An Empirical Model for Dynamic Friction in Microfabricated Linear Microball Bearings

Xiaobo Tan, Alireza Modafe, and Reza Ghodssi

Abstract—The frictional dynamics of microball bearings needs to be understood for effective design and control of micromachines using such bearings. In the prior work by the authors and coworkers, a vision-based, non-intrusive measurement method was developed for characterizing the microtribological behavior of linear microball bearings. In this paper a novel scheme is proposed for computing the frictional force and the relative velocity between bearing elements using noisy position information extracted from the video. It enables one to reveal essential frictional characteristics despite the low capture rate (30 frames/second) of the camera. Experimental and numerical results show that the dynamic friction behavior is captured by a memoryless viscous model.

I. INTRODUCTION

The field of microelectromechanical systems (MEMS) has witnessed rapid advances over the past two decades, and the trend of miniaturization could revolutionize virtually all aspects of our life [1]. This presents unique opportunities and challenges in modeling and control designs, since dominant players in system dynamics differ in micro/nano and macro domains [2]. Many physical phenomena are important yet not well understood at micro and nano scales, a well-known example of which is friction.

In order to provide low-friction and robust support in micromachines, Ghodssi et al proposed a linear microball bearing structure [3] consisting of microfabricated V-grooves and commercially available stainless-steel microballs. Understanding of tribological behaviors of microball bearings is essential as it is directly related to the proper design and effective control of micromachines based on such bearings. Experiments were conducted to measure the static coefficient of friction (COF) of a linear microball bearing [3]. Its dynamic COF was investigated with a vision-based experimental setup [4], where a Coulomb friction model was used. Recently the authors and their coworkers upgraded the experimental system to capture motions of all bearing elements (the slider, the stator, and the microballs) with infrared imaging. This led to characterization of some important tribological behaviors such as rolling vs slipping and effects of surface roughness [5]. However, a faithful mathematical model was not available for the microball bearings since the Coulomb friction was a crude approximation, as evidenced by the scattering of data points on the friction vs. velocity plot [4].

Classical frictional models typically express the frictional force as a static function of the relative velocity between the contact surfaces, which may include the static friction, Coulomb friction, Stribeck friction, and viscous friction [6], [7], [8]. The Dahl model describes the friction in the stiction regime (also called presliding regime or micro-slip regime) in terms of the micro-displacement [9]. The relationship between the friction and the displacement can be hysteretic [10]. A dynamic friction model (called LuGre model) was presented by Canudas de Wit et al with an internal state representing the average deflection of contacting asperities [11], and was further extended by others [10], [12]. This model demonstrated, among other properties, the hysteresis between the friction and the sliding velocity for unidirectional sliding. Such hysteresis was also reported for unidirectional, unsteady sliding velocities and modeled through a time lag by Hess et al [13]. Similar hysteresis behavior was studied for a forced oscillator with a compliant contact [14]. The experimental results reported in [10], [11], [13], [14] were all based on macroscopic systems.

This paper aims to model the dynamic friction in linear microball bearings. The term “dynamic friction” (as opposed to the static friction) refers to the friction when a relative motion between the slider and the stator is present. In the experiment the stator is fixed to a forced oscillator, and the motion of the slider is driven only by the friction. Hence the friction is directly linked to the acceleration of the slider, which is determined based on successive video images. In general the bearing dynamics is very complex due to fabrication-related surface irregularities, ball-to-ball or ball-to-wall collisions, and oxide growth on contact surfaces [5], and the slider may occasionally demonstrate stick-slip, slow drift, or sudden impacts. However, the emphasis of this paper is on the frictional behavior when the slider is (relatively) steadily sliding. The reason is that steady sliding is supposed to be the normal operating condition of these bearings once the fabrication processes are refined. Due to the constraint of camera speed, the (relative) velocity regime is about [-0.02, 0.02] m/s. Note that this limitation does not undermine the importance of the study here since interesting friction-related dynamics typically takes place at low-velocities and during velocity reversals [15]. In particular, the friction behavior in this regime will prove very important for some applications (e.g., sub-micron positioning) envisioned for micromotors based on the bearing.

