsymbol{\Lambda}\boldsymbol{B})$ are P-matrices [i.e. all principal minors of the matrix are positive definite], where $\boldsymbol{\Lambda}$ is any diagonal matrix each of whose diagonal elements[3] is either $+f$ or $-f$.

Since $\boldsymbol{A}$ is a positive-definite stiffness matrix, this condition is clearly satisfied for $f=0$, and for given matrices $\boldsymbol{A}, \boldsymbol{B}$ [with $\boldsymbol{B}$ non-null] there therefore exists a critical coefficient of friction $f_{\rm cr}$ above which the P-matrix criterion fails and the rate problem is ill-posed. In the special case where $N=1$, each of the matrices reduces to a single element $\boldsymbol{A}=k_{11}$, $\boldsymbol{B}=\boldsymbol{B}^T=k_{12}$, $\boldsymbol{C}=k_{22}$, and we recover $f_{\rm cr}=A_{11}/|B_{11}|=k_{11}/|k_{12}|$, as in §8.2.2.

For $N=2$, Ahn (2010) has shown that, as in the Klarbring model, a dynamic solution with damping for $f>f_{\rm cr}$ approximates to a unique quasi-static solution with displacement jumps. The present author is unaware of any proof of this result for $N>2$.

8.4 Periodic loading

Many engineering applications involve periodic [cyclic] loading, due to vibration or to repetitive operations, and the energy dissipated in friction under these circumstances can be an important measure of system performance. It has been estimated that frictional hysteresis in assembled structures accounts for more energy dissipation than internal material damping, but the effective damping in such cases is notoriously difficult to quantify. Also, the energy dissipated in microslip regions produces a favourable environment for the initiation of fretting fatigue cracks (Lovrich and Neu, 2006, Nowell *et al.*, 2006.), which are a primary mode of failure in many nominally static contacting systems subject to vibration, including notably the blade root contact in aero engines (Murthy *et al.*, 2004).

We anticipate that the system will eventually reach a steady periodic state under cyclic loading, but since frictional systems are history-dependent, we might reasonably ask whether this steady state is a unique function of the loading, or whether it also depends on the initial conditions, or on an initial period of transient loading leading to the steady state. For example, for a bolted joint, there might be a dependence on the bolt-tightening protocol. We shall prove here that important features of the steady state are indeed unique in uncoupled systems [$\boldsymbol{B}=\boldsymbol{0}$], but that in coupled systems we must anticipate a dependence on initial conditions.

8.4.1 A uniqueness proof for uncoupled systems

If the problem is uncoupled, the complementarity conditions $w_i \geq 0,\ p_i \geq 0,\ w_i p_i = 0$ determining the normal contact reactions are independent of the frictional problem

[3] Klarbring allows the possibility that the coefficient of friction f may be node-dependent, in which case the elements of $\boldsymbol{\Lambda}$ must be taken as $\pm f_i$.

and hence are not history-dependent. Thus, the normal forces are unique functions of the instantaneous external forces $\boldsymbol{p}^w(t)$.

Suppose that there exist two time-varying solutions $\boldsymbol{q}_1(t), \boldsymbol{v}_1(t)$ and $\boldsymbol{q}_2(t), \boldsymbol{v}_2(t)$, corresponding to the same external loading scenario $\boldsymbol{p}^w(t)$, $\boldsymbol{q}^w(t)$, but different initial conditions $\boldsymbol{v}_1(0), \boldsymbol{v}_2(0)$ respectively. The following argument applies to both two and three-dimensional problems. In the latter case the tangential nodal tractions $\boldsymbol{q}_i$ and displacments $\boldsymbol{v}_i$ are two-dimensional vectors, whereas in two-dimensional problems, they are scalars.

Following Klarbring *et al.* (2007), we define an 'energy norm'

$$\mathcal{E} = \frac{1}{2}(\boldsymbol{v}_1 - \boldsymbol{v}_2)^T \cdot \boldsymbol{A}(\boldsymbol{v}_1 - \boldsymbol{v}_2) \ , \tag{8.18}$$

where the stiffness matrix $\boldsymbol{A}$ is defined in (8.12). The norm $\mathcal{E}$ can be regarded as a scalar measure of the difference between the two solutions. Differentiating with respect to time, we obtain

$$\dot{\mathcal{E}} = (\dot{\boldsymbol{v}}_1 - \dot{\boldsymbol{v}}_2)^T \cdot \boldsymbol{A}(\boldsymbol{v}_1 - \boldsymbol{v}_2) = (\dot{\boldsymbol{v}}_1 - \dot{\boldsymbol{v}}_2)^T \cdot (\boldsymbol{q}_1 - \boldsymbol{q}_2) \ , \tag{8.19}$$

using (8.12) with $\boldsymbol{B} = \boldsymbol{0}$ and recalling that $\boldsymbol{q}^w(t)$ is the same for both solutions. Now if a given node i is stuck at time t in both solutions, $\dot{\boldsymbol{v}}_i^1 = \dot{\boldsymbol{v}}_i^2 = 0$, and the contribution of that node to $\dot{\mathcal{E}} = 0$ is zero. In the two-dimensional case[4], if the given node is slipping in the same direction, or separated in both solutions, $\boldsymbol{q}_i^1 = \boldsymbol{q}_i^2$ and we can therefore conclude that $\dot{\mathcal{E}} = 0$ during any period in which the nodes in the two solutions are in the same state.

Now consider the case where node i is stuck at time t in solution 2 ($\dot{\boldsymbol{v}}_i^2 = \boldsymbol{0}$), but slipping in solution 1. There will then generally be a non-zero contribution to $\dot{\mathcal{E}}$ given by $\dot{\boldsymbol{v}}_i^1 \cdot (\boldsymbol{q}_i^1 - \boldsymbol{q}_i^2)$. Now the friction law (8.1–8.3) demands that $|\boldsymbol{q}_i^2| < fp_i$ and $|\boldsymbol{q}_i^1| = fp_i$, with direction diametrically opposed to $\dot{\boldsymbol{v}}_i^1$. This situation is illustrated in Figure 8.9, where it is clear geometrically that $\dot{\boldsymbol{v}}_i^1 \cdot (\boldsymbol{q}_i^1 - \boldsymbol{q}_i^2) < 0$, and hence that the contribution to $\dot{\mathcal{E}}$ from node i is negative.

