Frictional Hertzian contact problems under cyclic loading using static reduction

Jae Hyung Kim, Yong Hoon Jang*

School of Mechanical Engineering, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-749, Republic of Korea

ARTICLE INFO

Article history:
Received 5 June 2013
Received in revised form 15 September 2013
Available online 12 October 2013

Keywords:
Hertzian contact
Cyclic loading
Frictional energy dissipation
Static reduction

ABSTRACT

In this study we investigate an axisymmetric Hertzian contact problem of a rigid sphere pressing into an elastic half-space under cyclic loading. A numerical solution is sought to obtain a steady state, which demands a large amount of computer memory and computing speed. To achieve a tractable problem, the current numerical model uses a “static reduction” technique, which employs only a contact stiffness matrix rather than the entire stiffness of the problem and is more accurate than the approach used by most finite element codes. Investigation of the tendency of contact behavior in the transient and steady states confirms that a steady state exists, showing converged energy dissipation. The dependence of dissipation on load amplitude shows a power law of load amplitude less than 3, which may explain some deviations in the experimental findings.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

In many practical applications, contacts between elastic components are subjected to cyclic normal and tangential loads due to the vibration of machine components. It is well known that microslip can occur in localized regions. The resulting cyclic microslip can generate fretting damage, which is a form of failure known as fretting fatigue crack, and associated frictional energy dissipation, which provides effective structural damping (Nowell et al., 2006; Jang and Barber, 2011a,b).

Numerous theoretical and experimental studies of this process have been conducted, both for idealized geometries such as Hertzian contacts and for more practical systems such as bolted lap and flange joints. Specifically, Mindlin et al. (1952) considered a Hertzian contact problem where the tangential load is oscillatory and cyclic microslip is predicted in an annulus surrounding a central stick zone. Johnson (1955, 1961) provided classical experimental evidence of the resulting fretting damage, showing that the extent of visibly damaged regions in the contact areas of spheres correlates well with Mindlin’s prediction of the microslip zone, but the dependence of energy dissipation on load amplitude showed a lower power dependence than the theoretical value. More recently, Barber et al. (2011) and Putignano et al. (2011) also investigated the contact of elastically similar half-spaces under oscillatory loading to show the variation in frictional energy dissipation per cycle in terms of system parameters such as the normal incremental stiffness of the contact, the external forces, the local coefficient of friction, the relative phase between oscillation under normal and tangential loads (including Jang and Barber, 2011b), and the amplitude at sufficiently small oscillatory loads. The lower power dependence of dissipation was also reported. In the present study, we examine the dependence of dissipation on load amplitude and hypothesize that it is less than the theoretically-predicted cubic amplitude.

For a simpler normal loading problem, Spence (1968, 1973, 1975) solved the axisymmetric Hertzian contact problem, the two-dimensional Hertzian problem, and the axisymmetric flat punch problem, all for loading only where the stick and slip zones were present in the same ratio. These contact problems are governed by three parameters: the relative region of the stick-slip radius, the coefficient of friction, and Poisson’s ratio. Comprehensive solutions for the relative region of stick-slip were obtained from many researchers (Eлагуine et al., 2006; Jelagin and Larsson, 2008). Specifically, they investigated the change of contact states during loading and unloading cycles in a numerical analysis of the frictional contact cycle and showed that the relative stick region in the loading cycle varies and the additional slip annulus spreads towards the center. An analytical solution only for the axisymmetric flat punch unloading problem was obtained by Turner (1979). A numerical analysis recently followed by Ahn and Barber (2008) and afterwards by Lee et al. (2012), was performed to investigate a model of an elastic block pressed against a frictional rigid plane under oscillating loading, showing continuous variation of the contact state such as stick and slip boundaries. One important result was that the “receding contact” system approaches a steady contact state after transient behavior. More recently, Stingl et al. (2013) solved the periodic loading problem for a two-dimensional...
Hertzian contact, stating that it is “computationally less demanding” than the axisymmetric problem. This is because they used the Kalker algorithm “CONTACT” (Vollebregt E.A.H., 2012), which can only be applied to the axisymmetric problem by treating it as fully three-dimensional. Kalker’s algorithm is more efficient than the straight finite element method because it only manipulates the surface values and hence requires only discretization of the surface. In the present study we aim to achieve the same reduction in computing effort while using an axisymmetric finite element model in the process of static reduction.

