Stick–slip oscillations: Dynamics of friction and surface roughness

M. T. Bengisu
General Motors Corporation, GM Technical Center, Warren, Michigan 48090

Adnan Akay
Mechanical Engineering Department, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

(Received 11 May 1997; accepted for publication 2 October 1998)

While its classical model is relatively simple, friction actually depends on both the interface properties of interacting surfaces and on the dynamics of the system containing them. At a microscopic level, the true contact area changes as the surfaces move relative to each other. Thus at a macroscopic level, total friction and normal forces are time-dependent phenomena. This paper introduces a more detailed friction model, one that explicitly considers deformation of and adhesion between surface asperities. Using probabilistic surface models for two nominally flat surfaces, the stick–slip model sums adhesive and deformative forces over all asperities. Two features distinguish this approach from more traditional analyses: (i) Roughness distributions of the two interacting surfaces are considered to be independent, (ii) Intersurface contacts occur at both asperity peaks, as in previous models, and on their slopes. Slope contacts, in particular, are important because these oblique interactions produce motion normal to the plane of sliding. Building the model begins by analyzing local friction forces as composites of resistance to elastic deformation and shear resistance arising from adhesion between asperity surfaces. By extending the expressions obtained for normal and tangential friction forces over the macroscopic surfaces, the model then describes the stick–slip behavior frequently observed in dynamic systems and permits simulating a rigid body on a moving platform. Numerical results for several surface and system parameters illustrate both time-dependent and time-averaged frictional forces. These analyses also show that, although total averaged friction remains constant with respect to sliding velocity for the cases considered, the relatively small deformation component exhibits resonancelike behavior at certain speeds. Stick–slip occurs only within a narrow range around these critical speeds of a system. External damping can prevent stick–slip motion, and both deformative and adhesive frictional forces must be present for it to occur at all. [S0001-4966(99)03701-7]

PACS numbers: 43.40.Ga, 43.40.Dx [CBB]

INTRODUCTION

In modeling friction-induced vibration and noise problems, friction force is often treated phenomenologically, as a function of relative velocity between surfaces. These functional relationships mostly originate from measured values of coefficient of friction on test devices. Friction values obtained from such measurements are normally averaged over time and surface areas on which they are measured.

Traditionally, functional relationships between friction force and velocity are confined only to the direction of sliding, neglecting any contribution of normal components of contact forces during sliding. In reality, sliding between, even nominally flat, surfaces also develops a time-dependent normal force component and, therefore, a response in a direction normal to sliding. The forces developed in the normal direction, combined with the continuous change of true contact area during sliding, make friction force to depend also on system dynamics. Such interaction between friction and the response of the system within which it exists can not be modeled through the traditional use of coefficient of friction.

Friction and normal forces develop at the true contact areas between surfaces. Between two nominally flat surfaces, true contact takes place among the asperities. Each asperity adheres and deforms as it slides over another. The direction and magnitude of the deformation and adhesion forces at each contact change during sliding. Thus the topography of the surfaces and the dynamic response of the system together determine the distribution of the contact positions and direction of the forces at each contact. Modeling of friction within a dynamic system thus involves solving coupled equations that describe the system dynamic response and the distribution of contact forces during sliding. Such a combination relates microscopic-scale contact processes to the macroscopic-scale system response.

Many physical and chemical processes contribute to friction. They take place primarily at or near the true contact areas on the sliding surfaces. The vector sum of resistive forces at the contact areas constitutes the friction force. Although a detailed description of all the processes that contribute to friction is beyond the scope of this paper and the current state of the art, it is possible to describe some components of friction and relate them to dynamic behavior of the system in which they exist. We present a model that includes two such dynamic processes that contribute to friction: elastic deformation of asperities and adhesion between them.

Many of the previous models of contact between nomi-

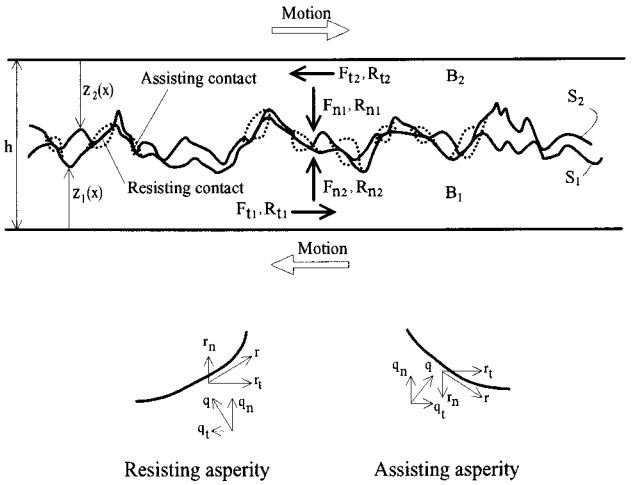

FIG. 1. Cross section of two rough surfaces in contact. F and R represent sums of the local deformation and adhesion forces, respectively. Subscripts n,t denote normal and tangential components. Local contact forces f and r are decomposed with respect to local coordinates.

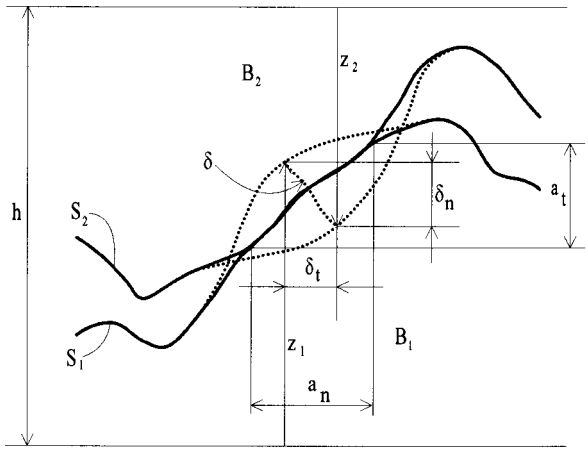

FIG. 2. Cross section of a typical asperity contact with projections of deformation and contact areas. Solid lines represent deformed surfaces and dotted lines represent undeformed surfaces.

nally flat surfaces utilize the assumption that contact between two rough surfaces can be described by contact between a rough surface and a smooth one.[1–6] This assumption dictates that contacts may occur only at surface peaks with contact areas parallel to the mean plane of the surface. In reality, however, contacts between rough surfaces also occur obliquely. In particular, for surfaces with uncorrelated surface asperity distributions, the probability for contacts to occur at asperity peaks is very small.

Oblique contacts couple sliding motion to normal motion. The slope at an oblique contact partitions the contact forces to their components in the direction of sliding and normal to it. As the contact position between two asperities moves, its slope also changes, altering the direction of the contact forces and their projections onto friction and normal forces. Thus as a result of oblique contacts between asperities, sliding motion between two surfaces can also generate oscillations in the normal direction.

The kinematic relationship between friction, $\mathscr{F}_t$, and normal, $\mathscr{F}_n$, forces and the contact area projections resulting from oblique contacts was previously developed by the authors for generalized contact forces in two dimensions,[7]

$$\mathscr{F}_t(t)=F_t+R_t=\frac{A_t}{A_n}\mathscr{F}_n+\left(1+\frac{A_t^2}{A_n^2}\right)R_t, \tag{1}$$

$$\mathscr{F}_n(t)=F_n+R_n. \tag{2}$$

Sums of the tangential components of all local deformation forces, F_t, and adhesion forces, R_t, over the interface constitute the total friction force. Similarly, the corresponding normal components, F_n and R_n, make up the normal contact force at the interface. The sign convention regarding their directions is shown in Fig. 1.

The model in this paper, based on Eq. (1), relates the time-dependent contact area projections, $A_n(t)$, $A_t(t)$, to the motion of the dynamic system through a set of differential equations (Sec. I). The model also expresses contact forces in Eq. (1), F due to deformation and R due to adhesion, in terms of contact area projections (Sec. II). Simultaneous solution of differential equations for the contact areas and for the motion of the bodies then yield both friction $\mathscr{F}_t$ and normal $\mathscr{F}_n$ forces as well as the response of the system to friction (Sec. III).

I. DYNAMICS OF CONTACT AREAS

True contact areas change as a result of both relative motion of surfaces and deformation of asperities during this motion. For example, time-rate of deformation, $\dot{\delta}$, of material point and area, a, at each contact depend on the relative motion of two surfaces through a kinematic relationship,[7]

$$\begin{aligned} -\dot{h}a_n^i+\dot{s}a_t^i&=\dot{\delta}_t^i a_t^i+\dot{\delta}_n^i a_n^i, \quad i=1,...,k,\\ -\dot{h}a_n^j-\dot{s}a_t^j&=\dot{\delta}_t^j a_t^j+\dot{\delta}_n^j a_n^j, \quad j=1,...,l, \end{aligned} \tag{3}$$

where indices i and j indicate resisting and assisting contacts, respectively, and subscripts t and n refer to components in tangential and normal directions to the mean plane. (The forces that result from elastic deformation of asperities either assist or resist the relative tangential motion, depending on the slope of the contact area relative to the direction of motion. *Resisting contacts* have a tangential component that opposes relative motion. *Assisting contacts* possess tangential components in the same direction as the relative motion.[7]) The relative velocity of the surfaces has components $\dot{h}$ in normal direction and $\dot{s}$ in tangential direction, as indicated in Fig. 2. *Normal contact* and *tangential contact areas*, a_n and a_t, refer to local contact area projections that are parallel and perpendicular to the mean planes of the surfaces, respectively, as illustrated in Fig. 1.