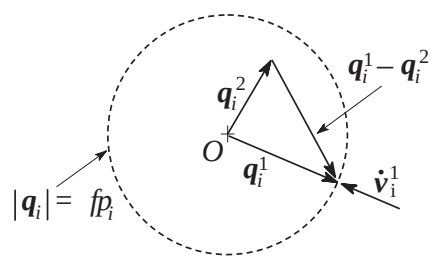


Figure 8.9: The nodal force $\boldsymbol{q}_i^2$ lies strictly within the friction cone $|\boldsymbol{q}_i| = fp_i$, whereas $\boldsymbol{q}_i^1$ lies on the cone since the node slips in solution 1.

[4] For the three-dimensional case, node i may slip in different directions in the two solutions, in which case $\boldsymbol{q}_i^1 \neq \boldsymbol{q}_i^2$. However, it can then be shown that the contribution to the norm $\dot{\mathcal{E}} < 0$ [see Problem 8.5].

Similar arguments can be applied to other combinations of nodal states and all show that the contribution to $\dot{\mathcal{E}}$ is negative from any node at which the states in solutions 1 and 2 differ (Andersson *et al.*, 2014). Thus $\mathcal{E}$ must continue to decrease with time during any period in which one or more nodes in the two solutions are in different states. Eventually, either (i) $\mathcal{E}=0$ indicating that the two solutions converge on a unique steady state, or (ii) $\mathcal{E}$ converges on a constant non-zero value, implying a situation where the states of the nodes at any given time are the same in the two solutions, but the displacements may still differ.

It can then be shown that case (ii) can arise if and only if there exists a set $\mathcal{T}$ of 'permanent stick nodes', such that for $i \in \mathcal{T}$, $\dot{\boldsymbol{v}}_i(t) = 0$ for all t in the steady state. If $\mathcal{T}$ is null, the system always converges on a unique steady state. This is consistent with the concept advanced in §8.3.3 that the memory of the system resides in nodes that are stuck, so that long term memory depends on the existence of nodes that are permanently stuck. Also, at nodes $j \notin \mathcal{T}$, the slip velocities $\dot{\boldsymbol{v}}_j^1$, $\dot{\boldsymbol{v}}_j^2$ and tractions $\boldsymbol{q}_j^1$, $\boldsymbol{q}_j^2$ must be equal, so the energy dissipated in friction at each node is unique, which as explained above has important consequences for fretting fatigue and damping.

8.4.2 Shakedown

A special case of the above result applies to the case where all nodes belong to the set $\mathcal{T}$ so that there is no slip during the steady state. This is analogous with the state known as 'shakedown' for elastic-plastic structures. We conclude that if there exists a state such that, once reached, all nodes would lie within the friction cone for all $\boldsymbol{q}^w(t)$, $\boldsymbol{p}^w(t)$, then the system will eventually shake down from any initial condition (Klarbring *et al.*, 2007). This is the frictional equivalent of Melan's theorem in plasticity (Melan, 1936).

8.4.3 Coupled systems

The above proof depends on the uncoupled assumption $\boldsymbol{B}=0$, and indeed counter-examples to Melan's theorem can be found for all cases where $\boldsymbol{B}\neq 0$. We recall that even in the uncoupled state, different initial conditions can lead to different values of the locked-in slip displacements in the permanently stuck nodes $\mathcal{T}$. If $\boldsymbol{B} \neq 0$, these will imply different normal reactions p_i at the remaining nodes, and consequent differences in frictional [slipping] tractions and displacements.

If shakedown is possible for a coupled two-dimensional system, there must exist some 'safe shakedown space' $\mathcal{S}$ in the slip displacement diagram of Figure 8.7, comprising all points that are never excluded by any of the frictional constraints at any time t during the steady state. Ahn *et al.* (2008) have shown that a two dimensional N-node system will shakedown from all initial conditions if all the frictional inequalities (8.17) are active constraints in defining $\mathcal{S}$.

If the periodic loading is expressed in the form

$$\boldsymbol{p}^w(t) = \boldsymbol{p}_0 + \lambda \boldsymbol{p}_1(t)\ ; \quad \boldsymbol{q}^w(t) = \boldsymbol{q}_0 + \lambda \boldsymbol{q}_1(t)\ , \tag{8.20}$$

where λ is a scalar load factor, there will generally exist some critical value λ_1 such that for $\lambda < \lambda_1$ Ahn's criterion is satisfied, and a larger value λ_2 above which $\mathcal{S}$ becomes null. For $\lambda_1 < \lambda < \lambda_2$, shakedown will depend on the initial conditions, whereas for $\lambda > \lambda_2$ shakedown is not possible. For the simple two-node case of Figure 8.5, $\mathcal{S}$ will be a quadrilateral for $\lambda < \lambda_1$, but this region shrinks as λ is increased and becomes a triangle for $\lambda > \lambda_1$ which collapses to a point at $\lambda = \lambda_2$.

8.4.4 Asymptotic approach to a steady state

The steady periodic state may be reached after a few loading cycles, or it may be approached asymptotically, depending on the loading history and other characteristics of the system. For example, Figure 8.10 shows the trajectory for a two node system in which, during each cycle, constraint IV moves from the dashed position to the solid position and then recedes, after which constraint I excercises a similar motion.

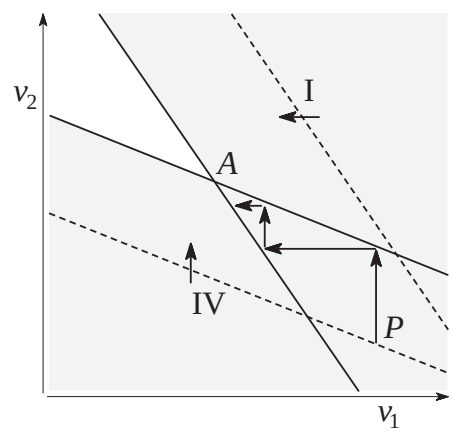


Figure 8.10: Motion of the point P to the final stuck point A by a series of geometrically decreasing slips.