In this paper, we investigate an axisymmetric Hertzian contact problem of a rigid sphere pressing into an elastic half-space under an oscillating load to explore the tendency of contact behavior in the transient and steady states, and the dependence of dissipation on load amplitude. The proposed numerical model is intrinsically “a computationally more demanding problem” and it can be a barrier to achieving accuracy, particularly in cyclic loading problems where a sufficient number of cycles is run to reach a steady state. A possible attractive numerical procedure known as static reduction (Thaitirarot et al., 2013), which uses only a contact stiffness matrix rather than the whole stiffness of the problem, is adopted.

2. Problem description

We consider the axisymmetric contact problem shown in Fig. 1, in which a rigid spherical indenter with a radius $R$ is pressed against an elastic plane surface by a time-varying force $F(t)$. Note that the contact area is denoted as $a$. The interface between the rigid indenter and the plane is subjected to Coulomb friction conditions with the friction coefficient $f$. The quasi-static analysis is performed with the consideration of coupling between normal and tangential effects.

2.1. Equilibrium equations using reduced contact stiffness

Since the finite element model is often rather large and may be a burden to achieving accuracy, particularly in cyclic loading problems, we would need to reduce the finite element model to a contact focused model defined on $N$ contact nodes alone. The procedure known as static reduction (Thaitirarot et al., 2013) is well explained when the finite element body comprises three node sets: the contact nodes, the externally loaded nodes, and the unloaded nodes. However, since the current model is intended to solve the contact problem in the absence of externally loaded nodes, we summarize and extend the process to form the contact stiffness matrix without consideration of the externally loaded nodes.

Suppose the elastic body is discretized using a total of $N$ contact nodes and $M$ non-contact nodes. For clarity, the contact nodes are identified by the superscript $C$ and the non-contact (internal) nodes by the superscript $I$. The contact nodal forces $f_i$ and the corresponding nodal displacements $u_i$ in two dimensions are denoted by

$$
\mathbf{f}_i = \{q_i, p_i\}, \quad \mathbf{u}_i = \{v_i, w_i\},
$$

where $p_i, q_i$ are the normal and tangential forces in the contact surface, respectively and $v_i, w_i$ are the normal and tangential (slip) displacements of the contact surface, respectively. The sign convention of the contact forces and the displacements are shown in Fig. 2.

Then the contact nodal force and nodal displacement vectors for the linearly elastic system satisfy the following equation:

$$
\mathbf{f}^c = \mathbf{f}^a + K^c \mathbf{u}^c,
$$

where the force and displacement vectors are aligned as

$$
\mathbf{f}^c = [f_1^c, f_2^c, \ldots, f_N^c]^T; \quad \mathbf{u}^c = [u_1^c, u_2^c, \ldots, u_N^c]^T
$$

and $K^c$ is a reduced stiffness matrix and $\mathbf{f}^a$ comprises the contact nodal forces that would be developed by the given external loading if the contact nodal displacements were all constrained to be zero ($\mathbf{u}^c = \mathbf{0}$). It is noted that the vector $K^c$ is a positive definite and symmetric matrix.

If the contact nodal forces and displacements have been determined, the total energy dissipation per cycle can be obtained in one complete cycle of loading as:

$$
W = \sum_{i=1}^{N} \int Q_i(t) v_i(t) dt
$$

where the dot denotes the derivative with respect to time $t$. Since the frictional force always opposes the motion during slipping, the dissipation is always positive.

It is beneficial to represent contact pressure $p_n$ and contact shear traction $q_n$ instead of contact forces since the nodal forces depend on the mesh, so are not really of any fundamental interest and it is better to compare directly with Spence’s analytical solution (Spence, 1975), which plots pressure. It is possible to convert the contact nodal forces to the contact tractions by using the shape functions for the finite element model.