The total numbers of resisting and assisting contacts, k and l, in Eqs. (3), can be approximated using the results of the Greenwood and Williamson (1967) model for random surfaces,

$$k=\gamma A_n^r, \quad l=\gamma A_n^a, \tag{4}$$

where γ is a constant and superscripts (r,a) denote resisting and assisting asperities.

Deformation velocities $\dot{\delta}_n$ and $\dot{\delta}_t$ in Eq. (3) can also be expressed in terms of contact areas through contact deforma-

tion models. In elastic contact problems involving rough surfaces, Hertz contact theory is used commonly with the assumption that asperities are spherical and contact is elastic;[9] the contact area, a, between two spherical objects of radius ρ_1 and ρ_2, is expressed as

$$a=\pi\rho\delta, \tag{5}$$

where $1/\rho=1/\rho_1+1/\rho_2$ and δ is the maximum value of deformation.

Differentiating Eq. (5) with respect to time and substituting its normal and tangential components into Eqs. (3) yields a direct relationship between local contact areas and relative velocities. Summing the resulting equations over the total numbers (k and l) of resisting and assisting local contact areas, respectively, leads to expressions describing the dynamic relationship between the total contact areas and the normal and sliding velocities,

$$\begin{aligned} -\dot{h}A_n^r+\dot{s}A_t^r &\cong \frac{1}{k\pi\rho}\,(A_t^r\dot{A}_t^r+A_n^r\dot{A}_n^r),\\ -\dot{h}A_n^a-\dot{s}A_t^a &\cong \frac{1}{l\pi\rho}\,(A_t^a\dot{A}_t^a+A_n^a\dot{A}_n^a), \end{aligned} \tag{6}$$

where A is the total contact area summed over all asperities on one of the surfaces S_1. (The combined radius ρ now represents the combination of the average asperity radii on surface S_1: $\rho_1=[1/(k+l)]\Sigma_{i=1}^{k+l}\rho_1^i$.)

Further simplification of Eq. (6) may be made by expressing A_t in terms of A_n through space-averaged (but time-dependent) slope of contact areas. Such a relationship between the sums of the normal and tangential projections of contact areas may be obtained by approximating deformation velocity $\dot{\delta}$ at each contact by a uniform velocity distribution. This approximation, also used in the development of Eq. (3), implies that the local contact areas are plane surfaces and the contact slope $Z'_\iota=\partial Z/\partial\chi_1^\iota$ $(\iota=i,j)$ is constant throughout a given (local) contact surface. With these assumptions the tangential components of the contact areas can be expressed in terms of the normal components,

$$a_t^i=Z'_i a_n^i, \quad a_t^j=-Z'_j a_n^j, \tag{7}$$

where the slopes of contact areas are with respect to the direction of motion.

Summing expressions in Eq. (7) over the interface gives the following relationships between summations of tangential and normal contact areas:

$$\begin{aligned} A_t^r=\sum_{i=1}^{k} a_t^i &\cong \left(\frac{1}{k}\sum_{i=1}^{k} Z'_{1i}\right)\left(\sum_{i=1}^{k} a_n^i\right)=\langle Z'_+\rangle A_n^r,\\ A_t^a=\sum_{j=1}^{l} a_t^j &\cong -\left(\frac{1}{l}\sum_{j=1}^{l} Z'_{1j}\right)\left(\sum_{j=1}^{l} a_n^j\right)=-\langle Z'_-\rangle A_n^a. \end{aligned} \tag{8}$$

The average resisting and assisting contact slopes, $\langle Z'_+\rangle$ and $\langle Z'_-\rangle$, in Eqs. (8) are the expected values of the positive and negative contact slopes and they can be calculated for given probability densities of the asperity slopes.

Finally, substituting the approximate relationships given by Eqs. (4) into Eqs. (6) yields two ordinary differential equations for the total resisting and assisting normal contact areas in terms of relative velocities between the surfaces,

$$\begin{aligned} \alpha(\langle Z'\rangle^2+1)\dot{A}_n^r &= -(\alpha\langle Z'\rangle\langle\dot{Z}'\rangle+\dot{h}-\langle Z'\rangle\dot{s})A_n^r,\\ \alpha(\langle Z'\rangle^2+1)\dot{A}_n^a &= -(\alpha\langle Z'\rangle\langle\dot{Z}'\rangle+\dot{h}+\langle Z'\rangle\dot{s})A_n^a, \end{aligned} \tag{9}$$

where $\alpha=1/\gamma\pi\rho$ and $\langle Z'\rangle=\langle Z'_+\rangle=-\langle Z'_-\rangle$ for a random distribution of surface roughness.

The solutions of Eq. (9) for resisting and assisting contacts are combined to give the projections of total contact area required for the friction force in Eq. (1),[7]

$$A_n=A_n^r+A_n^a=\sum_{i=1}^{k} a_n^i+\sum_{j=1}^{l} a_n^j. \tag{10}$$

The tangential component of the total contact area, A_t, on the other hand, is defined as the difference between resisting and assisting contacts. (Considering only the resisting contacts overestimates the friction force because the presence of assisting contacts offsets some of the resistance.)

$$A_t^r-A_t^a=\sum_{i=1}^{k} a_t^i-\sum_{j=1}^{l} a_t^j. \tag{11}$$

Equations (8) and (11) lead to an expression that relates the tangential component A_t to the normal component obtained from Eqs. (9),

$$A_t=\langle Z'\rangle(A_n^r-A_n^a). \tag{12}$$

Although the expression in this section are derived for surface S_1, similar ones can be written for S_2.

In addition to A_t from Eq. (12) and A_n from the solution of Eqs. (9), calculation of friction force in Eq. (1) requires expressions for the contact force components F_n, R_n, and R_t. In the next section expressions are developed for deformation and adhesion forces at each contact and related to the sum of true contact areas A.

II. CONTACT AND FRICTION FORCES

Because of the cause-and-effect relationship between elastic deformation forces and contact areas, deformation force at each asperity contact can be expressed in terms of the resulting contact area. Similarly, because adhesion forces depend independently on both the normal pressure, or the deformation force, and the contact area, they also can be expressed in terms of contact areas.

A. Deformation forces

In cases such as metals, where the local tangential adhesive force r is typically smaller than the normal force f by an order of magnitude or more,[8] its contribution to elastic deformation of an asperity may be neglected to simplify the calculations. Thus only the local normal contact forces, f, cause elastic deformation and they are obtained using Hertz contact theory with the assumption that asperity tips are spherical.[1,3,9]

The physics of the contact problem treated here also requires consideration of energy dissipation from the contact area during deformation. Because use of complex elasticity modulus in nonlinear transient problems leads to physically

unrealizable results,[10] energy dissipation during elastic deformation is modeled here as proportional to the deformation rate of an asperity. Such a consideration is equivalent to the loss of deformation energy by means of propagation of elastic waves into the bulk of the contacting bodies, i.e., radiation loss away from contact areas. Hence, the sum of the Hertz contact force and a dissipation force that depends on the rate of deformation constitutes the elastic deformation force

$$\mathrm{f}=\mathscr{K}a^{3/2}+\eta\dot{\epsilon}. \tag{13}$$

The constant $\mathscr{K}$ is defined as

$$\mathscr{K}=\frac{4E}{3\rho\pi^{3/2}},$$

where η is the loss factor of the material and ρ is composite asperity radius as defined before. E is the composite modulus of elasticity related to the moduli of the two materials and their Poisson's ratios ν_1 and ν_2,

$$\frac{1}{E}=\frac{1-\nu_1^2}{E_1}+\frac{1-\nu_2^2}{E_2}. \tag{14}$$

Recalling that the contact surfaces are assumed to be planar, such that $a^2=a_t^2+a_n^2$ and $a_t/a_n=Z'$, the rate of deformation of a material at a contact surface can be expressed as

$$\dot{\epsilon}=\frac{1}{2\pi\rho}\frac{d}{dt}\,(a_t^2+a_n^2)=\frac{1}{\pi\rho}\,\dot{a}a_n(1+Z'^2)^{1/2} \tag{15}$$

and

$$a^{3/2}=a a_n^{1/2}(1+Z'^2)^{1/4}. \tag{16}$$

After substituting Eqs. (15) and (16) into Eq. (13), the sum of the normal components of the deformation contact forces at the interface can be expressed as

$$F_n=\frac{\mathscr{K}}{\gamma^{1/2}}\,A_n\left(1+\frac{1}{4}\,\langle Z'^2\rangle+\cdots\right)+\frac{\eta}{\pi\rho}\,\dot{A}_n\left(1+\frac{1}{2}\,\langle Z'^2\rangle+\cdots\right), \tag{17}$$

where

$$\langle Z'^2\rangle=\frac{1}{k}\sum_i^k Z_i'^2=\frac{1}{l}\sum_j^l Z_j'^2$$

is the mean-square value of either the positive or the negative contact slopes. For asperity profiles such that $a_t/a_n<1(Z'^2\ll 1)$, the series of even moments in Eq. (17) can be truncated after the first term.

