Proc. of SPIE Vol. 1998, Optomechanical Design, ed. D Vukobratovich, P R Yoder, V L Genberg (Sep 1993) Copyright SPIE

The influence of elastic hysteresis on combined (rolling-sliding) friction

Peter Hoa Vo

Hughes Danbury Optical Systems, Inc.
100 Wooster Heights Road, Danbury, Connecticut 06810-7589

ABSTRACT

An investigation of the effects of elastic hysteresis on certain ball bearing/plate configurations used in optics mounting (when subjected to rolling-sliding friction forces) has been performed. The energy losses attributable to surface friction breakaway coefficients were measured. The elastic strain work in the ball and plate material has been proven to be the primary cause of the inherent, measurable energy loss of hysteresis.

This paper compares the hysteresis work loss of different ball bearing sizes and materials when interacting with tungsten carbide plate. In addition, the Dahl[3] friction force (i.e., the friction force required to move a ball within the linear region of the static friction curve) is demonstrated and measured. Ball materials of stainless steel, ceramic, and tungsten carbide were subjected to laboratory testing with results indicating that the forces of rolling and/or sliding friction are influenced by, and dependent upon the ball size and material used in the ball bearing support mount.

Also, a methodology for performing calculations of the rolling friction breakaway forces extending beyond the Dahl friction regime (at the transition point of static to dynamic) is demonstrated.

1. INTRODUCTION

The precision metrology mount used on the Advanced X-Ray Astrophysics Facility (AXAF) program to support the large X-ray optics employs a combination of rolling-sliding friction. An instrument ball bearing interface is used to support the X-ray optics, as a limited friction influence is necessary to minimize distortion on the optic due to gravitational force. (A typical optic support interface concept is illustrated in Fig. 1.)

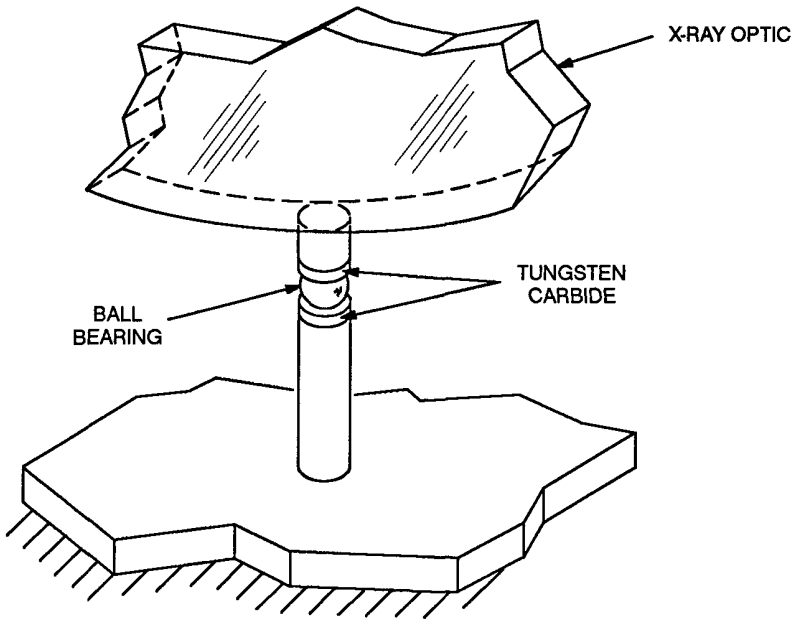

Fig. 1. A typical x-ray optic support interface.

There is no easy method to calculate the combined friction force as a function of different materials and ball diameters. As a result, an experiment to collect empirical data for combined friction forces was conducted.

During the process of obtaining data indicative of the combined friction forces, the energy losses due to elastic hysteresis were measured, as well as the breakaway coefficients of friction of a ball moving between two plates.

The energy loss due to the elastic hysteresis of a ball moving between plates was measured in the "spring region", where the ball travel distance was small enough so that actual rolling had not yet occurred. In addition, the breakaway coefficients of rolling friction were measured at the point when the ball first started rolling. In other words, the ball's breakaway friction forces, from a static to a dynamic condition, were measured.