2.2. Partitioning the stiffness matrix

The last term in (2) defines the contact nodal forces required to generate contact nodal displacements $\mathbf{u}^c$ in the absence of external loads (since the effect of the external loads is included in the first term $\mathbf{f}^a$). More specifically, if the non-contact boundary contains a region $\Gamma_f$ at which external tractions are prescribed, and/or a region $\Gamma_u$ at which displacements are prescribed, we should use

![Fig. 1. Axymmetric contact problem in which a rigid spherical indenter with radius $R$ is pressed against an elastic plane surface by a time-varying force $F(t)$.](image1)

![Fig. 2. Sign convention for contact forces and displacements.](image2)
the boundary conditions \( f_j = 0 \) for \( j \in \Gamma_f \) and \( u_j = 0 \) for \( j \in \Gamma_u \) when solving for the reduced stiffness matrix \( K^r \). The overall stiffness matrix \( K \) of the finite element model so defined can be partitioned such that

\[
f = \begin{bmatrix} f^i \\ f^c \end{bmatrix} = Ku = \begin{bmatrix} K^i \quad K^{ic} \\ K^{ci} \quad K^c \end{bmatrix} \begin{bmatrix} u^i \\ u^c \end{bmatrix},
\]

or

\[
f^i = K^i u^i + K^{ic} u^c; \quad f^c = K^{ci} u^i + K^c u^c. \tag{5}
\]

Since the full stiffness matrix \( K \) is symmetric, it follows that \( K^T \) is the transpose of \( K^C \). Now in the problem under consideration there are no nodal forces at the non-contact nodes, so \( f^i = 0 \), and thus

\[
K^i u^i = -K^{ic} u^c.
\]

from (5). We can solve this matrix equation obtaining formally

\[
w^i = -\left[ K^i \right]^{-1} K^{ic} u^c.
\]

Substituting into the right-hand side of (5), we then obtain

\[
f^c = K^c u^c \tag{6}
\]

where

\[
K^c = \left( -K^T \left[ K^i \right]^{-1} K^{ic} + K^C \right)
\]

is the required reduced stiffness matrix.

In addition, if the force and displacement vector of (3) are reordered as

\[
\begin{align*}
\{ u^i \\ u^c \} &= \{ v_1, v_2, v_3, \ldots, v_N, w_1, w_2, w_3, \ldots, w_N \}^T \\
\{ f^c \} &= \{ q_1, q_2, q_3, \ldots, q_N, p_1, p_2, p_3, \ldots, p_N \}^T,
\end{align*}
\]

Eq. (3) can be constructed as

\[
\begin{bmatrix} Q \\ P \end{bmatrix} = \begin{bmatrix} A & B' \\ B & C \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} + \begin{bmatrix} Q^w \\ P^w \end{bmatrix} \tag{10}
\]

where the stiffness matrix with partitioned sub-matrices \( A, B, C \) should be symmetric. It is also noted that the current contact problem, a frictional elastic system under cyclic loading, has \( B \neq 0 \), which makes the contact problem a “coupled” problem.

2.3. Contact boundary conditions for the Hertzian problem with an initial gap between the bodies

The current contact problem, known as an incomplete contact problem, has an intrinsically small gap between two contact surfaces. It is usually accepted that the gap is compared with the expected dimensions of the contact area. This restriction can ensure that the derivative of the gap is also small compared with unity. Thus, it is appropriate that a set of nodes on each of the two contact surfaces form a line joining two adjacent nodes that is approximately perpendicular to each surface. With the aid of this conjecture, the nodal gap at node \( j \) after deformation is given by \( g_j + w_j \), where \( g_j \) is the length of the line joining two nodes on the perpendicular surfaces \( w_j \). Then, the contact boundary condition for the Hertzian contact problem should be

\[
\begin{align*}
w_j = -g_j; & \quad i_j = 0; \quad P_j \geq 0: \quad |Q_j| \leq |P_j|: \quad \text{Stick} \tag{11} \\
g_j + w_j > 0; & \quad P_j = 0; \quad Q_j = 0: \quad \text{Separation} \tag{12} \\
w_j = -g_j; & \quad i_j > 0; \quad P_j \geq 0: \quad Q_j = fP_j: \quad \text{Forward slip} \tag{13} \\
w_j = -g_j; & \quad i_j < 0; \quad P_j \geq 0: \quad Q_j = fP_j: \quad \text{Backward slip}. \tag{14}
\end{align*}
\]