B. Adhesive contact forces

The local adhesive forces are considered here to be similar to the rheological shear resistance of the third phase that forms between the contact surfaces of two materials. Following Kragelsky *et al.*,[8] the local adhesive contact force, r, is expressed with a binomial expression, as a combination of two shear terms

$$\mathrm{r}=\tau a+\beta\mathrm{f}, \tag{18}$$

where the constants τ and β are determined empirically. (The signs of τ and β depend on the relative velocity at the interface and are negative for positive relative velocity.) The adhesion force r in Eq. (18) is in the tangential direction to the local contact surface. Its first part is independent of normal contact force, whereas the second part is proportional to the contact pressure and contact area: $\mathrm{f}=\int q\,da$.

By separating Eq. (18) into its normal and tangential components with respect to the mean planes and summing over all resisting and assisting contact areas, the normal and tangential components of the total adhesive contact force can be expressed in terms of sums of contact area projections,

$$R_n=-\tau A_t-\beta F_t, \quad R_t=\tau A_n+\beta F_n. \tag{19}$$

R_n and R_t are uncoupled because of the orthogonality of the local elastic deformation and adhesion forces. In cases where, for example, the influence of adhesive forces on elastic deformation is not negligible, R_n and R_t will be coupled. The negative signs in the first of Eqs. (19) are a consequence of the convention of directions shown in Fig. 1.

C. Friction and normal force

Friction force, defined here as the force that resists sliding motion, is obtained by summing the tangential forces at the interface. Similarly, the normal force is the sum of all the forces in a direction normal to the interface. Expressions for friction force and normal force are obtained by substituting Eqs. (19) into Eqs. (1) and (2). The results explicitly show the relationships of the friction and normal forces to deformation and adhesive forces and to the normal and tangential projections of the contact areas,

$$\begin{aligned}\mathscr{F}_t(t)&=\left(\frac{A_t}{A_n}+\beta\right)F_n+\tau A_n,\\ \mathscr{F}_n(t)&=\left(1-\beta\,\frac{A_t}{A_n}\right)F_n-\tau A_t,\end{aligned} \tag{20}$$

where, F_n, sum of the normal components of deformation forces, is given by Eq. (17). Thus Eqs. (20), together with Eq. (17), define the friction force, and the resulting normal force, in terms of contact areas and even-movements of the contact slope distribution.

F_n, A_n, and A_t are based on statistical distributions of the slopes of asperities as described, earlier; τ and β are experimentally determined constants. Both $\mathscr{F}_t(t)$ and $\mathscr{F}_n(t)$ are time dependent because of the changes in F_n, A_n, and A_t due to relative motion. Thus true friction force is obtained in conjunction with the dynamic response of the system within which it exists, as shown in the next section.

III. FRICTION FORCE AND SYSTEM RESPONSE

An example of the interaction of friction force with system dynamics is demonstrated with a simple, idealized dynamic system: a rigid block (with deformable surface asperities) connected to a linear spring and a viscous damping element, is subject to frictional force at its interface with a flat platform that moves at a constant speed, as depicted in Fig. 3. The rigid block is free to move both in normal and

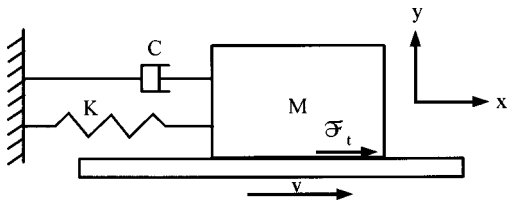

FIG. 3. Block on a moving platform. Block is free to move in x and y directions.

tangential directions with respect to the platform but without rotation. Both the block and the platform are assumed to have nominally flat surfaces, each with a random asperity height distribution statistically independent of the other.

The equations of motion of the dynamic system are written in terms of the coordinates x and y which also describe its displacements in tangential and normal directions with respect to the moving platform,

$$M\ddot{x}+C\dot{x}+Kx=-\mathscr{F}_t,$$
$$M\ddot{y}=-Mg+\mathscr{F}_n. \tag{21}$$

Parameters M, C, and K describe the mass, viscous damping, and the stiffness, respectively. The differential Eqs. (9) for the contact areas are rewritten by substituting for $\dot{h}=\dot{y}$ and $\dot{s}=\dot{x}-v$, where, v is the speed of the platform,

$$\alpha(\langle Z'\rangle^2+1)\dot{A}_n^r=-\{\alpha\langle Z'\rangle\langle\dot{Z}'\rangle+\dot{y}-\langle Z'\rangle(\dot{x}-v)\}A_n^r,$$
$$\alpha(\langle Z'\rangle^2+1)\dot{A}_n^a=-\{\alpha\langle Z'\rangle\langle\dot{Z}'\rangle+\dot{y}+\langle Z'\rangle(\dot{x}-v)\}A_n^a. \tag{22}$$

Numerical solutions of Eqs. (21) and (22) are obtained using expressions for friction and normal forces in Eq. (20) and for the contact areas in Eqs. (10) and (12). The mean value of the contact slopes, $\langle Z'\rangle$, and its relation to separation, h, between the surfaces are obtained from the description of surface roughnesses for the rigid body and the platform as shown in the next section.

A. Surface model

Calculating contact areas in Eq. (22), and normal contact force F_n in Eq. (17), requires mean value of contact slopes and their even moments for each surface. Randomly varying, uncorrelated continuous functions Z_1' and Z_2', each with a Gaussian distribution, represent the slopes of asperities on each surface as a function of their separation, h. Because the subset of slopes at the contact areas is also Gaussian, the mean value of the positive contact slopes can be found using Z_1' and Z_2'. Then, for slopes ranging between zero and some maximum contact slope, $Z'_{\max}$,

$$\langle Z'\rangle=\frac{1}{\sigma\sqrt{2\pi}}\int_0^{Z'_{\max}}Z'e^{-Z'^2/2\sigma^2}\,dZ', \tag{23}$$

where σ is the standard deviation. The even moments of the slopes are

$$\langle Z'^m\rangle=\frac{1}{\sigma\sqrt{2\pi}}\int_0^{Z'_{\max}}Z'^m e^{-Z'^2/2\sigma^2}\,dZ' \quad m=2,4,..., \tag{24}$$

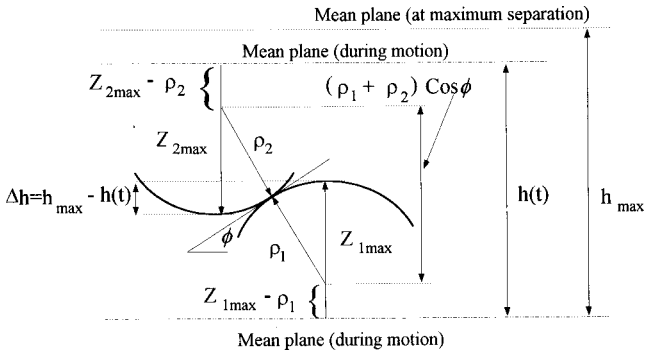

FIG. 4. Schematic representation of the highest asperities in contact and the relationship of the maximum contact slope to separation.

where the value of the upper limit $Z'_{\max}$ changes with interference between the surfaces. These expressions are the same for both surfaces, but with different values of $Z'_{\max}$ for each. The integral in Eq. (23) can be evaluated in closed form to give

$$\langle Z'\rangle=\frac{\sigma}{\sqrt{2\pi}}(1-e^{-Z'^2_{\max}/2\sigma^2}). \tag{25}$$

The second and higher moments about origin have negligible effects on the normal force.

The mean value of contact slopes, $\langle Z'\rangle$, relates to separation h through the maximum contact slope $Z'_{\max}$ in expression (25) which varies with separation during sliding. When separation is large, contacts occur closer to asperity tips. At the maximum value of separation, contact area is parallel to the mean plane. For smaller values of separation, contacts are oblique and contact slopes are larger. Accordingly, the mean value of the contact slopes is also a function of separation between the surfaces.