Starting from the point P, the system will then follow the trajectory shown, and in particular, the final 'safe' point A will be reached by an infinite series of geometrically decreasing slips. This figure also illustrates that a node that slips during the transient period but is stuck in the steady state will generally experience at least one instant of incipient slip[5] per cycle, since the motion of the constraint that pushed it to its final position will be repeated in each subsequent cycle. We shall find this category of behaviour exhibited by continuous systems in subsequent chapters. It is also a commonly observed feature of finite-element solutions for frictional systems subject to periodic loading. Since the slips decrease geometrically during the final asymptotic approach to the steady state, it is possible to reduce computing time in

[5] Meaning that the tractions just reach the limiting condition $|\boldsymbol{q}_i| = fp_i$, usually only momentarily, but no slip occurs.

such cases by summing the geometric series, based on results from a modest number of cycles.

8.5 A simple continuum frictional system

To illustrate some of the features of frictional contact problems in a simple continuum context, we consider the one-dimensional system illustrated in Figure 8.11. A rectangular elastic strip of cross-sectional area A is pressed against a rigid plane surface by a uniform force w per unit length, after which a monotonically increasing tensile force F_0 is applied to one end. The coefficient of friction between the strip and the plane is f.

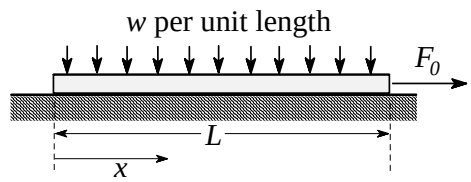


Figure 8.11: An elastic strip pressed against a rigid plane surface.

If $|F_0| < fwL$, the entire contact area cannot slip, and hence there must exist a stick region in which the sliding velocity $\dot{u}_x = 0$. For this region, integrating in time, we obtain $u_x = 0$ and hence the axial strain

$$e_{xx} = \frac{du_x}{dx} = 0 \ . \tag{8.21}$$

Neglecting Poisson's ratio effects, it then follows that the axial stress $\sigma_{xx} = 0$ and hence the local value of tensile force in the stick region is

$$F = A\sigma_{xx} = 0 \ . \tag{8.22}$$

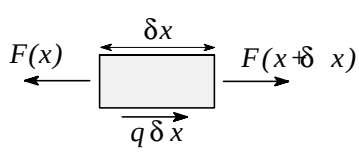


Figure 8.12: Equilibrium of a segment of the strip.

Consider now the equilibrium of a small element of the strip of length δx, as shown in Figure 8.12. If the frictional traction is q per unit length, we have

$$F(x + \delta x) - F(x) + q\delta x = 0 \tag{8.23}$$

and hence, dividing by δx and taking the limit as $\delta x \to 0$,

$$\frac{dF}{dx} + q = 0 \ . \tag{8.24}$$

Now in the stick region, we have shown that $F = 0$ and hence q must also be zero. Thus, paradoxically, there is no frictional traction in the stick region.

In a slip region, the friction force q must oppose the relative motion, so $q = -fw \ \mathrm{sgn}(\dot{u}_x)$ and

$$\frac{dF}{dx} = fw \ \mathrm{sgn}(\dot{u}_x) \ . \tag{8.25}$$

Solving for F, we have

$$F = fwx \ \mathrm{sgn}(\dot{u}_x) + C \ , \tag{8.26}$$

where C is a constant of integration.

For the given loading, it is reasonable to expect stick at the left end of the bar and slip at the right end, with $\dot{u}_x > 0$. It follows that the force F must vary with x according to Figure 8.13 in order to preserve equilibrium of elements at the ends of the bar and at the transition between stick and slip.

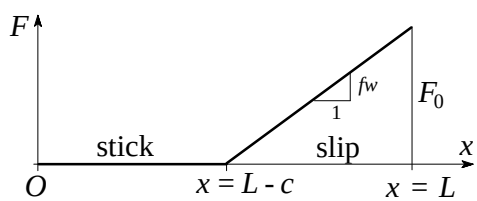


Figure 8.13: Axial force distribution for the problem of Figure 8.11.

In particular, the extent of the slip region is

$$c = \frac{F_0}{fw} \ , \tag{8.27}$$

and the tensile force $F(x)$ in the slip region is

$$F(x) = fw(x + c - L) \qquad L - c < x < L \ . \tag{8.28}$$

The configuration shown, with a region of stick and a region of microslip, can only occur as long as $c < L$ and hence $F_0 < fwL$. For larger values of F_0 there will be gross slip and the strip will accelerate to the right.

The displacement u_x must satisfy the equation

$$\frac{du_x}{dx} = e_{xx} = \frac{\sigma_{xx}}{E} = \frac{F}{AE} \ , \tag{8.29}$$

and since $u_x = 0$ in the stick zone $0 < x < L - c$, continuity at the transition point demands that

$$u_x(x) = \int_{L-c}^{x} e_{xx} dx = \frac{1}{AE} \int_{L-c}^{x} F dx \qquad L - c < x < L \ . \tag{8.30}$$

In other words, the displacement u_x is proportional to the area under the $F(x)$ curve. Substituting for $F(x)$ from (8.28) and performing the integration, we obtain

$$u_x(x) = \frac{fw}{2AE}\left[x-(L-c)\right]^2 = \frac{\left[fw(x-L)+F_0\right]^2}{2fwAE} \qquad L-c<x<L \tag{8.31}$$

for the case shown in Figure 8.13. In particular, the end displacement is

$$u_x(L) \equiv u_0 = \frac{F_0^2}{2fwAE}\ . \tag{8.32}$$

8.5.1 Unloading

The slip *velocity* can be obtained by differentiating equation (8.31) with respect to time, giving

$$\dot{u}_x = \frac{\left[fw(x-L)+F_0\right]\dot{F}_0}{fwAE}\ . \tag{8.33}$$

It follows that forward slip ($\dot{u}_x > 0$) strictly only occurs as long as F_0 is increasing and all points instantaneously stick when the maximum force $F_{\max}$ is reached.