Rolling of the ball between two plates is a cyclic process of changing stresses at the ball's contact zones. As the ball moves forward, the front end of the ball is under compressive stress while the rear end is under tensile stress. Under this loading, the ball and the plates temporarily deform at the contact zones. The deformation is elastic on a macro-phenomena scale, provided that the material elastic limit has not been exceeded.

The energy loss in the moving of a ball between two plates is caused by the energy required to produce the elastic strain in the material. This energy is then lost by conversion into heat and dissipated to the surroundings. Since the energy loss is an extremely small quantity, the precise measurement of the changes in temperature is very difficult.

Different ball sizes (ϕ0.50", ϕ0.75", ϕ1.00") and materials such as: tungsten carbide (Ra 91), 440C stainless steel (Rc 58), ceramic (Rockwell 45N81) and 52100 hardened steel (Rc 60-65), rolling against tungsten carbide plates (Ra 91) were all evaluated.

The results show that the magnitude of the work required for elastic hysteresis is indeed a function of different materials and different ball sizes. Generally, for the same ball size and under the same normal load within a given travel distance, more energy is required to roll a low-elastic modulus ball than a high-elastic modulus ball. The reason for this is that the low modulus ball requires more "work of deformation" (or elastic strain) in the material than the high modulus ball.

2. TEST APPARATUS

The test setup is shown in Fig. 2. The test fixture consists of two equilateral triangular plates. Both top and bottom plates have three tungsten carbide pads bonded to them. The three tungsten carbide pads are ground and lapped with diamond abrasive in random directions to the final configuration of one-microinch surface finish, and a fraction of an arcsecond coplanarity to each other.

The bottom plate has three built-in fine adjustment screws, one at each of the three corners of the plate. These adjustments provide the tilt alignments of the test fixture to gravity.

Initially the bottom plate was placed on a flat granite surface with the three polished-tungsten carbide pads facing up, and leveled to gravity to within ±3 arcseconds. Three balls were then placed between the top and bottom plates at the polished-tungsten carbide locations.

Static (normal) loads were then applied to the balls by placing dead weight in the center of the top plate.

The lateral movement of the top plate (with respect to ground) was monitored by a laser gage mounted to ground, with a retroreflector mounted on the edge of the top plate and facing the laser source. As the top plate moves, the retroreflector moves with it and this motion is detected by the laser receiver.

Adjustment in elevation of any of the three adjustment screws on the bottom plate will cause the test fixture to tilt about the pivot line, which is defined by the location of the other two adjustment screws. Changes in tilt of the bottom plate were measured by a displacement gauge (Federal gauge).

As the bottom plate tilts, the top plate is subject to a component force due to gravity, which causes the balls to slide-roll and the top plate to move. The distance the top plate travels, with respect to the bottom plate, is measured by the laser gauge.