In cases where the relationship between the maximum contact slope and separation can be expressed explicitly, e.g., for spherically shaped asperities, it is possible to evaluate $\langle Z'\rangle$ in Eq. (25). For spherical asperities, the maximum contact slope occurs on the tallest asperities of the surfaces. The asperities located farthest from the mean planes on each surface define the maximum value of the separation, $h_{\max}$ (or the minimum value of geometric approach) between the two surfaces. With these considerations, a relationship can be defined for separation h in terms of average asperity size on each surface. Using the geometric relationship for such a contact condition as illustrated in Fig. 4, separation between surfaces is expressed as

$$h=Z_{1\,\max}-\rho_1+(\rho_1+\rho_2)\cos\,\phi+Z_{2\,\max}-\rho_2, \tag{26}$$

where ρ_1 and ρ_2 represent the average asperity radii for surfaces S_1 and S_2, respectively, $Z_{1\,\max}$ and $Z_{2\,\max}$ are the maximum values of the functions describing the surface profiles, and ϕ is the contact slope as defined in Fig. 4. At the maximum value of h, by definition, the corresponding maximum contact slope is zero, $\phi=0$.

Substituting $\cos\phi\cong1-\phi^2/2$ in Eq. (26) yields $\phi^2=\Delta h/\varrho$, where the value of $\varrho=(\rho_1+\rho_2)/2$ is the average of the mean asperity radius for each surface. The difference between the maximum and instantaneous values of separation is

TABLE I. Numerical values of the parameters used in calculations. The tangential damping value of $C=2$ kg/s is used during dynamic simulations, whereas $C=200$ kg/s is used for averages. The friction pair chosen for all cases is a steel–copper pair having a composite elasticity modulus of 77 800 MPa.

Case	σ (rad)	$(1/2\pi)\sqrt{K/M}$ (Hz)	τ (MPa)	β
1	0.3	10	0	0
2	0.2	10	0	0
3	0.1	10	0	0
4	0.1	31.6	0	0
5	0.1	100	0	0
6	0.1	10	9.8	0.075
7	0.1	31.9	9.8	0.075
8	0.1	100	9.8	0.075

M (kg)	C (kg/s)	R (mm)	η (kg/m^2 s)
10	2–200	0.06	200

$$\Delta h(t)=Z_{1\,\max}+Z_{2\,\max}-h=h_{\max}-h(t).$$

For small values of ϕ, $\phi\cong\tan\phi=Z'_{\max}$. Thus the maximum contact slope can be expressed in terms of separation,

$$Z'_{\max}\cong\sqrt{\frac{\Delta h}{\varrho}}. \tag{27}$$

By replacing $Z'_{\max}$ in Eq. (25) with Eq. (27), the expected value $\langle Z'\rangle$ of contact slopes can be expressed in terms of separation of the surfaces.

The model is now complete and can be solved numerically to investigate the interaction of friction and the dynamic response of the system that contains it.

B. Computations

Dynamic response of the system and the friction force that excites it mutually depend on each other through contact area changes as shown in Eqs. (21) and (22). In the following sections, equations describing the system response and the corresponding friction and normal forces are solved for the cases listed in Table I. The values given in Table I are for steel–copper pairs for which empirical values of the parameters τ and β are available.[8] Surface properties used in the examples reported below come from published surface data;[1] asperity tip radii have a Gaussian distribution with a mean radius of 60 μm and the maximum asperity heights reach 3–4 μm, typical for ground mild steel.[1] Maximum value of average slope is determined by setting $Z'_{\max}$ in Eq. (27) to 3σ.

By numerically integrating the expressions in (21) and (22), first, the time-averaged friction force is examined as a function of platform speed. Then, the time-dependent friction force is analyzed along with the corresponding response of the dynamic system. In both cases, friction force is normalized with respect to normal contact force.

Because of the asymptotic nature of $\langle Z'\rangle$ in Eq. (25), to circumvent computational difficulties, in all the computations reported here contact slope is taken to be zero when relative approach reduces to less than 0.001 μm.

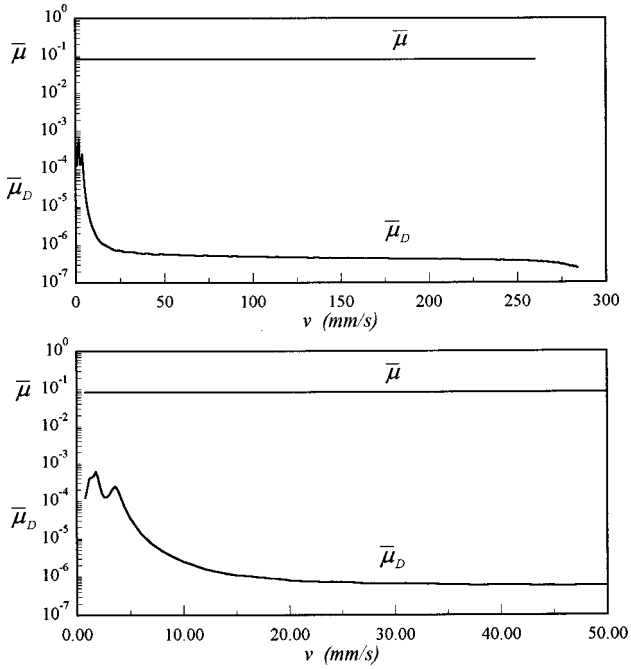

FIG. 5. Change of $\bar{\mu}$ and $\bar{\mu}_D$ with platform velocity.

C. Averaged normalized friction force

Time averaging the ratio of the instantaneous values of friction and normal forces gives $\bar{\mu}$, the time-averaged normalized friction force,

$$\bar{\mu}=\lim_{T\to\infty}\frac{1}{T}\int_0^T\frac{\mathscr{F}_t}{\mathscr{F}_n}\,dt.$$

Because adhesion depends on the deformation force, $\bar{\mu}$ is a complex combination of the similarly time-averaged normalized deformation ($\bar{\mu}_D$) and adhesion ($\bar{\mu}_A$) components. $\bar{\mu}$ reduces to $\bar{\mu}_D$ when constants τ and β are zero. (Use of the term coefficient of friction is avoided because of its diverse connotations and different definitions in the literature.)

An examination of the averaged normalized friction force, $\bar{\mu}$, in Fig. 5 shows that, for cases 3 and 6 in Table I, a negligibly small part of $\bar{\mu}$ results from elastic deformation, leaving adhesion as the dominant source of friction. This is not an unexpected result since the effects of resisting and assisting contacts largely offset each other.

Numerical results show that $\bar{\mu}$ with respect to platform speed is nearly constant, whereas $\bar{\mu}_D$ shows significant variation as a function of platform speed. In particular, at some platform speeds, the deformation component $\bar{\mu}_D$ shows resonancelike peaks. The platform speeds, at which resonancelike peaks occur, increase with tangential natural frequency, f_0, of the dynamic system. This dependence is illustrated in Fig. 6 with plots of $\bar{\mu}_D$ corresponding to $f_0=10$, 3.16, and 100 Hz of the system (cases 3, 4, and 5 in Table I); the speeds corresponding to the peak values of $\bar{\mu}_D$ are tabulated in Table II.

Asperity slope distributions also influence the velocities at which the peak values of $\bar{\mu}_D$ occur. For example, for smoother surfaces which have higher standard deviations, as indicated by Eq. (25), peak values of $\bar{\mu}_D$ occur at higher

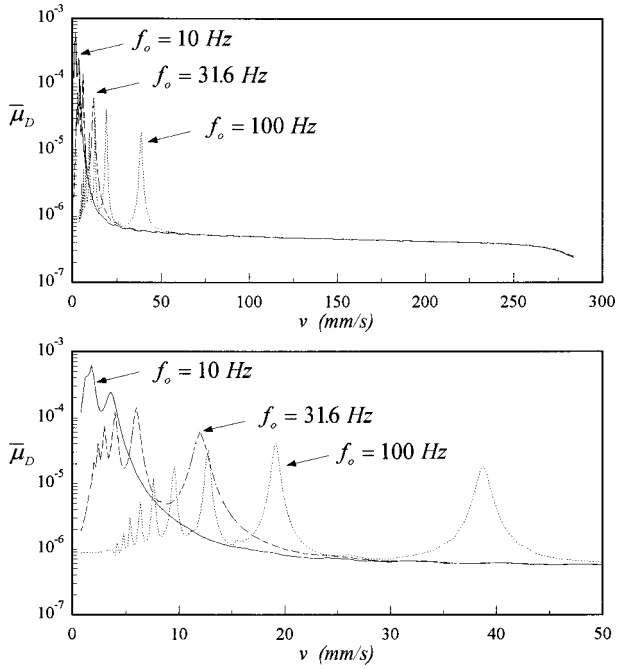

FIG. 6. Variation of $\overline{\mu}_D$ with platform velocity for various tangential natural frequencies of the dynamical system.