If the force is then reduced, a region of reverse microslip will be developed near the end, in which $\dot{u}_x < 0$, so

$$\frac{dF}{dx} = -fw \tag{8.34}$$

from (8.25), and the tensile force F has the form shown in Figure 8.14. Axial forces and hence frictional tractions developed during the loading phase are 'locked in' except in those regions where reverse slip has occurred. In particular, if the force is reduced to zero, the backward slip zone will be exactly half of the original forward slip zone and we shall leave a residual axial force field of the form shown dashed in Figure 8.14.

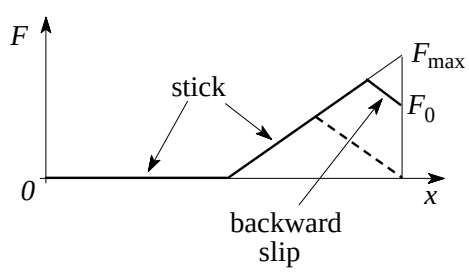


Figure 8.14: Axial force distribution during unloading.

At this point, the extension of the bar is again proportional to the area under the $F(x)$ curve and is therefore half that at $F_{\max}$, being given by

$$u_1 = \frac{F_{\max}^2}{4fwAE}\ . \tag{8.35}$$

Although the force has now been reduced to zero, the strip is left in a state of residual stress and hence retains some memory of the loading cycle. In particular, strain measurements in the residual state would enable us to deduce the value of $F_{\max}$ during the loading cycle. However, if unloading were continued into negative values with $F_0 \leq -F_{\max}$, memory of the loading phase would be 'erased'.

8.5.2 Periodic loading

Consider now the case where the force oscillates between zero and $F_{\max}$. In the steady state, slip will occur only in the region that slipped backwards during unloading, so there is a region that slipped only during the first half cycle and thereafter remained stuck. This is characteristic of such problems — the residual stress developed during the first cycle tends to reduce the extent of slip during subsequent cycles.

Both loading and unloading phases involve axial force distributions with straight-line segments of slope $\pm f$ as in Figure 8.14 and elementary calculations show that the displacement increases quadratically with force as in equation (8.32). In particular, whilst the force is increasing from zero to $F_{\max}$ in the second and subsequent cycles, the displacement at the end is

$$u = u_1 + \frac{F_0^2}{4fwAE} \tag{8.36}$$

and the work done by the force is

$$W_1 = \int_{F_0=0}^{F_0=F_{\max}} F_0\, du = \int_0^{F_{\max}} F_0\, \frac{F_0}{2fwAE} dF_0 = \frac{F_{\max}^3}{6fwAE}\ . \tag{8.37}$$

During unloading we have

$$u = \frac{F_{\max}^2}{2fwAE} - \frac{(F_{\max} - F_0)^2}{4fwAE} \tag{8.38}$$

and the work done is

$$W_2 = \int_{F_0=F_{\max}}^{F_0=0} F_0\, du = \int_0^{F_{\max}} F_0\, \frac{(F_{\max} - F_0)}{2fwAE} dF_0 = -\frac{F_{\max}^3}{12fwAE}\ . \tag{8.39}$$

Thus only half of the work done during loading is recovered during unloading and there is a frictional dissipation

$$W = W_1 + W_2 = \frac{F_{\max}^3}{12fwAE}\ . \tag{8.40}$$

Notice also that the dissipation varies with the cube of the force amplitude. This is a fairly common relation for a range of elastic frictional systems, including those in two and three dimensions.

8.5.3 Discrete model of the strip problem

If the elastic strip of Figure 8.11 is modelled using the finite element method, the resulting discrete system is mathematically equivalent to that shown in Figure 8.15, where a set of rigid blocks make contact with a rigid plane and are connected by springs of stiffness k. If the entire strip is represented by N blocks separated from each other by a distance L/N, the system behaviour will approach that of the continuous strip if each block is pressed against the plane by a force $P = wL/N$ and the springs are of stiffness $k = AEN/L$.

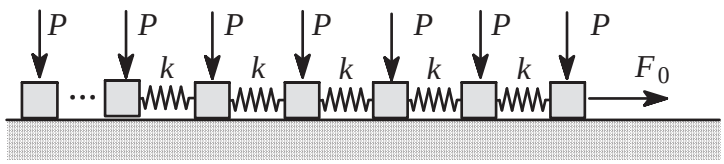


Figure 8.15: Discrete model approximating the strip of Figure 8.11.

Models of this kind were introduced by Iwan (1966, 1967) to approximate nonlinearities associated with material plasticity, and are therefore sometimes known as Iwan models. In particular, they have been used to explore the behaviour of bolted joints between engineering components (Song *et al.*, 2004, Quinn and Segalman, 2005), which are believed to account for a significant proportion of the damping in built-up structures.

One of the characteristics of Iwan models, shared by continuum models of frictional microslip, is that the slip displacements at nodes (regions) that are instantaneously stuck retain some memory of the loading history, but in most cases this is restricted to a memory of the extreme values of the force.

8.5.4 The inverse problem

We showed in §8.5.1 that the system of Figure 8.11 generally remains deformed when the tangential force F_0 is removed. The residual tractions must satisfy the frictional condition $|q(x)| \leq fp(x)$, but we might reasonably ask the question "Can a loading history be devised such as to develop any desired self-equilibrated distribution of tangential traction satisfying this condition?"

The answer is certainly 'yes' for the discrete model of Figure 8.15, since we could proceed as follows:-

1. If the desired residual friction forces at all the blocks are given, equilibrium arguments can be used to determine the forces in the springs, and hence also their extensions Δ_i.
2. If $\Delta_1 > 0$, we increase the force monotonically to

$$F_0 = (N-1)fP + k\Delta_1 \ .$$

   This will cause all the blocks except the first to slip to the right, and will achieve the required extension Δ_1.

3. We now *decrease* F_0 monotonically until

$$F_0 = -(N-2)fP + k\Delta_2 \ ,$$

at which point all except blocks 1 and 2 are slipping to the left and the desired extension is achieved in the second spring.