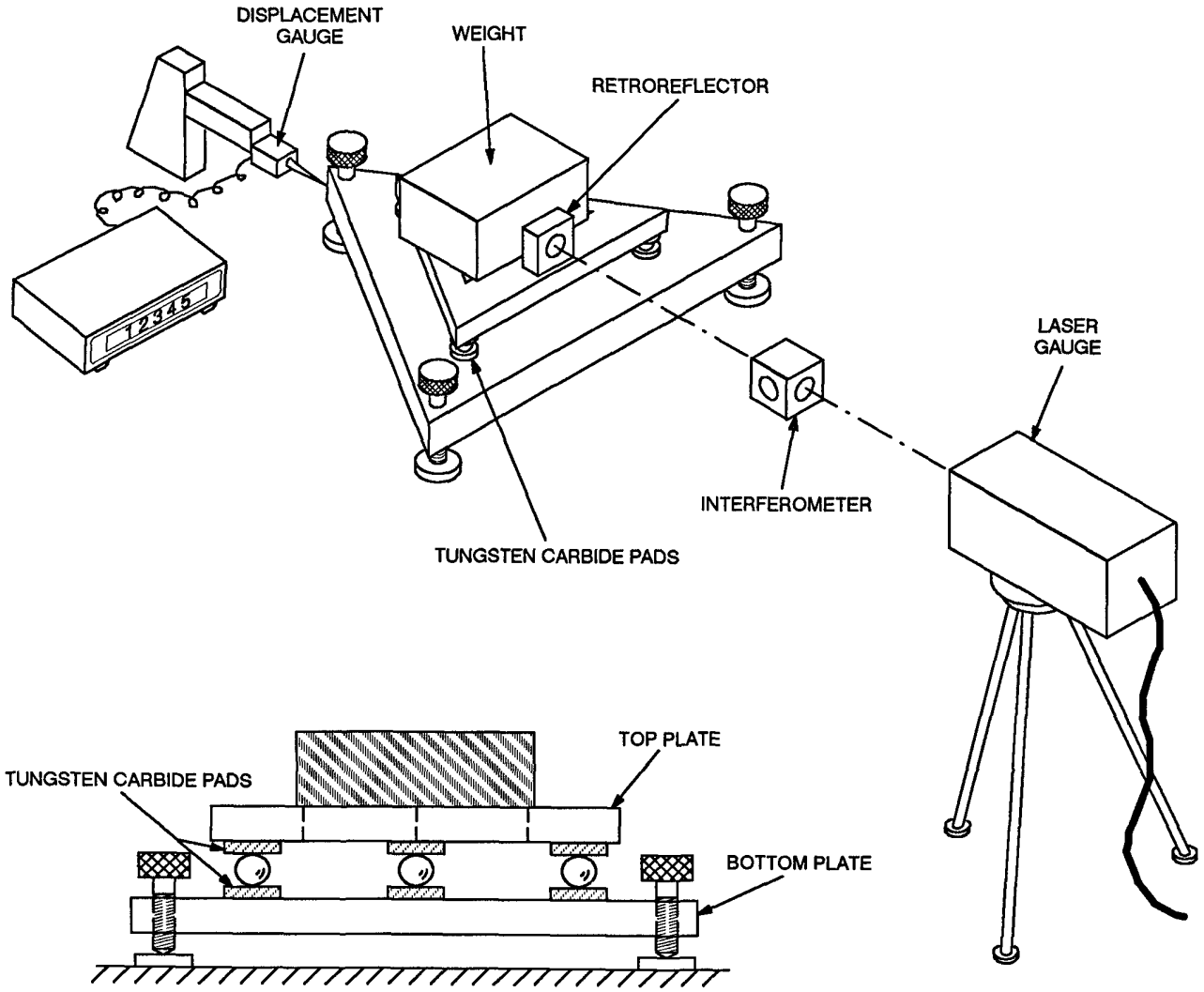

Fig. 2. Test setup for elastic hysteresis on combined friction.

Knowing the normal load and the tilt angle of the tungsten carbide plate with respect to gravity, the moving friction force can be calculated as follows:

$$F = W \cdot \tan \phi$$

for small angle, $\tan \phi = \phi$

therefore, $F = W \cdot \phi$

where,

F: moving friction force in "spring region" or Dahl[3] friction force (pounds)

ϕ: tilt angle of the flat plates with respect to gravity (radians)

W: weight on top of the three balls (pounds).

3. RESULTS

3.1. Dahl friction

A typical hysteresis loop was constructed with one forward and one reverse path, as shown in Fig. 3. For the particular test setup as shown in Fig. 2, the energy that is inputted to the system in the forward path is created by tilting the bottom plate from zero to an angle theta, which causes the shift in center of gravity of the weight that sits on top of the balls, and forms the potential energy in the system that causes the balls to move. As long as the potential energy is smaller than the moving friction force between the balls and the plates, the balls will not roll.

The data of the energy losses due to elastic hysteresis of the balls was collected, analyzed, calculated, and plotted out for different ball sizes and different materials. For Figs. 4, 5, 6, 7, and 8, the ordinates are the Dahl frictions and the abscissas are the displacements. The displacement is the movement of the top plate with respect to the bottom plate. The actual average of three data points are displayed as the square notations in the curves. The curve fitting method is not used for these curves because the scattering of data was minimum.

Since the quantity of data is large and the gaps between the data points are small, the total energy loss in hysteresis and the total input energy can be calculated by using a discrete integration method. The total input energy is the areas that formed by the forward path projected down to the abscissa. The total energy *loss* in hysteresis is the area *inside* of the hysteresis loop.