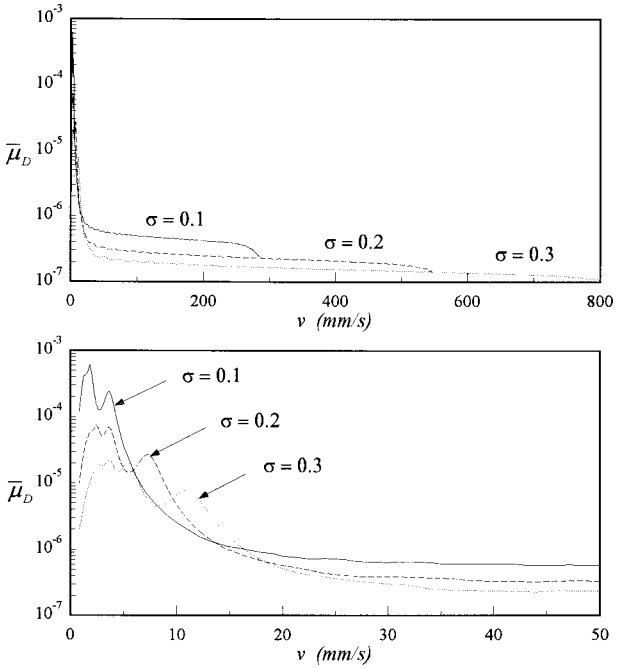

FIG. 7. Variation of $\overline{\mu}_D$ with platform velocity for different standard deviation of surface roughness.

platform velocities; Fig. 7. The parameters used in Fig. 7 and the results tabulated in Table III correspond to cases 1–3 in Table I.

The composite modulus of elasticity of the materials does not have a detectable influence on the variation of $\overline{\mu}_D$ or on the resonancelike peaks at lower platform speeds. However, it has some influence at higher platform speeds.

D. Instantaneous friction and stick–slip oscillations

Stick–slip oscillations of the system occur at platform speeds within a narrow band of the peak values of $\overline{\mu}_D$ shown in Fig. 6. Oscillations at platform speeds corresponding to different peaks have significantly different spectra. Oscillations at platform speeds outside of these bands are sustained, but without stick–slip.

Phase planes and spectra of the motion of the block and the corresponding instantaneous friction force, given in Figs. 8–14, help explain stick–slip oscillations and their relationship with the resonancelike peaks of $\overline{\mu}_D$. The instantaneous normalized friction force, plotted in Figs. 8–14, is defined as $\mu(t)=\mathcal{F}_t(t)/\mathcal{F}_n(t)$.

Figure 8 demonstrates an example (case 6; $\sigma=0.1$) for which the platform speed (2.7 mm/s) is away from the speed bands that lead to peaks. By comparison with the corresponding case given in Fig. 9, Fig. 8 shows a smaller amplitude of oscillations that are not repetitive (transient parts of the solutions are excluded in all the results). Also, the tangential velocity in its phase plane never reaches the velocity of the platform and, thus, does not achieve ''stick'' condition; the relative velocity is always larger than zero. In the same figure, friction force varies through many cycles of oscillations of the dynamic system and exhibits multiple values for a given relative tangential velocity. The spectrum corresponding to the response velocity of the block exhibits several harmonics and side bands, indicative of nonlinearities.

At platform speeds, 3.6, 1.8, and 1.2 mm/s, corresponding to the peak values of $\overline{\mu}_D$ shown in Fig. 6, response of the same system as in Fig. 8 (case 6; $\sigma=0.1$) exhibits higher amplitudes that are periodic. At these speeds both the friction force and response of the mass show that the oscillator reaches the platform speed, achieving the condition of ''stick;'' Figs. 9–11. In terms of phase-plane diagrams, system response shows a doubling and tripling of its period as the platform speed is decreased from 3.6 mm/s to the lower speeds 1.8 and 1.2 mm/s, respectively. The corresponding changes also manifest themselves in the spectrum for each case as additional harmonics and one-half subharmonic. In all cases, friction-relative velocity trajectory follows a differ-

TABLE II. Some platform speeds where averaged normalized deformation friction forces reach their peak values for the surface with $\sigma=0.1$.

f_0 (Hz)	v_1 (mm/s)	v_2 (mm/s)	v_3 (mm/s)	v_4 (mm/s)	v_5 (mm/s)	v_6 (mm/s)
10	3.6	1.8	1.2	⋯	⋯	⋯
31.6	12.0	6.0	4.0	3.0	2.4	2.0
100	38.75	19.15	12.8	9.6	7.6	6.4

TABLE III. Platform speeds where averaged normalized deformation forces reach their peak values for the dynamic system with 10 Hz tangential natural frequency.

σ	v_1 (mm/s)	v_2 (mm/s)	v_3 (mm/s)	v_4 (mm/s)
0.1	3.6	1.8	1.2	⋯
0.2	7.3	3.65	2.43	1.83
0.3	11.0	5.5	3.66	2.75

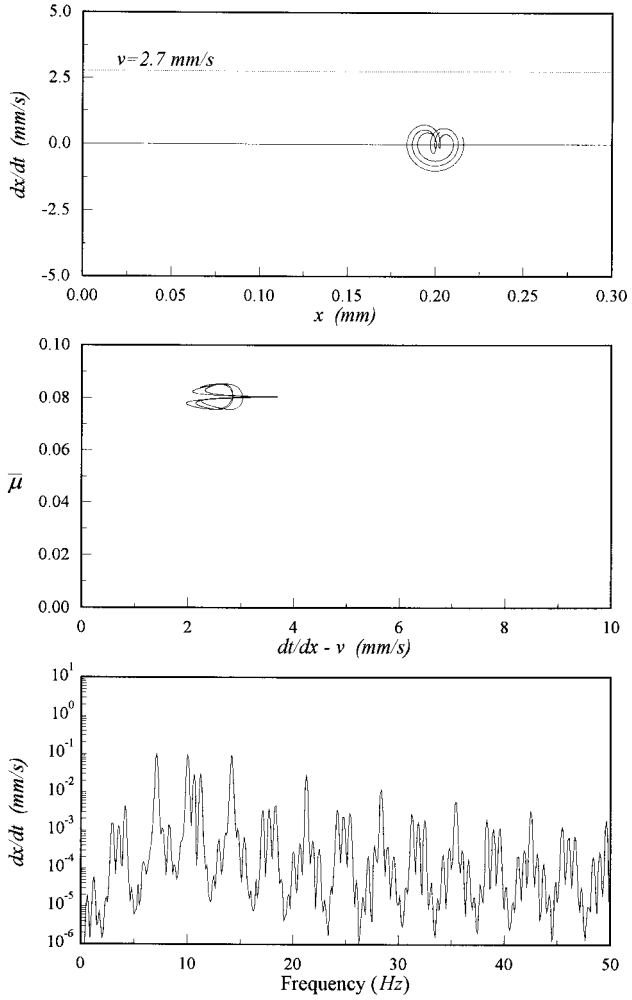

FIG. 8. Response of the dynamic system with a tangential natural frequency of $f_0 = 10$ Hz displays sustained oscillations at a platform speed of 2.7 mm/s.

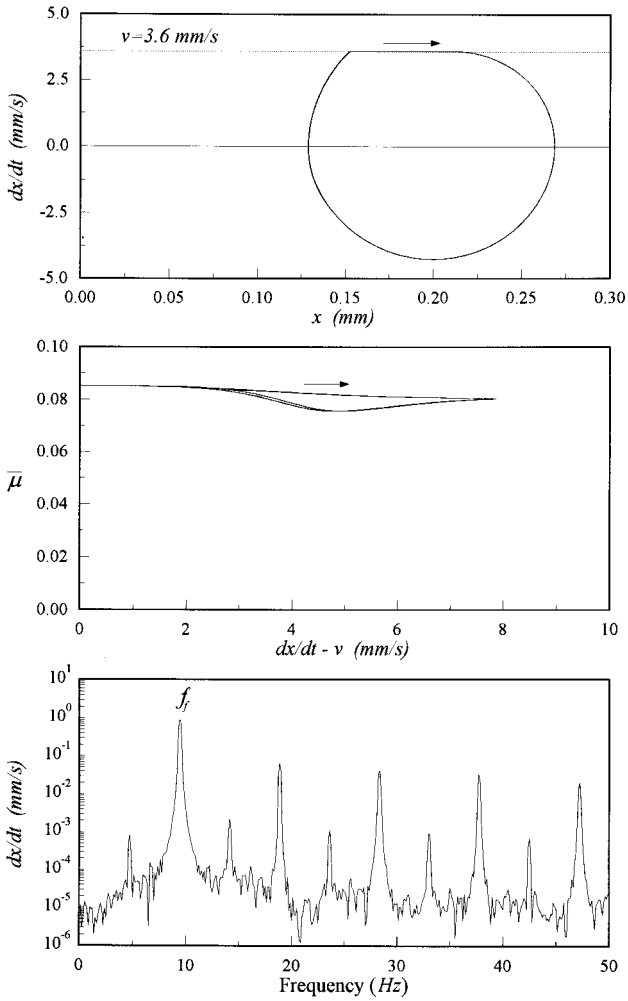

FIG. 9. Response of the dynamic system with a tangential natural frequency of $f_0 = 10$ Hz displays stick–slip oscillations at a platform speed of 3.6 mm/s.

ent path when it moves in the direction of the platform than against it.