Repeating this procedure, it is clear that all the desired spring extensions can be achieved by applying alternating tensile and compressive forces F_0 of successively decreasing amplitude, such that at each step one fewer blocks slip[6].

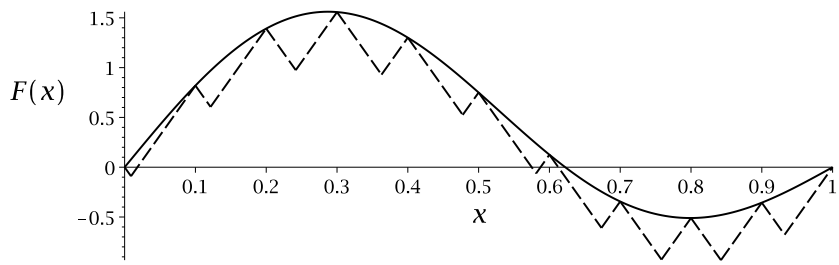


Figure 8.16: Residual force distribution [dashed line] fitting a desired distribution [solid line] at discrete points.

The same technique could be applied to the continuous strip of Figure 8.11, but the resulting distribution of residual tensile force $F(x)$ will consist of straight line segments of slope $\pm fw$, as in Figure 8.14. The desired distribution is therefore achieved only at a finite set of points, as illustrated in Figure 8.16. By increasing the number of force reversals, we can achieve a closer fit, but the residual stress curve remains non-dfferentiable. A related result is that this one-dimensional system 'remembers' only a denumerable set of discrete extreme values of F_0, whereas the desired function $F(x)$ has a non-denumerable infinity of degrees of freedom.

8.6 More complex friction laws

Amontons' law of §8.1 is clearly a considerable oversimplification of the physics of friction, and numerous more sophisticated models have been proposed as improvements. For example, the friction coefficient may exhibit dependence on sliding velocity or normal contact pressure, possibly as a result of thermal or viscous effects. Also, creep at areas of actual contact may lead to a more complex dependence on previous loading or sliding history. Many of these models have been proposed as explanation of dynamic instabilities such as 'squeal' in automotive brakes, or fault rupture of tectonic plates, leading to earthquake events.

[6] If $\Delta_1 < 0$, the first loading should be monotonically increasing *compression* to the value $F_0 = -(N-1)fP + k\Delta_1$, after which the forces alternate in sign as before.

8.6.1 Instabilities during steady sliding

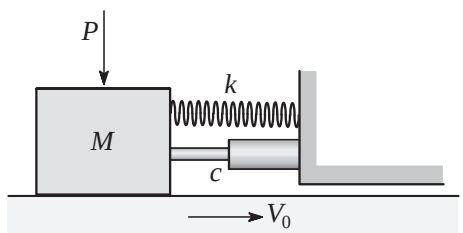


Figure 8.17: A system prone to frictional vibrations.

Figure 8.17 shows a dynamic system comprising a mass M, loaded by a normal force P, and supported by a spring of stiffness k and a damper of coefficient c. The mass rests on a rigid plane surface that translates to the right at constant speed V_0.

With an appropriate friction law, this system provides a simple explanation of phenomena such as brake squeal, stick-slip motion of machine-tool slides, creaking doors and the squeak of chalk on a blackboard.

8.6.2 Velocity-dependent friction coefficient

Experimental measurements of friction forces under steady sliding conditions often take the form of the *Stribeck curve*, shown in Figure 8.18, where the friction coefficient first decreases and then increases with sliding speed. The rising portion of this curve is associated with viscous effects, particularly if the contact is lubricated, in which case the equilibrium lubricant film thickness increases with sliding speed. At very low speeds, the roughness of the surfaces exceeds the theoretical film thickness and the frictional behaviour becomes dominated by solid-to-solid contact, resulting in a higher coefficient of friction.

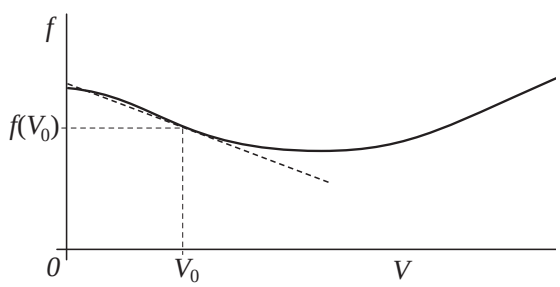


Figure 8.18: The Stribeck curve: relation between friction coefficient f and sliding speed V. The dashed line represents the linear approximation (8.41).

For small perturbations from the steady-state velocity V_0, we can approximate this curve by the linear function

$$f(V) = f(V_0) + f'(V_0)(V - V_0) \ , \tag{8.41}$$

shown by the dashed line in Figure 8.18.

If the rightward elastic displacement of the mass in Figure 8.17 is denoted by $u(t)$, the instantaneous sliding velocity is

$$V = V_0 - \dot{u} \ , \tag{8.42}$$

and the equation of motion takes the form

$$M\ddot{u} + c\dot{u} + ku = Pf(V) = Pf(V_0) - Pf'(V_0)\dot{u} \ , \tag{8.43}$$

or

$$M\ddot{u} + [c + Pf'(V_0)]\,\dot{u} + ku = Pf(V_0) \ . \tag{8.44}$$

This equation has a steady-state solution in which the mass is stationary at the point $u = Pf(V_0)/k$, but the transient solution also includes exponentially growing perturbations from the steady state if the composite damping term is negative, or

$$f'(V_0) < -\frac{c}{P} \ . \tag{8.45}$$

In other words, steady sliding is unstable if the negative slope of the Stribeck curve is sufficiently steep near the steady-state velocity.

In most practical cases, the negative effective damping term is relatively small, so that the exponential growth is oscillatory at a frequency close to the natural frequency of the system. The system may then reach a non-sinusoidal limit cycle as the amplitude of vibrations carries the instantaneous speed into regions where the negative slope of the Stribeck curve is less steep.

Numerous other physical mechanisms have been proposed to account for frictional vibrations (Ibrahim, 1994), but in most cases the result is to destabilize a lightly damped natural resonance of the system (Moirot and Nguyen, 2000). One consequence is that accurate predictions of experimentally observed frequencies do not provide evidence in favour of any particular theory.