2.4. Determining the reduced loading vector \( f_w^c(t) \)

The variable \( f^w_0 \) is previously mentioned as the contact nodal forces that can be obtained from the external loading when the contact displacements are all zero. The simple way to determine these forces is to superpose the solution of two separate problems: (1) The forces \( f^w_0 \) that would be generated by the external loads if there had been no initial gap \( g = 0 \), if the contact had been incomplete and (2) the forces \( f^w_0 \) that are required to close the initial gap in the absence of external loads, which may be obtained using (10) as

\[
\begin{bmatrix} Q^w_0 \\ P^w_0 \end{bmatrix} = \begin{bmatrix} A & B' \\ B & C \end{bmatrix} \begin{bmatrix} 0 \\ -g \end{bmatrix}. \tag{15}
\]

Thus, the contact nodal forces in the incomplete contact can be stated as

\[
f^w_0 = f^w_0 + f^w_0. \tag{16}
\]

2.5. Solution algorithm

We adopt the numerical technique for Coulomb friction contact problem proposed by Ahn and Barber (2008). They describe an algorithm in which the nodal displacements are initially assumed to have the same values as in the previous time step. The nodes are examined one-by-one in a Gauss–Seidel sense and the state and the nodal displacement at the node under examination are updated in accordance with the above conditions. The algorithm cycles through the entire set of nodes several times and the iteration is terminated when the changes during one such cycle are less than a designated convergence criterion. This algorithm requires the reduced stiffness matrix to be configured as in (10).

3. Results and discussion

Several results are obtained under the cyclic loading state where the external load \( F(t) \) initially increases with time to a maximum value \( F_{\text{max}} \), after which it is monotonically reduced to \( F_{\text{min}} \). This loading/unloading process continues until steady state. The ratio of maximum and minimum load, defined as \( T = F_{\text{max}}/F_{\text{min}} \) will be present to the results. Note that the amplitude of the cyclic load \( F_{\text{amp}} \) is \( (F_{\text{max}} - F_{\text{min}})/2 \) and the mean load \( F_{\text{mean}} \) is \( (F_{\text{max}} + F_{\text{min}})/2 \). The ratio of the amplitude to the mean load \( F_{\text{amp}}/F_{\text{mean}} \) can be expressed as \((1 - T)/(1 + T)\). The coefficient of friction was taken to be \( \mu = 0.15 \) and Poisson’s ratio \( v = 0.25 \) and \( R = 1400a \).

The finite element body is initially established to form the contact stiffness matrix. Fig. 3 shows the finite element mesh generated with axisymmetry. The mesh is arranged radially and four-node isoparametric quadrilateral elements are used. In the vicinity of the contact boundary, the mesh is dense to allow for a sufficiently high degree of freedom to retain high accuracy. There are 19,201 and 18,800 whole nodes and elements, respectively, which results in 38,402 degrees of freedom. However, the size of the contact stiffness matrix that we addressed after static reduction is dramatically reduced, showing 101 and 100 nodes and elements, respectively, and 202 degrees of freedom.

3.1. Validation using the Spence problem during loading/unloading phases

The accuracy of the current numerical model heavily depends on mesh discretization. Fig. 4 shows the typical convergence of energy dissipation of the first cycle \( W_1^e \), and the relative error of the dissipation according to the number of contact nodes when
The results show that the dissipation at the first cycle reduces dramatically after the contact nodes of 80, at which the corresponding relative error is under 0.002. Thus sufficient accuracy is ensured if the number of elements is above 80, except for larger values of $T$. The current numerical model uses more than 100 elements for numerical convergence. In addition, the number of steps in a cycle demands greater than 100 in this model to obtain a good resolution of results. However, care should be taken in order to select an appropriate number of steps because the computational memory size required to store all of the information up to the steady state is considerably large and the calculation speed is also affected.