A similar, but not easily observed, result relates to the total period of oscillation of the block as shown in the phase-plane plots (and the corresponding frequency spectra). For example, in Fig. 9 (top) the phase plane appears to have a single path where, in reality, it consists of two paths that are very close to each other. Evidence of this is seen in the middle figure where friction force follows two separate paths. This slight variation every other cycle is seen in the tangential velocity spectrum of the mass as a half-frequency subharmonic of the fundamental frequency corresponding to the single loop observed in the phase plot in Fig. 9.

The apparent self-intersection of the trajectories in the phase planes in Figs. 9 through 14 results from the projection of the multidimensional phase-space trajectory onto the tangential velocity-displacement plane of the block; the actual six-dimensional phase-space trajectory of the system does not self-intersect.

The stick–slip behavior of the dynamic system presented here is very much reminiscent of the response of a friction-excited system even when the friction force-velocity relationship is predefined.[11,12]

IV. STICK–SLIP OSCILLATIONS: DISCUSSION

A. Periodicity and nonlinearity

The assumption that the spatial distribution of contact slopes at a given time is a stationary random function implies that the average contact slope, $\langle Z' \rangle$, is independent of the tangential position of the mass with respect to the platform. On the other hand, the average contact slope, $\langle Z' \rangle$, depends on the the separation between the mean planes, as described in Eqs. (25) and (27). As a result, the average contact slope, $\langle Z' \rangle$, remains independent with respect to tangential position of the mass but changes periodically with its normal motion. Thus the influence of asperities on the motion of the system is nearly periodic, albeit nonlinear.

The periodic nature of the average contact slope $\langle Z' \rangle$ leads to nearly periodic solutions of the differential Eq. (9) for A_n^r and A_n^a. Hence, in accordance with Eqs. (10), (12), and (20), the resulting net normal and tangential forces are also nearly periodic, causing nearly periodic motion of the block both in normal and tangential directions. Both the normal and tangential components of the contact force exhibit the same fundamental frequency even though they may have different overall spectra.

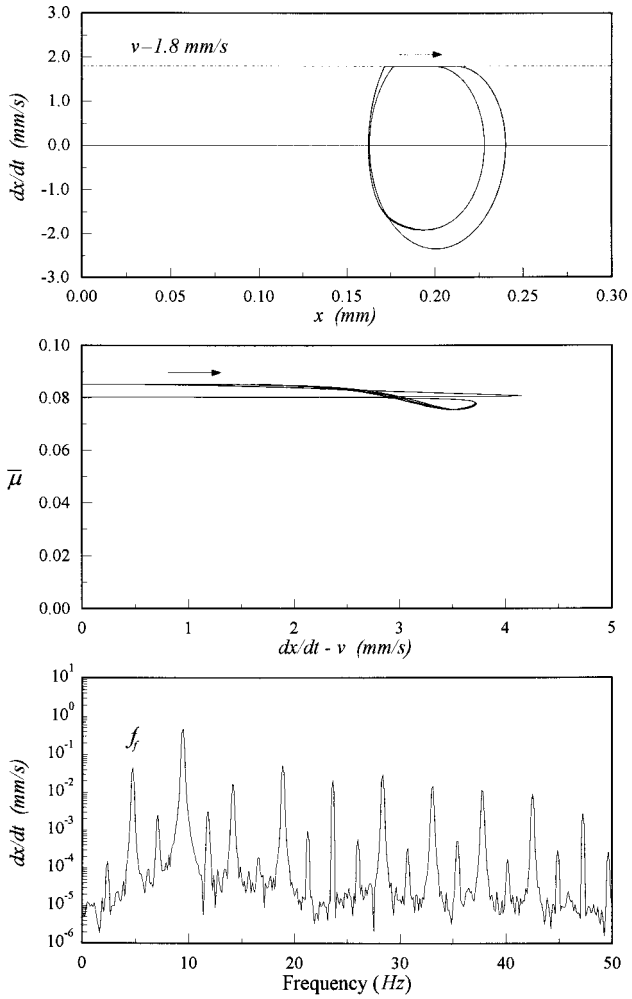

FIG. 10. Response of the dynamic system with a tangential natural frequency of $f_0=10$ Hz displays stick–slip oscillations at a platform speed of 1.8 mm/s.

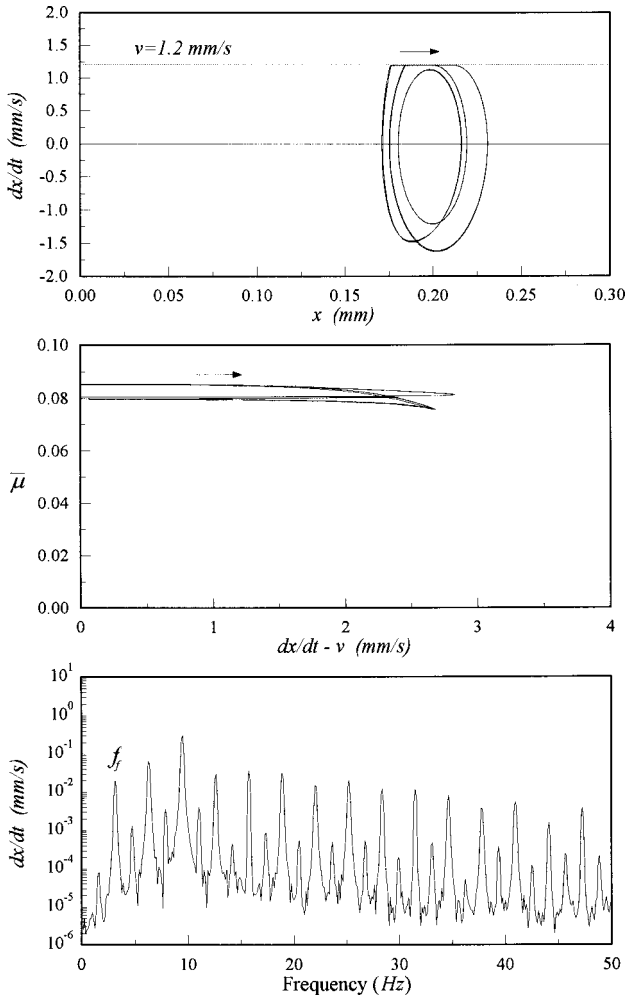

FIG. 11. Response of the dynamic system with a tangential natural frequency of $f_0=10$ Hz displays stick–slip oscillations at a platform speed of 1.2 mm/s.

In cases of nonstationary contact slope distributions, however, average contact slope changes from one cycle to the next as the block and the platform slide against each other.

B. Critical speeds and fundamental frequency

Stick–slip oscillations, and the corresponding peak values of $\overline{\mu_D}$, occur when the tangential natural frequency, f_0, of the system coincides with the fundamental frequency, f_f, of the friction force or one of its harmonics. The fundamental frequency of the friction force approximately relates to the platform speed as

$$f_f \cong (v/\lambda), \tag{28}$$

where λ is the wavelength corresponding to the relative tangential displacement of the block during a cycle of its combined normal and tangential oscillation. The platform speeds at which integer multiples of f_f equal f_0, i.e., $nf_f=f_0$, are referred to as the critical speeds, $v=v_{c_n}$.

The *fundamental critical speed*, v_{c_1}, is the highest platform speed at which stick–slip occurs and the fundamental frequency of the friction force equals the natural frequency $f_f=f_0$. At a lower critical speed, roughly described as v_{c_1}/n, the harmonic multiple of f_f becomes equal to the natural frequency; $nf_f=f_0$; Figs. 9–14.

The wavelength λ represents the average distance between resisting and assisting contacts. Its value can be obtained using the average slope of contacts, $\langle Z' \rangle$, and the normal displacement, or approach Δh, during the oscillations,

$$\lambda = 2\int_0^{\Delta h_{\max}} \frac{dh}{\langle Z' \rangle}, \tag{29}$$

where $\langle Z' \rangle$ is given by Eqs. (25) and (27) and $\Delta h_{\max}$ is the maximum relative approach during the motion of the block. Both $\Delta h_{\max}$ and $\langle Z' \rangle$ depend on surface roughness as well as system response through the solutions to differential equations Eqs. (21) and (22).