8.6.3 Stick-slip vibrations

In some cases, the oscillation may grow until the instantaneous sliding speed falls to zero at some point in the cycle. Generally the non-linearity associated with the frictional inequalities will then cause the system to achieve a limit cycle involving periods of stick and slip (Moirot *et al.*, 2003).

Experiments show that the force required to initiate slip is often greater than that required to sustain motion (Rabinowicz, 1995), in which case we can define a *coefficient of static friction* f_s, and of a *coefficient of dynamic friction* f_d. In this case, stick-slip vibrations can occur even if f_d is independent of sliding speed.

For the system of Figure 8.17, stick corresponds to $\dot{u} = V_0$ and, once initiated, it will persist until

$$cV_0 + ku = f_s P \quad \text{or} \quad u = \frac{f_s P - cV_0}{k} \ . \tag{8.46}$$

This equation and $\dot{u} = V_0$ define the initial conditions for the subsequent slip phase which is governed by the equation of motion

$$M\ddot{u} + c\dot{u} + ku = f_d P \ . \tag{8.47}$$

The response depends qualitatively on the dimensionless parameters

$$\lambda = \frac{\omega_n (f_s - f_d) P}{k V_0} \ ; \quad \zeta = \frac{c\,\omega_n}{2k} \quad \text{where} \quad \omega_n = \sqrt{\frac{k}{M}} \tag{8.48}$$

is the undamped natural frequency of the system. Figure 8.19 shows the velocity $\dot{u}$ during the slip phase for $\lambda = 1$ and two values of the damping factor ζ.

For $\zeta = 0.08$, the resulting oscillation decays sufficiently rapidly to ensure that the mass never again reaches the speed V_0 and hence slip continues indefinitely. Thus, stick-slip oscillations are not possible for this case.

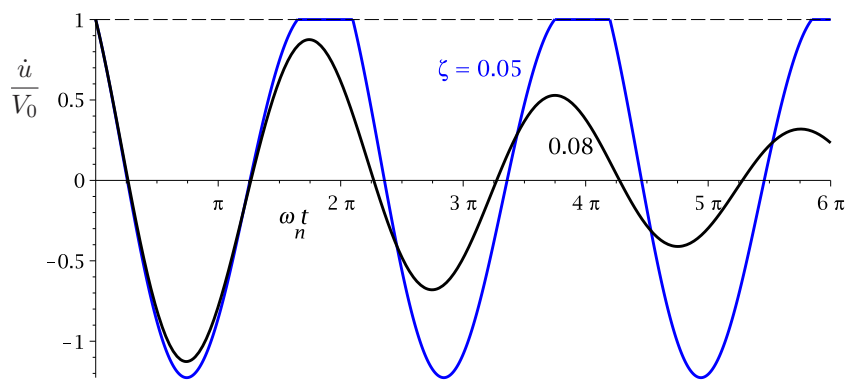


Figure 8.19: Slip velocity following a period of stick for the system of Figure 8.17.

For a lower level of damping $\zeta = 0.05$, the mass reaches V_0 before the end of the first cycle of oscillation and then remains stuck until the condition (8.46)$_2$ is once again satisfied. A steady-state stick-slip oscillation is therefore reached in the first cycle. However, under these conditions, the equation of motion also supports a continuous slip solution with $\dot{u} = 0$ and this state is stable, so the final state depends on the initial conditions. Figure 8.20 defines the range of values of λ and ζ in which stick-slip motion can occur. Notice that if the system is critically damped ($\zeta > 1$), the velocity $\dot{u}$ decreases monotonically with time and stick-slip is impossible even for arbitrarily large λ.

In Figure 8.19, the initial excursion to negative values of $\dot{u}$ is of the same order as λ. Since the average value of $\dot{u}$ over a steady-state cycle must be zero, this implies that the duration of the stick phase increases with λ. In particular, when V_0 is very small, the motion will comprise widely spaced slip events of duration slightly longer

than π/ω_n. This situation can arise in machine-tool slides. An even more striking example is the 'clanking' of heating pipes supported in frictional clamps, where gradual temperature changes cause the pipe to experience a set of widely separated slip events.

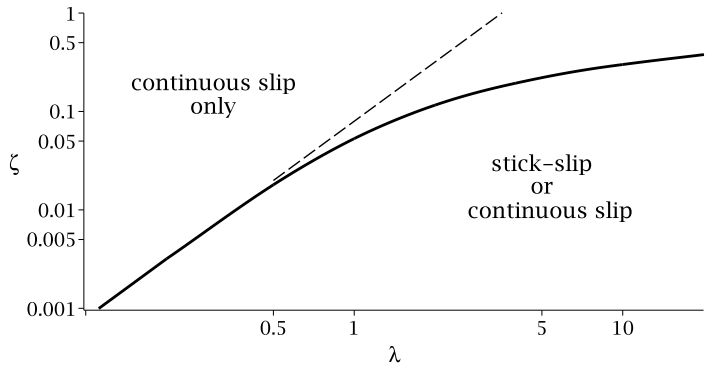


Figure 8.20: Stability boundary for the system of Figure 8.17 with $f_s > f_d$. The dashed line corresponds to the relation $\zeta = \lambda^2/4\pi$.

8.6.4 Slip-weakening laws

If the static and dynamic friction coefficients differ, we must postulate a physical change of some sort when motion commences, and it is unlikely that this can occur instantaneously. A more likely postulate is that some finite slip displacement must take place before the dynamic coefficient is achieved, and this suggests a law of the form $f = f(h)$, where h is the shift (slip displacement) defined in equation (7.1), and

$$f(0) = f_s \; ; \qquad f(h) \to f_d \; ; \quad h \to \infty \; . \tag{8.49}$$

We shall argue in Chapter 18, §18.6 that when two rough solids slide over each other, the typical event comprises the interaction of asperities in the two surfaces, and the shift h occurring during one such event will be of the order of the dimensions of typical actual contact areas, and hence very small. Rabinowicz (1951) described experiments in which the energy lost in friction during a frictional impact was measured and compared with independent measurements of f_s and f_d under steady-state conditions. He was therefore able to estimate the length scale Δ defined such that

$$(f_s - f_d)\Delta = \int_0^\infty [f(h) - f_d]dh \; . \tag{8.50}$$

For the dry sliding of metals, he obtained values of a few micrometers.