For validation compared to analytical results given by Spence (1975), the relative region of the stick and slip radius $c/a$ was calculated to be 0.37 for the current friction coefficient and Poisson ratio, which is close to the numerical result of 0.36 obtained in the present study.

Fig. 5 shows the evolution of stick-slip regions and the contact traction distributions during the first cycle of loading and unloading when $T = 0.2$. A single relative stick-slip region evolves in the loading phase, indicating self-similarity. When the load begins to decrease from its maximum value, self-similarity is lost, the original stick regions shrinks, and a forward slip annulus spreads towards the center. At the same time, a region of backward slip forms immediately at unloading and an additional stick region appears between the forward and backward slip regions. During unloading, the total contact area itself decreases monotonically relative to the fully loaded state. A similar trend is already reported (Elaguine et al., 2006; Jelagin and Larsson, 2008).

The space coordinate $r$ is also normalized by the contact radius $a_{\text{max}}$ at the maximum force $F_{\text{max}}$. As shown in Fig. 5, the normal contact between frictional spheres produces a slip annulus, where shear tractions acting in a radial sense arises, perturbing the pressure distribution. When the load increases, the shear traction satisfies the inequality of the Coulomb friction law, representing that the shear traction is lower than $f_p a$, in the stick region and equal to $f_p a$ in the slip region. During unloading, a similar trend of contact tractions of the loading phase occurs in the inner stick and slip region but the shear traction at the edge of contact area shows negative values in the backward slip region.

3.2. Contact behavior of the reloading/unloading process

The punch is reapplied into the half plane by a normal force up to $F_{\text{max}}$ and this force is then reduced to $F_{\text{min}}$. This loading/unloading process is repeated until there is a steady state. According to the ratio of $T$, the detailed contact features change since the minimum load $F_{\text{min}}$ changes. However, qualitatively similar results are obtained. We select a case of $T = 0.2$ as a typical contact phenomenon.

The evolution of the slip, stick and separation regions and the normalized contact traction distributions during the reloading phases after ten cycles of reloading and unloading are shown in Fig. 7. The interior stick region shrinks while a forward slip region

![Fig. 3. Finite element meshes for a whole part and a magnified contact region, respectively, before static reduction.](image)

![Fig. 4. Convergence of energy dissipation at the first cycle $W_{\text{diss}}$ (solid line) and the relative error (dotted line) according to the number of contact nodes when $T = 0.2$.](image)

![Fig. 5. Evolution of four contact states and normalized contact traction distributions in the first loading/unloading phase for $T = 0.2$. The contact states represent stick (1), separation (2), forward slip (3), and backward slip (4).](image)

![Fig. 6. The normalized contact pressure $p_{\text{norm}}(r)$ and contact shear traction $q_{\text{norm}}(r)$ distribution during the loading/unloading phases at $T = 0.6F_{\text{max}}$. The contact normal pressure $p_0$ and contact shear traction $q_0$ is normalized by the peak pressure in the frictionless Hertz contact problem for the same radius of indenter and the maximum force $F_{\text{max}}$, which is defined as $p_0 = \frac{3F_{\text{max}}}{2\pi a_{\text{max}}^3} = \left(\frac{6F_{\text{max}}E}{\pi R^2}\right)^{1/3}$.](image)
spreads toward the center. This contrasts with the first loading case shown in Fig. 5. The evolution of the contact status during unloading after ten cycles shows that four central zones (stick/forward slip/stick/backward slip) occur, similar to the initial unloading case shown in Fig. 5. However, the forward slip region between the inner and outer stick regions decreases dramatically.

Fig. 8 shows the normalized contact pressure and contact shear distribution during the reloading/unloading phase after ten cycles of reloading and unloading at $\frac{F(t)}{F_{\text{max}}}=0.6$. In the reloading phase, the shear traction in the stick region is lower than $p_A$ but decreases considerably and then increases abruptly in the vicinity of the edge of the inner stick region, approaching the value of $p_A$. When the load decreases, the contact tractions in the first cycle and 10 cycles behave in a similar manner except that the variation range of the shear traction in the inner and outer stick region and the inner forward slip region extends and shrinks, respectively.

