

Simultaneous Control of a Precision Linear Stage in Multiple Lubrication Modes with the Complex Lag

Daniel Helmick and William Messner

Department of Mechanical Engineering

Carnegie Mellon University

5000 Forbes Ave.

Pittsburgh, PA 15213

dhelmick@andrew.cmu.edu and bmessner@andrew.cmu.edu

Abstract—Precision linear stages can exhibit two markedly different dynamic behaviors during a single move as the lubrication condition of the bearings changes from elasto-hydrodynamic lubrication to boundary lubrication. For robustness purposes it is desirable to design a single compensator which provides stable behavior for both dynamic modes and transitions between them. One approach is to design a compensator for the elasto-hydrodynamic lubrication dynamics and then to cascade a lag compensator with that controller to accommodate the boundary lubrication dynamics. The lag compensator trades off phase margin of the loop in the elasto-hydrodynamic mode for increased performance when operating in the boundary lubrication mode. This paper demonstrates the superior performance of the complex lag compensator compared with a real lag compensator for simultaneous control of the two operating modes.

I. INTRODUCTION

Precision positioning systems operate over a wide range of velocities as they move distances on the order of tens of centimeters to precision positions. The velocity is high during point to point movements, and the bearing operates in the elasto-hydrodynamic lubrication (EHL) regime characterized by a relatively high DC gain and a relatively low frequency first resonance with modest damping, [1], [2], [3]. As the stage approaches the target position, the velocity decreases dramatically, and the bearing most likely operates in the boundary lubrication (BL) regime characterized by a low DC gain and a relatively high frequency first resonance having low damping.

For robustness purposes it is desirable to design a single compensator that stabilizes both modes. This is the simultaneous stabilization problem. Simultaneous stabilization itself, for the system considered in this paper, is straightforward. The challenging part of the design problem is to make the performance as good as possible in both modes. The approach taken in this paper is to design the controller for a 0 dB open-loop crossover above the first resonance in the EHL mode. The design then proceeds by cascading a lag compensator with the controller to trade off the phase margin in the EHL mode while making the 0 dB crossover for the BL mode as high as possible.

This paper demonstrates the superior performance achieved when employing a complex lag compensator for simultaneous control of the two operating modes compared

with using a double real lag compensator. The complex lag is a second order bi-proper transfer function with a common damping ratio $0 \leq \zeta \leq 1$ in the numerator and denominator polynomials. The real double lag compensator is a complex lag when the damping ratio is 1. Compared with a double real lag compensator, the complex lag's frequency response has a narrower phase notch, a steeper magnitude slope at the frequency of maximum phase lead, and a smaller difference between the high-frequency and low-frequency gain. Using the complex lag compensator, more aggressive control is possible without sacrificing controller robustness.

The complex lag with $\zeta < 1$ is not yet widely used in spite of a number of desirable characteristics. The continuous-time complex lead compensator, the reciprocal of the complex lag, was introduced in [4]. The first application of the discrete-time complex lead compensator appeared in [5]. Oboe and Messner demonstrated the application of the complex lag compensator for phase stabilization in [6]. In [7] Messner et al. employed the complex lag compensator for trading off phase margin for increased gain margin.

This paper is organized as follows. Section II briefly reviews the properties of the complex lag compensator. Section III covers some of the basic tribology important to this paper. Section IV describes the positioning system. Section V shows the controller design. Concluding remarks appear in Section VI.

II. COMPLEX LAG COMPENSATOR PROPERTIES

The complex lag compensator provides an additional design degree of freedom through selection of the damping ratio ζ , while the only degrees of freedom available in double real lag compensators are the frequency of maximum phase lag, ω_m , and the amount of lag at the point of implementation, $2\phi_m$. The formulas for the real lag compensator with a single pole and a single zero are referenced in numerous undergraduate text books on control, e.g. [8].