Papangelo *et al* (2015) have shown that in the limit where Δ is very small compared with the macroscopic contact area dimensions, the transition from f_s to f_d is mathematically equivalent to a shear fracture governed by a stress-intensity factor

$$K_{\mathrm{II}} = \sqrt{2E^*(f_s - f_d)p\Delta}\ , \tag{8.51}$$

where p is the local contact pressure. There is therefore a close connection between slip-dependent friction laws and fault propagation laws based on fracture mechanics arguments (Abercrombie and Rice, 2005, Ben Zion, 2008, Svetlizky and Fineberg, 2014).

8.6.5 Rate-state laws

The slip-weakening law is one of a class of friction models known as *rate-state models*, based on the premise that if the contact pressure p or the sliding velocity V change suddenly, the friction force does not immediately reach its new steady-state value, but approaches it through some transition process (Ruina, 1983, Rice *et al.*, 2001). These models have been used to eliminate unphysical predictions of elastodynamic instability of arbitrarily small-wavelength perturbations predicted by Amontons' law [see §19.5], but they also provide a convenient way of quantifying experimental results that are not well characterized by the simpler law.

We express the magnitude of the frictional traction $\boldsymbol{q}$ as a function

$$|\boldsymbol{q}| = g_1(S, p, V)\ , \tag{8.52}$$

where S is a state variable introduced to describe history-dependent changes in the local contact condition. This relation is then supplemented by an evolution law for S, typically of the form

$$\dot{S} = g_2(S, p, V)\ . \tag{8.53}$$

The state variable S can be identified with a physical property of the system, or alternatively, equations (8.52, 8.53) can be regarded as defining a 'black box' characterization of the system, whose parameters are determined by appropriate experiments.

As an example of the former approach, we can recreate the slip weakening law of §8.6.4 by treating the slip distance h as the state variable during slip periods, so that

$$g_1(S, p, V) = f(S)p\ ; \quad g_2(S, p, V) = |V| \qquad V \neq 0\ , \tag{8.54}$$

since in this case $\dot{S} = \dot{h} = |V|$.

For completeness, we require a further equation defining the rate at which the static friction coefficient is recovered during periods of stick. For example, if the static coefficient is approached exponentially on a time scale t_0, we might define

$$g_1(S, p, V) = f(S)p\ ; \quad g_2(S, p, V) = -\frac{S}{t_0} \qquad V = 0\ . \tag{8.55}$$

Alternatively, we could take the proportion of the nominal contact area in actual contact as the state variable. During static contact, creep allows this to increase asymptotically to a limiting value, but during sliding, contacts are continually being established and broken, so the growth of individual contacts is limited by the average dwell time (Bar Sinai *et al.*, 2012).

Using the black-box approach, some information about the functions g_1, g_2 can be obtained by (i) conducting steady-state sliding experiments with constant p, V, and then (ii) measuring the transient that is obtained when p or V is suddenly changed to a new constant value. Equation (8.52) suggests that there will then be an immediate discontinuity in $|\boldsymbol{q}|$, followed by an asymptotic approach to the new steady state governed by the evolution of the state variable. Experimental results exhibiting this behaviour were discussed by Ruina (1983). Other experiments suitable for system characterization include steady sinusoidal variation of p or V about some mean value (Cabboi *et al.*, 2016), and static friction tests where the traction needed to initiate motion is measured as a function of the duration of static contact under constant p.

PROBLEMS

1. The system of Figure 8.2 (a) is loaded by the periodic force

$$F_1 = F\cos(\omega t)\ ; \quad F_2 = F\sin(\omega t)\ ,$$

where the frequency ω is sufficiently slow for quasi-static conditions to apply.

Identify the times during each steady-state cycle that transitions occur between the various possible states, if the coefficient of friction $f = 0.5$ and

$$\boldsymbol{K} = \begin{bmatrix} 10 & 4 \\ 4 & 10 \end{bmatrix}\ .$$

Would the behaviour be qualitatively different if the coefficient of friction were lower?

2. The system of Figure 8.2 (a) is subjected to a force $\boldsymbol{F}(t) = \{C_1 t, C_2 t\}$ starting at time $t = 0$, where $f > k_{11}/k_{21} > 0$ and the constants C_1, C_2 are defined such that $\boldsymbol{F}$ moves through the multiple solution segment in Figure 8.5.

(i) Determine the quasi-static displacements $v(t), w(t)$ under the assumption of backward slip.
(ii) Now set up the dynamic equations of motion for backward slip, including the inertia term $-M\ddot{\boldsymbol{u}}$.
(iii) Show that the time-varying displacements from (i) satisfy the equations of motion, provided the mass has an appropriate initial velocity at time $t = 0$, and find this velocity.
(iv) Show that if the initial velocity differs very slightly from the quasi-static value, this perturbation will then grow without limit. In other words, that the backward slip state is unstable.

3. The contact stiffness matrix (8.13) for a two-node frictional system is defined such that

$$\boldsymbol{A} = \begin{bmatrix} 10 & 3 \\ 3 & 20 \end{bmatrix}\ ; \quad \boldsymbol{B} = \begin{bmatrix} 3 & 10 \\ 10 & 3 \end{bmatrix}\ ; \quad \boldsymbol{C} = \begin{bmatrix} 20 & 10 \\ 10 & 20 \end{bmatrix}\ .$$

Use Klarbring's criterion [§8.3.4] to find the critical coefficient of friction above which the rate problem is ill-posed.

4. A two-node frictional system is defined by the contact stiffness matrix of Problem 3. Find the minimum coefficient for which the system is capable of being wedged in a state with no external forces $[p_i^w = q_i^w = 0]$ but non-zero nodal displacements. Notice that the critical state may involve both nodes in contact $[w_1 = w_2 = 0]$, or only one node in contact.