Figures 4 and 5 show respectively the hysteresis curves of a ϕ0.75-inch ball made of 440C, and a ϕ0.75-inch ball made of 52100 steel rolled against tungsten carbide pads. For a given displacement, the Dahl friction of the 440C ball is about 25% higher than the 52100 ball. However, the percentage of energy loss compared to total input energy (reference Table 1), for the 440C stainless steel ball (19%) is less than the 52100 chrome steel ball (30%). This is true even though the deformations and contact stresses of the balls are the same to begin with.

Consider two identical ball bearings, each having the inner and outer races made of tungsten carbide. One bearing has 440C balls and the other bearing has 52100 balls. As stated previously, the Dahl friction of the 440C ball is about 25% higher than the 52100 ball. For a small motion under the "spring region" (static), the 440C ball bearing requires more input torque to move the bearing than does the 52100 ball bearing. However, the 440C ball bearing will require less initial input torque than the 52100 ball bearing to break the bearing from the static to dynamic state. This is because the breakaway coefficient of rolling friction of a 440C ball is less than that of a 52100 ball (refer to Table 2).

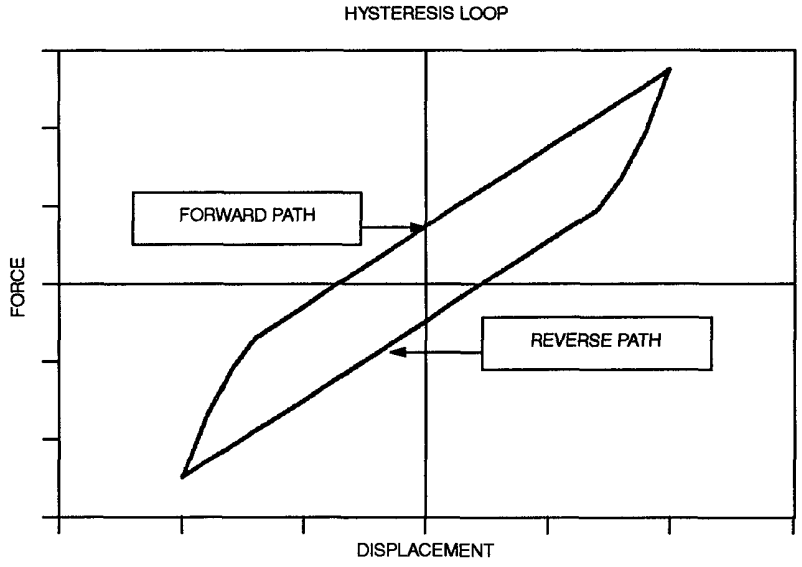

Fig. 3. Hysteresis loop.

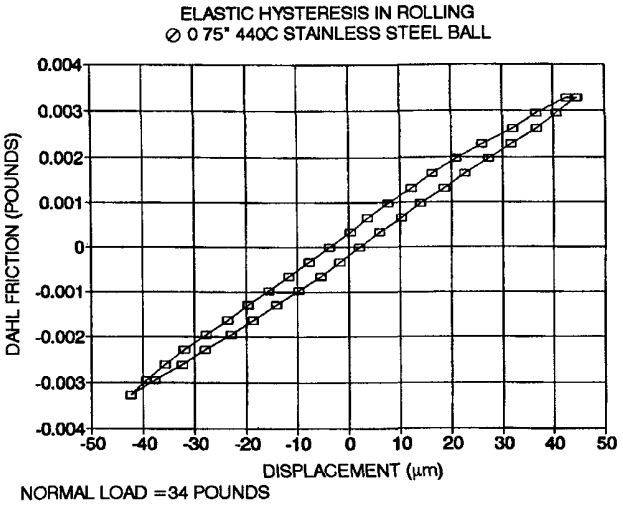

Fig. 4. Hysteresis of ϕ0.75-inch 440C stainless steel ball.

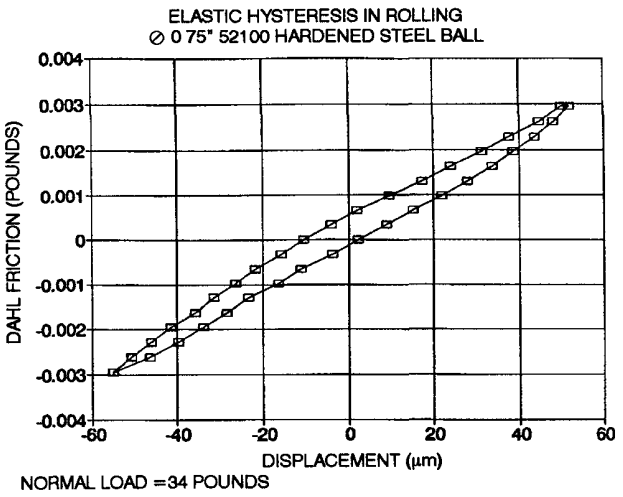

Fig. 5. Hysteresis of ϕ0.75-inch 52100 hardened steel ball.