The slip region in the unloading phase eventually decreases as the reloading and unloading evolves. In the long run, the energy dissipation does not change after cycling and the steady state is obtained at this cycle. In the next section, the specific behavior of dissipation is scrutinized. Fig. 9 shows the evolution of contact states during reloading and unloading in the steady state. During the reloading, the evolution of contact states is similar to the case shown in Fig. 7. For unloading in the steady state, the slip region between the inner and outer stick region described in Fig. 7 exists during the beginning of the unloading process but disappears early in the process, eventually leaving the two contact regions (stick/backward slip).

Fig. 10 shows the normalized contact pressure and the contact shear distribution during reloading/unloading in the steady state.
for stick and backward slip, as shown in Fig. 9. The shear traction in the stick region is zero at the origin and approaches a maximum value, which is less than \( f_{p_0} \), and then decreases to the negative values, which corresponds to the limit of slip traction \(-f_{p_0}\).

### 3.3. Frictional energy dissipation

To determine the dependence on \( T \), which is obviously very marked, we define the normalized dissipation per cycle in the first cycle and the steady state as

\[
W_{1st} = \frac{W_{1st}}{W_{ss}} \quad \text{and} \quad W_s = \frac{W_s}{W_{ss}}
\]

where \( W_{1st}, W_s, W_{ss} \) are the dissipation of the first cycle at a given \( T \), the steady state dissipation at a given \( T \), and the steady state dissipation when \( T = 0 \), respectively. Fig. 11 shows the normalized dissipation per cycle in the first cycle and the steady state \( W_{1st}, W_s, \) respectively, as a function of load amplitude \( F_{\text{amp}}/F_{\text{mean}} \). For \( W_s \) in small amplitude of cyclic load, the best fit values from power-law equations show the exponent is close to 2 and not 3 as suggested by theories of dissipation for elastically similar half-spaces (Johnson, 1961; Putignano et al., 2011). The normalized dissipation \( W_{1st} \) also shows that the difference between the steady state and the first cycle increases as \( T \) decreases, meaning that the first cycle dissipation is close to the steady state with a larger load amplitude. When \( F_{\text{amp}}/F_{\text{mean}} \) is small, i.e., a small load amplitude, the dissipation difference between the first cycle and the steady state is large. When the amplitude of the load is large, the dissipation of the first cycle is close to that of the steady state.

The energy dissipation per cycle decreases monotonically with each cycle since the size of the cyclic slip zones in the steady state is reduced compared with the earlier loading phase. The theoretical steady state is almost certainly only reached asymptotically except for \( T = 0 \) and \( T = 1 \) and hence the numerical convergence depends on the tolerance. According to the previous work, Ahn and Barber (2008), there was a region that always slipped in the same direction, so that it could not do so on the steady state, but the amount of slip in that zone decreased geometrically with each cycle. Thus, the total amount of slip was bounded, but the steady state was only approached asymptotically. Fig. 12 shows the number of cycles for \( T = 0.2 \) against error tolerance on a log–log scale. The error tolerance is defined as \( (W_{i-1} - W_i)/W_i \), where \( W_i \) is the dissipation at the cycle of \( i \). It is noticed that the number of cycles against the tolerance on a log–log scale is a straight line, indicating that if the numerical scheme permitted arbitrarily small tolerance, the number of cycles required would become arbitrarily large.
4. Conclusions

A numerical model was developed to investigate an axisymmetric Hertzian frictional contact under cyclic loading. By introducing a “static reduction” technique, a more computationally efficient approach than the standard finite element approach, with respect to memory and speed, was obtained. The results show that as the amplitude of the load decreases, the difference in dissipation between the first cycle and the steady state increases, leading to a power-law amplitude of 2. In the transient contact process, the contact status in the initial period of unloading changes considerably, approaching the steady state after the contact system experiences shakedown.

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (Grant no. 2012R1A1A2042106). The authors are grateful for the valuable suggestions of Professor J. R. Barber of the University of Michigan, Ann Arbor.

References