5. Construct a diagram analogous to Figure 8.9 for the three-dimensional case where node i is slipping in both solutions 1 and 2, but not in the same direction. Show that the contribution of this node to $\dot{\mathcal{E}}$ is strictly negative under these conditions. Hence show that in the steady state, the direction of slip at slipping nodes is independent of initial conditions.

6. Figure 8.21 shows the extreme positions of the frictional constraints for a two-node system, so that the central white quadrilateral represents the safe shakedown space $\mathcal{S}$. The periodic loading scenario causes the constraints to advance to and then retreat from these positions in the sequence I,III,II,IV,I,III.... Track the motion of the point P from an arbitrary starting point outside $\mathcal{S}$ and hence partition this space into regions corresponding to the final position of P lying on the lines AB, BC, CD, DA, or in one of the corners A, B, C, D. Also, explain which starting positions will involve the shakedown state being reached in one or two cycles, and which will involve an asymptotic approach to the steady state.

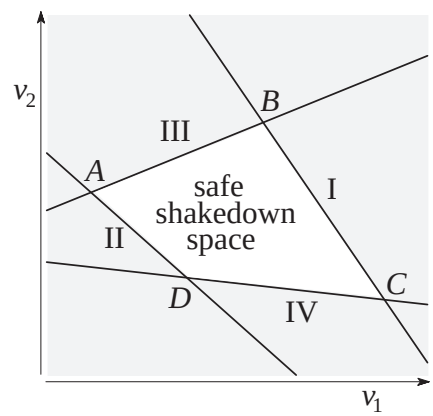


Figure 8.21:

7. A cylindrical bar of diameter D and length $L\,(\gg D)$ just fits inside a long cylindrical hole in a rigid body. The bar is now heated to a temperature T. If the coefficient of friction between the bar and the hole is f, find the distribution of contact pressure p and the mean axial stress σ_{zz} as functions of distance z along the bar. The mate-

rial of the bar has Young's modulus E, Poisson's ratio ν and coefficient of thermal expansion α. Assume plane cross-sections remain plane.

8. Figure 8.22 shows the geometry of a standard 'pull-out' test for fibrous composite materials, in which a tensile force F_0 is applied to the end of an exposed fibre and increased until failure occurs of the bond between the fibre and the matrix.

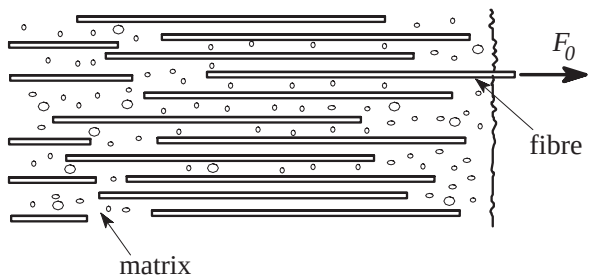


Figure 8.22: The 'pull out' test.

An approximate model of the process is shown in Figure 8.23, where the fibre is represented by a rod (a) and the matrix by a pair of rods (b). The stiffnesses of the rods are denoted by k_a, k_b respectively, defined through the relations

$$e_{xx} = \frac{du_x}{dx} = \frac{F(x)}{k} \ ,$$

where u_x is the local displacement of the rod in the x-direction and $F(x)$ is the tensile force in the rod at the point (x).

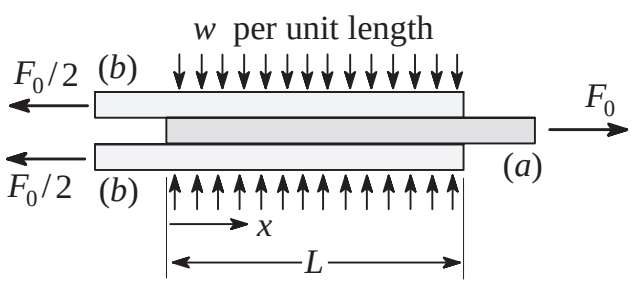


Figure 8.23: Rod model of fibre pull out.

We make the further assumption that the transfer of force between the rods is achieved solely by friction, through a coefficient of friction, f, and that the rods are pressed together by a normal force w per unit length.

The force F_0 is gradually increased from zero. Describe what happens and determine the location and magnitude of any regions of slip as functions of F_0. Also, give expressions for the axial forces in the matrix and the fibre as functions of x.

9. Suppose that each block in the discrete model of Figure 8.15 has mass M. At time $t=0$, all the springs are unstretched, and a force $F_0 > fP$ is suddenly applied to the last block as shown in the figure. Find the displacement u of this block assuming that all the remaining blocks remain stationary. Hence determine the time $t=t_1$ at which the second block will just start to move.

10. Suppose that the system of Figure 8.17 can support a steady-state vibration such that

$$u(t) = u_0 + A\cos(\omega t) \ .$$

Calculate (i) the energy absorbed in the damper and (ii) the work done by the friction force on the mass during one complete cycle $2\pi/\omega$, assuming the friction coefficient is given by equation

$$f = f_0 - f_1(V - V_0) \ ,$$

where f_0, f_1 are positive constants. Hence show that such a state is possible if and only if $f_1 = c/P$ and that in this case the amplitude A can take any value.

Repeat the calculation for the case where the friction coefficient is defined by the cubic relation

$$f(V) = f(V_0) - f_1(V - V_0) + f_3(V - V_0)^3 \ ,$$

where $f_3 > 0$. Show that the amplitude in the steady state is now determinate and find its value[7].

11. A state-variable friction law for unidirectional sliding at speed V is defined by the relations

$$\dot{S} = |V| - \frac{S}{t_0} \ ; \quad Q = f(S)P \ ; \quad f(S) = f_0 + f_1 \exp\left(-\frac{S}{L_0}\right) \ ,$$

where P, Q are the normal and frictional force respectively, S is the state variable which has dimensions of length, and t_0, L_0 are constants with dimensions of time and length respectively.

(i) If the bodies slide at constant speed V, what will be the relation between the coefficient of friction Q/P and V.
(ii) If the sliding speed is held constant at V_0 for some time, but then changed suddenly to V_1, how will the friction force vary during the ensuing transient.

[7] Note that this is an approximate solution, since the equation of motion will not generally be satisfied by the sinusoidal function for all t.