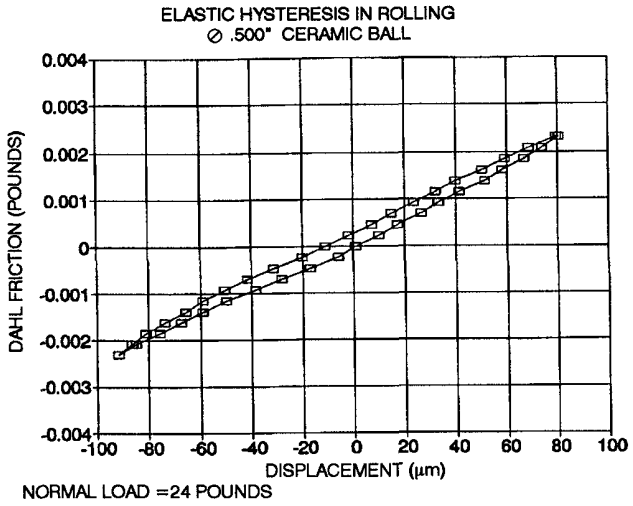

Fig. 6. Hysteresis of ϕ0.50-inch ceramic ball.

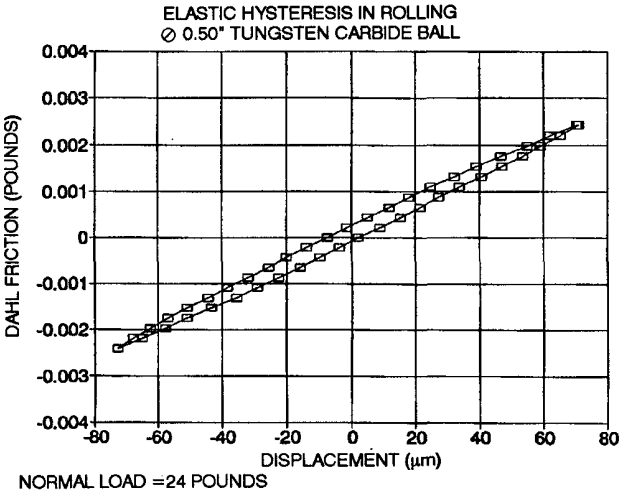

Fig. 7. Hysteresis of ϕ0.50-inch tungsten carbide ball.

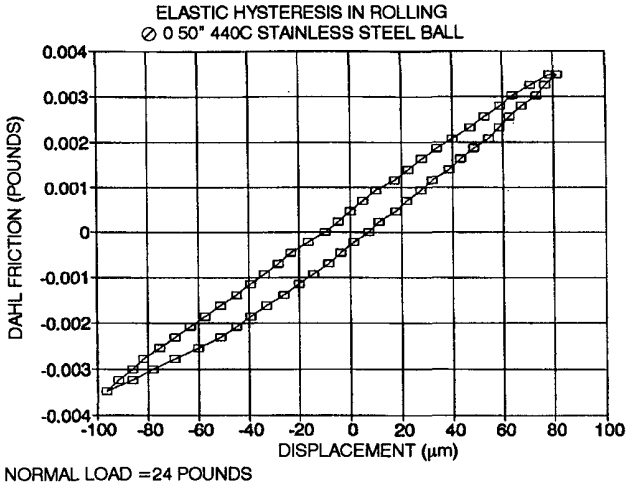

Fig. 8. Hysteresis of ϕ0.50-inch 440C stainless steel ball.