X. Tan is with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA. xbtan@msu.edu
A. Modafe and R. Ghodssi are with the Institute for Systems Research and the Department of Electrical & Computer Engineering, University of Maryland, College Park, MD 20742, USA. {modafe,ghodssi}@glue.umd.edu
In modeling the first question would be whether the friction-velocity relationship shows hysteresis (memory) or is static. And then one would like to know what exactly is the relationship. In the vision-based measurement system both the friction and the (relative) velocity are derived from the position information available through the camera. The camera captures 30 frames/second, and this (relatively) slow capture rate presents two serious challenges. First the images are typically blurred leading to errors in position identification. The second problem is that, without proper care, the derived velocities and acceleration will deviate significantly from the true values.

In this paper a three-stage data processing algorithm is presented to best extract the information of interest: a) approximation of derivatives (acceleration, velocities) using the center-difference method to minimize the phase error; b) amplitude-based filtering to reduce measurement noises; and c) frequency-weighted compensation to reconstruct the derivatives. Both numerical and experimental results show that the algorithm can effectively recover the derivative quantities based on the position snapshots. A much cleaner friction-velocity plot obtained through the algorithm shows that, (at least for the velocity range considered in the paper) the friction is a static function of the velocity. At low relative velocities the friction is approximately linear in the relative velocity thus demonstrating a viscous behavior; as the relative velocity increases, the friction approaches saturation. A Langevin function is used to approximate the relationship. The predictions based on the model agree well with the experimental results.

The remainder of the paper is organized as follows. The microball bearing and the experimental setup are described in Section II. The three-stage algorithm is presented in Section III. Identification of the friction model is discussed in Section IV. Finally conclusions are provided in Section V.

II. MICROBALL BEARINGS AND EXPERIMENTAL SETUP

A schematic of a linear microball bearing is shown in Fig. 1. The bearing consists of two silicon plates (slider and stator) and stainless-steel microballs of diameter 285 µm. Two parallel V-grooves, which house the balls, are etched on the plates using potassium hydroxide (KOH). Fig. 2 shows the picture of the experimental setup. The stator of the bearing is mounted on an oscillating platform driven by a servomotor through a “crank-slider” mechanism (for more details, see [4]). The platform (and hence the stator) moves only in the direction of the underneath guiding rails, so does the slider of the bearing. The position \( x(t) \) of the stator is approximately

\[
X_{\text{stat}}(t) = x_{\text{stat}}(\omega t),
\]

where \( \omega \) is the angular velocity of the servomotor, and the oscillating amplitude \( X_{\text{stat}} \) can be adjusted. Distinct marks are placed and also etched on the stator and slider of the bearing, respectively. A CCD camera (Sony DCR-TRV22) with a 24 × magnification lens captures motions of these marks, and position information of the slider and the stator can be extracted through image processing (National Instruments, IMAQ Vision Builder™). When switched to the night-shot mode (infrared), the camera can track the slider, the stator, and the microballs simultaneously [5].

III. THREE-STAGE DATA PROCESSING ALGORITHM

The equation of motion for the slider is:

\[
M_{\text{slid}} \ddot{x}_{\text{slid}}(t) = F_{\text{fric}}(t),
\]

where \( M_{\text{slid}} \) denotes the mass of the slider, \( a_{\text{slid}} \triangleq \ddot{x}_{\text{slid}}(t) \) denotes its acceleration, and \( F_{\text{fric}} \) is the friction between the stator and the slider. It is clear that the friction is equal to \( a_{\text{slid}} \) normalized by \( M_{\text{slid}} \).

The slider experiences periodic motion during steady sliding. To determine whether there is hysteresis or other memory effect in the friction vs. the relative velocity relationship, one can check the phase difference between \( a_{\text{slid}} \) and \( v_{\text{rel}} \), where \( v_{\text{rel}} \) denotes the relative velocity:

\[
v_{\text{rel}}(t) = v_{\text{slid}}(t) - v_{\text{stat}}(t) = \dot{x}_{\text{slid}}(t) - \dot{x}_{\text{stat}}(t)\]

If the phase difference is a multiple of \( \pi \), \( a_{\text{slid}} \) would be a static (memoryless) function of \( v_{\text{rel}} \), and it would be a dynamic function of \( v_{\text{rel}} \) otherwise. Note that due to the nonlinearity of friction, each of \( x_{\text{slid}}, v_{\text{slid}}, v_{\text{rel}}, \) and \( a_{\text{slid}} \) will typically contain more than one frequency components.
Therefore, one needs to look at their fundamental frequency components when calculating the phase difference.