Values of λ, plotted in Fig. 15 as a function of maximum relative approach, can be used to predict v_{c_n}. For example, for $\sigma=0.1$, a relative approach of 0.3 μm (the maximum relative approach considered in this paper), indicates a wavelength of $\lambda/2=0.175$ mm. For a system with a natural tan-

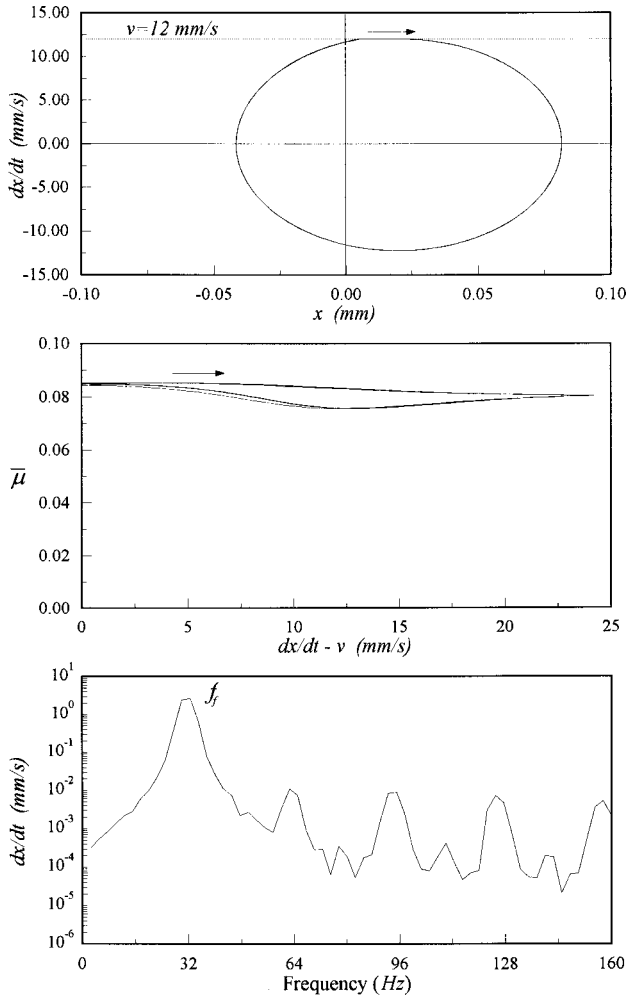

FIG. 12. Response of the dynamic system with a tangential natural frequency of $f_0 = 31.6$ Hz displays stick–slip oscillations at a platform speed of 12 mm/s.

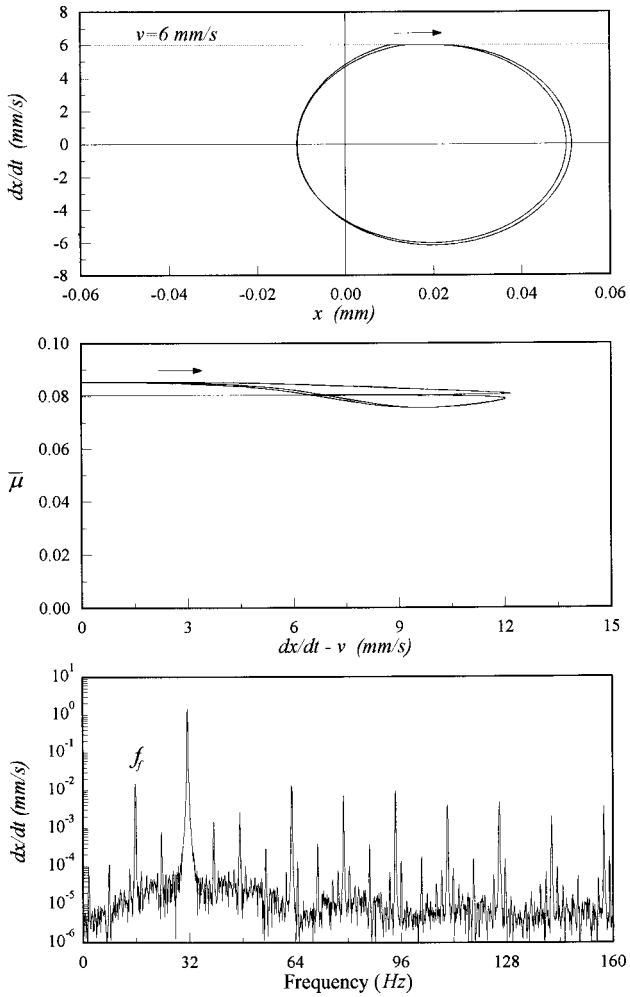

FIG. 13. Response of the dynamic system with a tangential natural frequency of $f_0 = 31.6$ Hz displays stick–slip oscillations at a platform speed of 6 mm/s.

gential frequency of 10 Hz, Eq. (28) indicates the fundamental critical speed to be 3.5 mm/s. The corresponding result from numerical simulations is 3.6 mm/s.

At $v = v_{c_1}$, oscillations exhibit a single loop. At the lower values of v_{c_n}, integer multiples of the fundamental period of oscillations appear, as described by the increased number of loops in the phase planes and corresponding subharmonics in the spectra. For example, the period is doubled at v_{c_2}, tripled at v_{c_3}, and so on. Although only a limited number of such platform speeds are revealed in these results (Tables II and III), there may be an infinite number of them for each combination of surface roughness and tangential natural frequency.

In accordance with Eq. (28), critical speeds at which stick–slip occurs (and $\overline{\mu}_D$ reaches its peak values) shift to higher values with increasing tangential natural frequency of the dynamic system. However, their values decrease with an increase in surface roughness, i.e., increased expected value of slopes. Further, values of the critical speeds are not affected by the changes in the tangential damping of the dynamic system or by the adhesive forces at the interface.

External damping inhibits stick–slip response of the system. Suppression of stick–slip by a small amount of damping is more acute for systems with higher tangential natural frequency or lower mean values of contact slopes.

Development of stick–slip also requires presence of both the deformation and adhesion forces. Without adhesion, direction of the deformation component of time-dependent friction force fluctuates, crossing zero-velocity state without reaching the condition of stick. Adhesion assures existence of the sticking condition when the relative velocity reaches zero, for at that moment the adhesive force changes direction and maintains, even for a short moment, zero relative velocity. In the absence of adhesion, deformation force can switch direction without sticking. On the other hand, adhesion, without the deformation forces, causes only static displacement of the block, without oscillations.

C. Stick–Slip bands

Numerical results indicate that around each critical speed there is a band of platform speeds within which stick–slip motion occurs, provided the external damping is sufficiently small. Outside of these *stick–slip bands*, stick–slip motion may not develop, regardless of the amount of damping. The widths of the bands differ at each critical speed: they appear to be the narrowest at the highest critical speeds

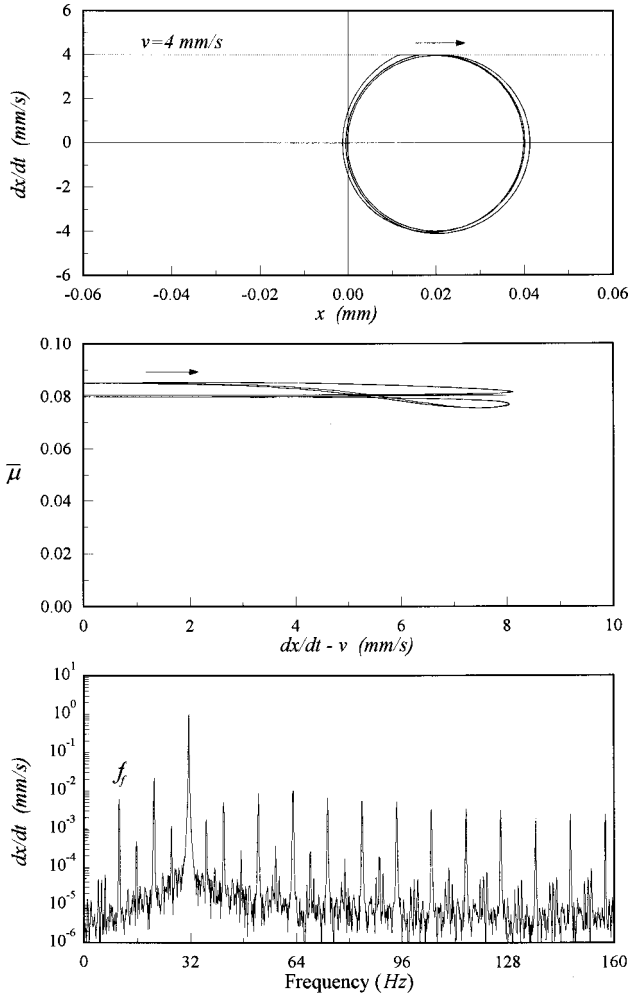

FIG. 14. Response of the dynamic system with a tangential natural frequency of $f_0 = 31.6$ Hz displays stick–slip oscillations at a platform speed of 4 mm/s.

and become wider at lower speeds. At very low speeds, however, the bandwidths increase again. The apparent dependence of stick–slip bandwidth on the critical speed could not be generalized with the limited number of results obtained here.

In some cases, stick–slip band is so narrow that external damping must be nearly zero for stick–slip motion to exist. In such cases, stick–slip response may not be sustained, but develops intermittently. Such stick–slip conditions are considered here to be ''unstable,'' unlike the ''stable'' conditions that lead to sustained stick–slip motion.