Table 1. Energy Loss Comparisons

Ball Size (in)	Ball Material	Normal Load (lbs)	Contact Stress (Ksi)	Ball/Plate Deformation (μ in)	Energy Loss Per Total Energy Input
ϕ0.75	440C	34	298	145	19%
ϕ0.75	52100	34	298	145	30%
ϕ0.50	440C	24	348	132	29%
ϕ0.50	Ceramic	24	453	100	20%
ϕ0.50	Tungsten Carbide	24	570	80	20%

Table 2. Breakaway Coefficient of Ball Bearings

Ball Size (in)	Normal Load (lbs)	Ball Material	Plate Material	Breakaway Coefficient of Rolling Friction
ϕ1.00	34	Tungsten Carbide	440C	0.00023
ϕ0.75	34	440C	Tungsten Carbide	0.00022
ϕ0.75	34	52100	Tungsten Carbide	0.00050
ϕ0.50	24	440C	Tungsten Carbide	0.00017
ϕ0.50	24	Ceramic	Tungsten Carbide	0.00011
ϕ0.50	24	Tungsten Carbide	Tungsten Carbide	0.00011

Heat that is generated in a typical bearing depends on the rolling friction force, or the coefficient of rolling friction and the preload. For the case of 52100 and 440C ball bearings, and assuming that both bearings are under the same preload and rotate at the same speed, the 52100 ball bearing is more likely to heat up faster and to a greater temperature than the 440C ball bearing. This is caused by the greater input torque to the 52100 bearing system. That is one reason why 440C stainless steel is considered a good material for ball bearings.

Figures 6, 7, and 8 show the hysteresis curves of a ϕ0.50-inch ceramic ball, a ϕ0.50-inch tungsten carbide ball, and ϕ0.50-inch 440C stainless steel ball, all against tungsten carbide pads.

The percentage of energy loss in hysteresis compared to total input energy of both ceramic and tungsten carbide balls are approximately 20%, while it is 29% for the 440C ball. This indicates that the 440C stainless steel ball requires more elastically strained work on the material than both ceramic and the tungsten carbide balls.

For applications which require small displacements and where rolling has not yet occurred, both ceramic and tungsten carbide balls are favored over the 440C ball. Since ceramic and tungsten carbide balls have less energy loss in hysteresis than the 440C ball, this means higher accuracy and better repeatability.

For a given displacement, the average Dahl friction of the ϕ0.50-inch diameter, 440C stainless steel ball is always higher than both tungsten carbide and ceramic balls. The ceramic ball has the lowest average Dahl friction of the three ϕ0.50-inch balls. As an example, to displace a distance of +60 μm, the ϕ0.50-inch, 440C ball requires an average 0.0026-pounds of force, while 0.0021-pounds of force are required for the tungsten carbide ball, and 0.0018-pounds of force for the ceramic ball.

The ball travel distance and the total energy loss in hysteresis (the area inside of the hysteresis loop) are shown in Table 3. For the purpose of studying the effect of hysteresis loss of the same ball size and with an identical normal load applied on each ball, let's consider rows 1 and 2 in Table 3 as group A (which has two ϕ0.75-inch balls with 34 pounds normal load), and rows 3, 4, and 5 as group B (which has three ϕ0.50-inch balls with 24 pounds of normal load).

Table 3. Ball Travel Distance Versus Energy Loss

Ball Size (in)	Normal Load (lbs)	Ball Material	Plate Material	Ball Travel Distance (μm)	Total Work Loss in Hysteresis (N-mm)	Hysteresis Work Loss Per Unit Micrometer (N-mm/μm)	Percentage of Non-Linearity in Hysteresis Curve
ϕ0.75	34	440C	Tungsten Carbide	87	0.00012	1.36×10^{-6}	3.89%
ϕ0.75	34	52100	Tungsten Carbide	107	0.00020	1.86×10^{-6}	5.81%
ϕ0.50	24	440C	Tungsten Carbide	178	0.00040	2.26×10^{-6}	4.95%
ϕ0.50	24	Tungsten Carbide	Tungsten Carbide	144	0.00016	1.09×10^{-6}	1.48%
ϕ0.50	24	Ceramic	Tungsten Carbide	172	0.00017	1.01×10^{-6}	4.15%

For group A, the ball travel distance is 87 μm for the ϕ0.75-inch 440C ball, and 107 μm for the ϕ0.75-inch 52100 ball. The difference in ball travel distance precludes direct comparison of the total energy loss in hysteresis from one ball to another.