A challenging problem here is that all derivative quantities, $a_{\text{slid}}(t), v_{\text{slid}}(t), v_{\text{stat}}(t)$, and $v_{\text{rel}}(t)$, can only be obtained through finite differencing of noisy position snapshots, $x_{\text{slid}}[n]$ and $x_{\text{stat}}[n]$ (note in practice, $x_{\text{stat}}[n]$ is also extracted through image processing instead of being calculated from (1)). Without proper care, such differencing operations may introduce spurious phase shift in addition to the usual approximation error. It is shown next that the widely used forward-difference method is inferior to the center-difference method in terms of phase-preservation.

A. Comparison of Finite-Difference Methods

The analysis is conducted for a generic sinusoidal function $x(t) = X \sin(\omega t + \phi)$. For signals with more than one frequency component, the analysis holds for each component. Calculate

$$v(t) = \dot{x}(t) = \omega X \cos(\omega t + \phi),$$

$$a(t) = \ddot{x}(t) = -\omega^2 X \sin(\omega t + \phi).$$

(3)\hspace{1cm}(4)

Sample $x(t)$ with time step $h$, and denote $x[n] \triangleq x(nh)$. One is then interested in approximating $v[n]$ or $a[n]$ based on the sequence $\{x[n]\}$. Note that to connect this to the microball bearing problem, $x(t)$ plays the role of $x_{\text{slid}}(t)$ in calculation of the slider acceleration, and it plays the role of $x_{\text{stat}}(t) - x_{\text{slid}}(t)$ in calculation of the relative velocity.

Two methods, forward-difference and center-difference, are compared for approximation of both the velocity (3) and the acceleration (4). Under the forward-difference method, the estimates are

$$v_f[n] = \frac{x[n] - x[n-1]}{h},$$

$$a_f[n] = \frac{v_f[n] - v_f[n-1]}{h} = \frac{x[n] + x[n-2] - 2x[n-1]}{h^2}.$$  

(5)\hspace{1cm}(6)

Under the center-difference method, the estimates are

$$v_c[n] = \frac{x[n+1] - x[n-1]}{2h},$$

$$a_c[n] = \frac{x[n+1] + x[n-1] - 2x[n]}{h^2}.$$  

(7)

1) Forward-difference method: Let $t = nh$ and $\theta = \omega t + \phi$. It then follows that

$$v_f[n] = \frac{1}{h}(x[n] - x[n-1]) = \frac{1}{h}[X \sin(\theta) - X \sin(\theta - \omega h)]$$

$$= \frac{X}{h} \left[ \cos(\theta) \sin(\omega h) + \sin(\theta)(1 - \cos(\omega h)) \right]$$

$$= \frac{\sqrt{2(1 - \cos(\omega h))X \cos(\theta)}}{h} \cos(\theta - \phi_1) = m_1 \omega X \cos(\theta - \phi_1) = m_1 \omega X \cos(\omega t + \phi - \phi_1),$$

where $m_1 = \frac{\sin(\omega h/2)}{\sin(\omega h/2)}$, and $\phi_1 = \tan^{-1} \left( \frac{1 - \cos(\omega h)}{\sin(\omega h)} \right)$. Comparing with (3), one can see that not only the magnitude of $v_f[n]$ is scaled from that of $v(nh)$ by $m_1$, but also its phase differs from that of $v(nh)$ by $\phi_1$. Similarly, one can compute $a_f[n]$:

$$a_f[n] = \frac{x[n] + x[n-2] - 2x[n-1]}{h^2} = \frac{X}{h^2} \left[ \sin(\theta) + \sin(\theta - 2\omega h) - 2\sin(\theta - \omega h) \right].$$