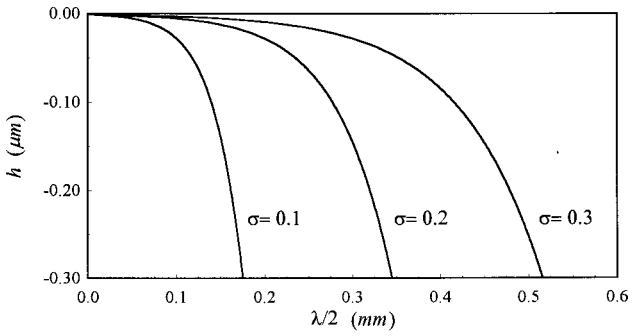

FIG. 15. Variation of λ/2 as a function of relative approach for different values of standard deviation σ of contact slope distribution.

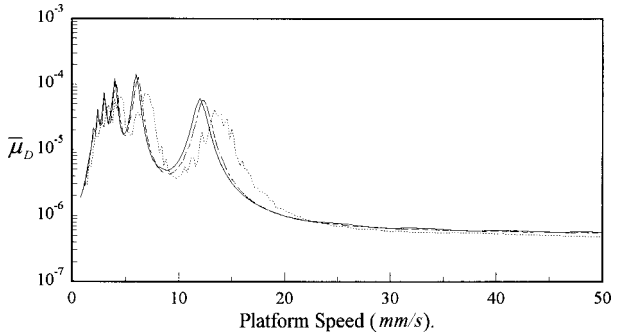

FIG. 16. Variation of $\overline{\mu}_D$ with platform speed for an unstable system characterized by a continuously changing variance. The dynamic system has a tangential natural frequency of $f_0 = 31.6$ Hz. Solid line indicates results for contact slope distribution with a constant variation. The dashed and dotted lines show unstable stick–slip resulting from changes of variance within 5% and 25% of the nominal.

When stick–slip oscillations are stable, system response settles to one stable limit cycle following an initial transient state. In cases of unstable stick–slip conditions, system response moves between two concentric, unstable limit cycles. Stick–slip takes place when the phase-plane trajectory follows the outer limit cycle. Following several stick–slip periods on or near the outer limit cycle, system response slowly diminishes and the phase-plane trajectory moves to the inner limit cycle. The inner limit cycle is also unstable; system response starts to grow until it reaches the outer limit cycle and stick–slip begins again. Figure 12 illustrates an unstable stick–slip response where the poor resolution of the velocity spectra results from the very short duration of stick and slip at the outer limit cycle.

D. Effects of surface roughness

Numerical results show that some critical speeds do not coincide with exact integer divisors of the fundamental critical speed. This is an indication of the existence of a modest nonperiodicity in the motion. Such behavior becomes more obvious when the distribution of contact slopes is nonstationary.

In cases of nonstationary contact slope distribution, the average contact slope, $\langle Z' \rangle$, becomes nonperiodic. The effect of nonstationarity of contact slopes on friction and system response is investigated numerically by randomly changing the variance of contact slope during oscillations. Results for which variance of slope distribution changes within 5% and 25% of its nominal value are plotted in Fig. 16. When compared with the corresponding stationary case, 5% variance in $\langle Z' \rangle$ has only minor effects on the value of critical speeds, while more significant changes develop with a 25% variance; Fig. 16. Further, the stick-slip oscillations could still be observed when the change in random variance is confined to within 5%, while it was not possible to obtain any stick–slip with variance changes within 25% as the surface roughness becomes less stationary.

In the numerical cases treated in this paper, resonance-like peaks of $\overline{\mu}_D$ occur at and below the fundamental critical

velocity v_{c_1}. As illustrated in Fig. 6, beyond v_{c_1}, $\overline{\mu}_D$ does not depend on the tangential natural frequency of the system, and smoothly wanes with increasing platform speed. At such high speeds compared with v_{c_1}, separation between surfaces remains large and approaches its maximum value, for there is not sufficient time for the surfaces to completely approach each other and fully undergo the effects of surface roughness. Near the speeds where $\overline{\mu}_D$ vanishes, computations become unstable and the current model is no longer applicable. For example, in the case of $\sigma = 0.1$ in Fig. 7, the $\overline{\mu}_D$ vanishes around a platform speed of 285 mm/s, whereas, for $\sigma = 0.2$ and $\sigma = 0.3$, it vanishes at platform speeds 550 and 800 mm/s, respectively. The numerical instability beyond the limiting platform speeds is ascribed to contact loss at the interface.

V. SUMMARY AND CONCLUDING REMARKS

The model developed in this paper demonstrates that friction depends on both the interface properties of the surfaces and on the dynamic response of the system that embodies them. The model relates macro-scale friction force to micro-scale forces developed at the true contact areas between surfaces. Expressing the contact forces in terms of contact areas and summing them statistically establishes this relationship. This study also considers contacts at the slopes of asperities, producing normal forces resulting from tangential relative motion of surfaces that sets this model apart from the previous ones.

A simplified mathematical model of a dynamic system is used to demonstrate the frequently observed stick–slip behavior in dynamic systems. An important result found in the cases considered above points to the significance of the deformation component of the friction force even if it is dominated by the adhesive component. Although the average friction force is essentially constant with respect to the mean sliding velocity of a friction platform, the deformation component of friction force shows a resonancelike behavior, reaching peak values at certain critical speeds. Numerical results show that stick–slip occurs only at and within a narrow band of each critical speed of the platform, defined here as *the stick–slip bands* of the critical speeds. The widths of the stick–slip bands are observed to depend on the critical speed. It is notable that at the *fundamental critical speed* of the platform, the system response is periodic, and at the progressively lower values of critical speeds, system response exhibits multiplicity of periods, such as doubling at the next lower speed, tripling at the following, and so on.

Stick–slip vibrations occur only in the presence of both deformation and adhesion components of frictional forces. An increase in surface roughness increases the strength of stick–slip motion, making it possible to develop even in the presence of large damping. On the other hand, changes in the magnitude of adhesive forces do not affect development of stick–slip as long as adhesive force magnitudes are above a certain threshold. This threshold roughly corresponds to the magnitude of the deformation forces.

The existence of stick–slip is also related to the stationarity of the contact slope distribution. For a stationary contact slope distribution, the average contact slope is a periodic function of separation, and stick–slip can be generated at certain sliding velocities. On the other hand, nonstationary contact slope distributions lead to nonperiodic average contact slopes, which do not produce stick–slip vibrations unless the nonstationarity of the distribution, or the deviation from the nominal, is small. The implication here is that surfaces that have a stationary distribution of asperity slopes (roughness), as may be the case for machined surfaces, have a higher propensity to exhibit stick–slip than surfaces that have nonstationary roughnesses.

The friction force expression developed in this paper may be expanded to include other processes that contribute to friction to the extent that they can also be expressed in terms of contact areas.

ACKNOWLEDGMENTS

One of the authors (A.A.) would like to thank the AlliedSignal Foundation and the Chrysler Challenge Fund for their support for this project.

[1] J. A. Greenwood and J. B. P. Williamson, ''Contact of nominally flat surfaces,'' Proc. R. Soc. London, Ser. A **295**, 300–319 (1966).
[2] J. A. Ogilvy, ''Numerical simulation of friction between contacting rough surfaces,'' J. Phys. D **24**, 2098–2109 (1991).
[3] A. Majumdar and B. Bhushan, ''Fractal model of elastic-plastic contact between surfaces,'' ASME J. Tribol. **113**, 1–11 (1991).
[4] W. R. Chang and I. Etson, ''An elastic-plastic model for the contact of rough surfaces,'' ASME J. Tribol. **109**, 257–263 (1987).
[5] A. Soom and C. Kim, ''Interactions between dynamic normal and frictional forces during unlubricated sliding,'' J. Lubr. Technol. **105**, 221–229 (1983).
[6] E. Rabinowicz, *Friction and Wear of Materials* (Wiley, New York, 1965).
[7] M. T. Bengisu and A. Akay, ''Relation of dry-friction to surface roughness,'' ASME J. Tribol. **119**, 18–25 (1997).
[8] L. V. Kragelsky, M. N. Dobychin, and V. S. Kombalov, *Friction and Wear Calculation Methods* (Pergamon, New York, 1982).
[9] J. A. Greenwood and J. H. Tripp, ''The contact of two nominally flat rough surfaces,'' Proc. Inst. Mech. Eng. **185**, 625–633 (1970).
[10] L. Cremer and M. Heckl, *Structure-Borne Sound* (Springer-Verlag, New York, 1987).
[11] M. T. Bengisu and A. Akay, ''Stability of friction-induced vibrations in multi-degree-of-freedom systems,'' J. Sound Vib. **171**, 557–570 (1994).
[12] J. A. C. Martins, J. T. Oden, and F. M. F. Simoes, ''A study of static and kinetic friction,'' Int. J. Eng. Sci. **28**, 29–92 (1990).