A polynomial "best fit" curve was drawn over the raw data of the hysteresis loop. Also, linear regression is used to draw a "best straight line" through the raw data of the forward and reverse curves of the loop. Next, calculate the rms between the data points of the "best fit" curve and the hysteresis curve. The deviation of the rms for the "best fit" curve and the "best straight line" is the amount of non-linearity in the hysteresis curve.

Using the preceeding approach, it was found that the most non-linearity in the hysteresis curve is 5.81% for ϕ0.75" 52100 steel ball, as shown in Table 3.

Since the worst case of non-linearity in the hysteresis curve is considerably small (5.81%), it is acceptable to compare the amount of elastic hysteresis loss per unit micrometer by taking the total work loss in hysteresis and dividing it by the ball travel distance.

In Table 3, it is found that for the ϕ0.75-inch 440C ball, the amount of work loss in hysteresis per unit micrometer (1.36×10^{-6} N-mm/μm) is 27% less than that of the ϕ0.75-inch 52100 ball (1.86×10^{-6} N-mm/μm). This suggests that, for a ball bearing system where rolling has not yet occurred, the employment of 440C balls will yield better accuracy and repeatability than 52100 balls, despite the fact that the 440C ball bearing may require higher input torque (higher Dahl friction).

The force-distance hysteresis loop is similar to a typical spring-mass hysteresis loop. Hence, hysteresis is a principal source of damping in a vibrating system. This allows using the experimental hysteresis data to estimate the amount of damping of particular materials. Based on the test data given in Table 3, the 52100 ball is capable of damping out any vibrating source in a relatively shorter time than the 440C ball because the energy loss in hysteresis for the 52100 ball is more than that of the 440C ball.

For Group B (the ϕ0.50-inch ball) even though each ball experienced 24 pounds of preload, the least amount of work loss per unit micrometer belongs to; first the ceramic ball (1.01×10^{-6} N-mm/μm), next the tungsten carbide ball (1.09×10^{-6} N-mm/μm), and last the 440C ball (2.26×10^{-6} N-mm/μm). For ball bearing applications which required high stiffness versus weight, and with low bearing torque and electrical and corrosion resistances, the ceramic ball bearing is a better choice over tungsten carbide and 440C ball bearings.

3.2. Breakaway coefficient of rolling friction

Unlike the construction of the hysteresis loop, the measurement of breakaway coefficient of rolling friction was done by adjusting (in elevation) one of the three adjustment screws on the bottom plate until the top plate began to roll.

Knowing the required elevation distance (d), and the distance from the adjustment screw to the pivot line (L) (which is defined by the location of the other two adjustment screws in the bottom plate), the breakaway coefficient of rolling friction can be calculated as follows:

Breakaway coefficient of rolling friction = d/L

To bias out the initial tilt alignment error in the bottom plate with respect to gravity, another identical set of breakaway coefficient of rolling friction measurements was conducted. However, in this set the adjustment in elevation of the bottom plate was effected in the opposite direction and caused the top plate to roll in the opposite direction.

The summary of the measured breakaway coefficients of friction are shown in Table 2. These coefficients are the averaged results of several trials.

It is also worthy to mention that during the process to obtain the breakaway coefficients of friction, the magnitude of the first data point was always slightly higher than the rest of the data points. The cause for this anomaly is not clear, though it may be caused by cold welding, surface irregularities, and/or Johansen effect. Due to this reason, the first data point was omitted from the data set.

Under the normal load of about 24 pounds, the breakaway coefficient for the ϕ0.50-inch tungsten carbide ball against tungsten carbide plate is 0.00011. It is equal to that of the ceramic ball, and 35% lower than the 440C ball of 0.00017.

Drutowski[1] has shown by experiment that the coefficients of rolling friction for a ϕ0.50-inch diameter cemented-carbide ball, a stainless steel ball, and a 52100 steel ball rolling on 52100 steel plates (all under normal load of about 24 pounds) is 0.00010, 0.00007, and 0.00005, respectively.

Even though the coefficients of rolling friction found by Drutowski[1] do not reflect any correlation to the breakaway coefficients of friction reported in this paper (because of the differences in the ball and plate materials), they all indicate that the magnitude of the coefficients of rolling friction of metal against metal are extremely small.