(8)

Write $\sin(\theta - 2\omega h) = \sin(\theta) \cos(2\omega h) - \cos(\theta) \sin(2\omega h)$, write $\sin(\theta - \omega h)$ in a similar fashion, and plug them into (8). After some manipulations, one gets

$$a_f[n] = \frac{x[n] + x[n-2] - 2x[n-1]}{h^2} = \frac{X}{h^2} \left[ \sin(\theta + \omega h) - \sin(\theta - \omega h) \right].$$

It can be shown that

$$v_c[n] = \frac{\sin(\omega h)}{\omega h} \omega X \sin(\omega t + \phi) = \frac{\sin(\omega h)}{\omega h} v(nh),$$

(9)\hspace{1cm}(10)

and $v_c[n]$ is in phase with $v(nh)$. Furthermore, one calculates

$$a_c[n] = \frac{x[n+1] + x[n-1] - 2x[n]}{h^2} = \frac{X}{h^2} \left[ \sin(\theta + \omega h) + \sin(\theta - \omega h) - 2\sin(\theta) \right]$$

$$= \frac{-2(1 - \cos(\omega h))X \sin(\omega t + \phi)}{h^2}$$

$$= \left( \frac{\sin(\omega h/2)}{\omega h/2} \right)^2 \omega^2 X \sin(\omega t + \phi)$$

$$= \left( \frac{\sin(\omega h/2)}{\omega h/2} \right)^2 \omega^2 \sin(\omega t + \phi)$$

$$= \frac{m_2 \omega X \cos(\theta - \phi_2)}{h^2}$$

where $m_2 = (\frac{\sin(\omega h/2)}{\omega h/2})^2$, and $\phi_2 = \omega h$. Therefore, the magnitude of $a_c[n]$ is scaled from that of $a(nh)$ (compare (4)) by $m_2$, and its phase differs from that of $a(nh)$ by $\phi_2$.

2) Center-difference method: For the center-difference method,

$$v_c[n] = \frac{x[n+1] - x[n-1]}{2h} = \frac{X}{2h} \left[ \sin(\theta + \omega h) - \sin(\theta - \omega h) \right].$$

(11)

and again, $a_c[n]$ is in phase with $a(nh)$.

From the above discussions, although both difference methods introduce magnitude scaling, a signal generated through the center-difference method carries the correct phase while the one generated through the forward-difference method does not.

B. The Data Processing Algorithm

The analysis in Subsection III-A is based on the assumption that $x[n]$ is sampled from a noiseless sinusoidal signal $x(t)$ of single frequency. This is typically not the case with the position data extracted from video images. Hence the proposed data processing algorithm consists of three stages (Fig. 3): a) calculating a crude estimate of velocity or acceleration (depending on the context) using
the center-difference method; b) performing Fast Fourier Transform (FFT) and then filtering based on the amplitude of individual frequency components; and c) reconstructing the velocity or acceleration by amplitude compensation for remaining frequency components.

![The three-stage data processing algorithm.](image)

1) Stage 1. Center-difference: The estimate \( v_c[n] \) (or \( a_c[n] \)) is obtained through the center-difference method (5) (or (6)). Note that each frequency component of \( v_c[n] \) (or \( a_c[n] \)) carries the same phase as the original \( v(nh) \) (or \( a(nh) \)) but the amplitude undergoes certain scaling.

2) Stage 2. FFT and filtering: FFT is performed for two purposes: filtering and preparing for amplitude compensation (Stage 3). It is tempting to implement low-pass filtering for noise reduction. However, low-pass filtering is not appropriate here since the signals may carry high-order harmonics due to the nonlinear frictional dynamics. Filtering out all high-frequency components, therefore, can distort the true dynamics. Furthermore, plain low-pass filtering introduces additional phase shift for signals.

In this paper amplitude-based filtering is used instead. Let \( M_{\text{max}} \) be the maximum amplitude among all frequency components. Then a frequency component \( k \) will be eliminated if its amplitude \( M_k \leq \epsilon_0 M_{\text{max}} \), where \( \epsilon_0 \) is some small constant.

3) Stage 3. Amplitude compensation: Let \( k \) be the index of remaining frequency components after filtering. Then (see Fig. 3)

\[
\hat{v}_c[n] = \sum_k V_k \sin(\omega_k nh + \alpha_k),
\]

\[
\hat{a}_c[n] = \sum_k A_k \sin(\omega_k nh + \beta_k).
\]

Based on (10) and (11), amplitude scaling is performed to compensate for the frequency-dependent amplitude distortion introduced in the difference methods:

\[
\hat{v}_c[n] = \sum_k \frac{\omega_k h}{\sin(\omega_k h)} V_k \sin(\omega_k nh + \alpha_k),
\]

\[
\hat{a}_c[n] = \sum_k \left(\frac{\omega_k h/2}{\sin(\omega_k h/2)}\right)^2 A_k \sin(\omega_k nh + \beta_k).
\]

Simulation is conducted to compare the three-stage algorithms with the forward-difference algorithms. For illustration, two sinusoidal trajectories are specified first:

\[
x_1(t) = X_1 \cos(\omega t + \pi),
\]

\[
x_2(t) = X_2 \sin(\omega t),
\]

where \( x_1(t) \) mimics the relative position of the slider to the stator, and \( x_2(t) \) mimics the slider position (again, it is not implied here that true \( x_{\text{slid}} \) or \( x_{\text{stat}} \) is purely sinusoidal; refer to the footnote in Subsection III-A). One would like to calculate \( v_1[n] \) (analogy of relative velocity) and \( a_2[n] \) (analogy of acceleration) based on snapshots \( x_1[n] \) and \( x_2[n] \) with sampling time \( h \). Note that \( x_1(t) \) and \( x_2(t) \) are specified in such a way that there is a phase difference of \( \pi \) between \( v_1(t) \) and \( a_2(t) \) (hence no loop in the \( a_2(t) \) vs. \( v_1(t) \) plot).

In simulation, \( X_1 = 0.0011, X_2 = 3.1 \times 10^{-3}, \omega = 6\pi, h = 0.033 \), and the filtering level \( \epsilon_0 = 0.002 \) (although \( x_1(t) \) and \( x_2(t) \) are purely sinusoidal, one will get extra frequency components in FFT due to the finite number of data points).

Fig. 4 compares the \( a_2 \) vs. \( v_1 \) plots obtained through the three-stage algorithm and the forward difference algorithm, respectively. It is clear that the forward-difference algorithm leads to a spurious loop, introduced by the phase errors in approximating \( v_1 \) and \( a_2 \). The phase difference between \( a_2 \) and \( v_1 \) in Fig. 4 (a) is numerically calculated to be 2.83, which is consistent with the analytical value obtained through (7) and (9). On the other hand, the three-stage algorithm is able to show the phase difference of \( \pi \) between \( a_2 \) and \( v_1 \) correctly (Fig. 4).

![Comparison of the three-stage algorithm with the forward-difference algorithm.](image)

IV. IDENTIFICATION OF FRICTION MODEL

The three-stage algorithm is then used to process the actual position data collected from the microball bearing through imaging. Fig. 6 shows the trajectories of the stator, the slider, and their relative positions, respectively. Note that although the slider experiences slow, random drift due to fabrication-related surface irregularities, its periodic oscillation is dominant. The stator oscillates with frequency 2.95 Hz and amplitude 1.2 mm.
Fig. 6. Experimentally measured trajectories of the slider (top), the stator (center), and their relative position (bottom).

Fig. 7 shows the relationships between the slider acceleration $a_{slid}$ (namely, friction normalized by the slider mass) and the relative velocity $v_{slid} - v_{stat}$ when three different schemes are used to compute these quantities. Direct application of the forward-difference method results in a noise-corrupted loop (Fig. 7(a)). If low-pass filtering (cutoff frequency 6 Hz) is used followed by the forward-difference, the loop trend is more evident (Fig. 7(b)). However, there is no loop in the acceleration vs. relative velocity plot when the three-stage data processing algorithm ($\epsilon_0 = 0.08$) is applied (Fig. 7). The phase difference between the fundamental frequency components of $a_{slid}$ and $v_{slid} - v_{stat}$ is 3.572 in Fig. 7(a), 3.585 in Fig. 7(b), and 3.157 (very close to $\pi$) in Fig. 7 (c). Since analysis in Section III has shown that the three-stage algorithm conserves the phase of the original signal, one concludes that the friction is well approximated by a static but nonlinear function of the relative velocity, and the “hysteresis” loops in Fig. 7(a) and (b) are artifacts of data processing.