The reversal of the 440C ball and tungsten carbide plate showed no change in the coefficients of friction (first and second rows of Table 2). However, this may not be a fair comparison because the ball size was changed from ϕ1.00-inch to ϕ0.75-inch, even though the normal load was the same at 34 pounds for both balls.

As the 440C ball diameter decreases from ϕ0.75-inch to ϕ0.50-inch, the breakaway coefficient of rolling friction also decreases by 29%. This may be caused by the dropping of the normal load from 34 pounds for the ϕ0.75-inch ball to 24 pounds for ϕ0.50-inch ball. Also, this suggests that the breakaway coefficient of rolling friction is not simply related to the ball and plate materials, but involves other factors such as contact stress, normal load applied on ball, ball size, surface finish, etc.

The work done by Drutowski[1] had proved that the coefficient of rolling friction is indeed a function of the normal load.

The breakaway coefficients of friction of ϕ0.75-inch balls for 440C, and 52100 steel balls rolling against tungsten carbide plates, are 0.00022 and 0.00050, respectively. Based on this data, we can again conclude that the 440C stainless steel is a better choice for ball bearing material than the 52100 chrome steel.

Likewise, for the ϕ0.50-inch ball bearing, the ceramic and the tungsten carbide are equal in breakaway coefficient of rolling friction, 0.00011, and a better choice over the 440C ball which has a coefficient of friction of 0.00017.

4. EMPIRICAL DATA ACCURACY

As the normal loads were applied on the top plate, the bottom plate experienced bending and the three adjustment screws of the bottom plate experienced a small amount of backlash in the threads, which upset the alignment of the test fixture with respect to gravity. The measurements of the combination error between the bending of the bottom plate and the backlash in the adjustments showed that the maximum error was about 2% for the data.

Another source of error in the test setup was that when tilting the bottom plate with one of the adjustments, the top plate is subject to the component loads which forces it (the top plate) to move laterally. The lateral motion of the retroreflector on the top plate (with respect to the bottom plate) may not always be straight toward the laser detector, it may have some rotation. For a typical hysteresis loop, however, this type of error is induced in the forward path and may be biased out by the reverse path with the same error amount.

Each hysteresis curve was plotted based on the average of three data sets. The achieved repeatability of these data were excellent (5%). The air disturbance in the test lab was the primary cause of data scatter.

5. CONCLUSIONS

Rolling friction is significantly different from sliding friction. The basic laws of Coulomb friction still apply, however the level of complication is greater. The magnitude of the rolling friction force is also much smaller when compared to sliding friction.

Theoretically, for the applications where the ball's travel distance is small enough so that the actual rolling has not yet occurred, the ball with the smallest amount of energy loss in hysteresis is ideal because it gives a higher level of accuracy and repeatability, and lowest friction. On the other hand, for applications where actual rolling has occurred, the ball with the lowest coefficient of rolling friction is usually more suitable. Deciding which ball to choose should also be based on other factors such as contact stress, bearing load, life expectancy, etc.

The more we know about the influence of elastic hysteresis and the coefficients of rolling friction, the better will be the design of precise and efficient ball bearings systems.

6. ACKNOWLEDGMENTS

The author is indebted to Mr. D. Kittell - who provided significant help in the experimental work and his inspiring encouragement, Mr. G. Cheney - who provided guidance and technical support, Mr. C. Delp and Mr. C. McGlynn for their valuable comments, and especially to Mr. C. La Fiandra for his comments and contributions that led to significant improvements in the report.

7. REFERENCES

1. Drutowski, R.C., "Energy Losses of Balls Rolling on Plates", Friction and Wear, Elsevier, 1959, pp 16-35.

2. Nurre, G.S., "An Analysis of the Dahl Friction Model and Its Effect on a CMG Gimbal Rate Controller", NASA Technical Memorandum X-64934, 1974.

3. Dahl, P.R., "A Solid Friction Model", Technical Report SAMSO-TR-77-131, Aerospace Corporation, 1968.

4. Dahl, P.R., "Solid Friction Damping of Spacecraft Oscillations", Technical Report SAMSO-TR-75-285, Aerospace Corporation, 1975.

5. Todd, M.J. and Johnson, K.L., "A Model for Coulomb Torque Hysteresis in Ball Bearings", Int. J. Mech. Sci., Vol. 29 No. 5, pp 339-354, 1987.