From Fig. 7(c), the friction is almost linear with respect to the relative velocity when the latter is low; and as the magnitude of the relative velocity increases, the friction approaches a saturation level. Hence a modified Langevin function $\mathcal{L}(\cdot)$ is proposed to model the friction

$$ F_{fric}(t) = M_{slid} \mathcal{L}(v_{rel}(t)), \quad (12) $$

where for any $z$,

$$ \mathcal{L}(z) \triangleq A_0 \left[ \frac{1}{\alpha z} - \frac{e^{\alpha z} + e^{-\alpha z}}{e^{\alpha z} - e^{-\alpha z}} \right]. \quad (13) $$

To identify the parameters $A_0$ and $\alpha$, one calculates

$$ \frac{d\mathcal{L}(z)}{dz} = \alpha A_0 \frac{4}{(e^{\alpha z} - e^{-\alpha z})^2} - \frac{1}{(\alpha z)^2}. \quad (14) $$

By evaluating the derivatives at $v_{rel} = 0$ and $v_{rel} = 0.015$ in Fig. 7(c) and solving (14), the parameters are determined to be $A_0 = 0.112$, $\alpha = 128.65$. Fig. 8 plots the friction vs. velocity relationship obtained through this model, and it matches the experimental data well (Fig. 8).

Simulation is further conducted to validate the above model. Combining (1), (12), and (13), the equation of motion (2) is integrated using the fourth order Runge-Kutta method with a time step of 0.001 second. The oscillation amplitude of the stator $X_{stat} = 1.2$mm and $\omega = 2\pi f$ with $f = 2.95$Hz, both estimated from the experimental data. Fig. 9 shows the simulated trajectories of the slider, the stator, and their relative position. Note the absolute position values are of little relevance since the reference points are defined arbitrarily. The oscillation amplitude of the slider in Fig. 9 is 0.210 mm, while the amplitude of the fundamental frequency component of $x_{slid}[n]$ in the experiment (Fig. 6) was 0.206 mm. Furthermore, the phase difference between $x_{slid}$ and $x_{stat}$ in Fig. 9 is 1.384 radians, comparing to 1.416 radians measured in the experiment (Fig. 6). These agreements strongly support that the proposed model is capable of capturing the friction behavior.

V. CONCLUSIONS

This paper was focused on the modeling of dynamic friction in MEMS-based linear microball bearings. The only available information was the noisy position data of the slider and the stator extracted from videos captured at a low rate. To recover velocity and acceleration with high fidelity from these noisy data, a three-stage data processing algorithm was proposed. The signals (derivatives of position) generated through this algorithm preserve the phase as well as the amplitude of the true signals. An enabling assumption of the algorithm is that signals are periodic, which holds true when the slider oscillates and slides steadily.

The friction - (relative) velocity plot obtained through the data-processing algorithm showed that hysteresis or other memory effects, if existing at all, were insignificant at least for the velocity range examined in the paper. The
friction demonstrates a viscous feature at low velocities, and then approaches saturation. Such behavior was captured well through the Langevin function. Simulation predictions based on this model achieved good agreement with experimental measurements. Note that the finding is interesting in that classical friction models are typically discontinuous at zero velocity [15], [8].

Future work involves understanding of the observed friction behavior. It is of particular interest to study the viscoelastic dynamics at the contact surfaces during the rolling motion of microballs.

ACKNOWLEDGMENT

This work was supported in part by the NSF grant ECS-0224361. The first author would like to acknowledge useful discussions with Professors P. S. Krishnaprasad, Ben Shapiro, and Brian Feeny.

REFERENCES