Important properties of a complex lag compensator are the ratio of high frequency to low frequency gain, width of the phase notch, and slope of the magnitude at the frequency of maximum phase lag. The transfer function of a complex lag compensator having maximum phase lag $2\phi_m$ and unity gain

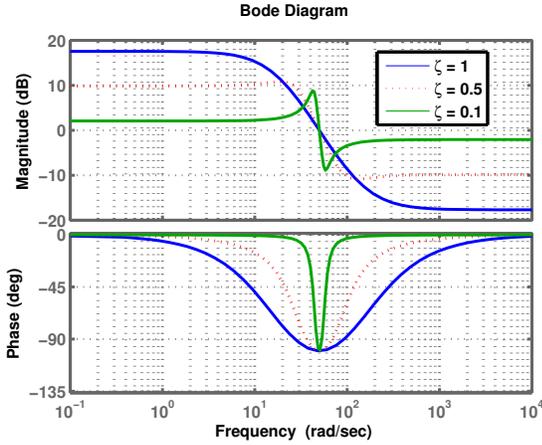


Fig. 1. Comparison of complex and real lag compensators achieving 100 degrees phase lag at 50 rad/sec. As the damping ratio decreases, the phase notch narrows, the slope of the magnitude plot in the transition region increases, and the ratio of low-frequency to high-frequency gain decreases. For $\zeta = 1$ the complex lag compensator is a real double lag compensator.

at ω_m is

$$C_{complex}(s) = \frac{K\omega_p}{\omega_z} \left(\frac{s^2 + 2\zeta\omega_z s + \omega_z^2}{s^2 + 2\zeta\omega_p s + \omega_p^2} \right). \quad (1)$$

The parameter ω_p is the undamped natural frequency of the poles. ω_z is the undamped natural frequency of the zeros, and ζ is the common damping ratio of the poles and zeros. The relationships between ω_m , ω_p , ω_z , ζ , and ϕ_m are

$$\omega_p = \omega_m \left(-\zeta \tan \phi_m + \sqrt{\zeta^2 \tan^2 \phi_m + 1} \right) \quad (2)$$

and

$$\omega_z = \omega_m \left(\zeta \tan \phi_m + \sqrt{\zeta^2 \tan^2 \phi_m + 1} \right) \quad (3)$$

The common damping ratio ζ in the numerator and denominator provides a symmetric phase notch.

Figure 1 shows the frequency responses of complex lag compensators each having maximum phase change of 100 degrees with several different ζ values. Holding the maximum phase lag constant at $2\phi_m$ several properties are clear. As the damping ratio decreases the phase notch narrows and the slope of the magnitude plot at ω_m increases. Both of these features are advantages compared to the real lag compensator.

However, the ratio of the low-frequency gain asymptote to the high-frequency gain asymptote also decreases with decreasing ζ , which is often a disadvantage compared to the real lag. Furthermore as the damping ratio decreases below 0.7 the frequency response increasingly resembles a lightly damped resonance/anti-resonance pair, as the $\zeta = 0.1$ example clearly shows. The use of the complex lag compensator requires balancing these design trade-offs.

III. BASIC TRIBOLOGY

During normal operations a linear stage will transition through multiple friction regimes as the velocity changes

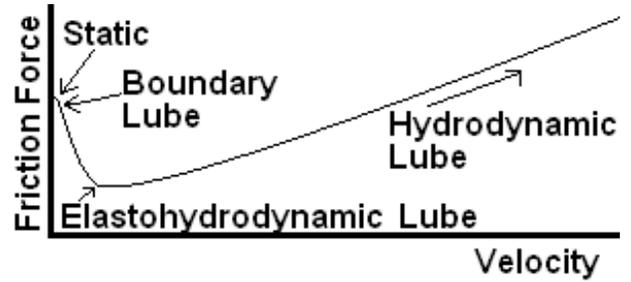


Fig. 2. The basic form of the Stribeck friction curve where friction force is the Y axis, and steady state velocity is the X axis. The modes of Static, Boundary Lubrication, EHD, and Hydrodynamic Lubrication are included in their general locations of the Stribeck Curve.

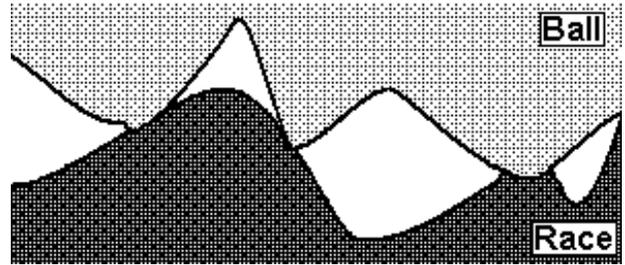


Fig. 3. Dry friction contact as viewed at the asperity level. Forces are carried directly by the asperities, and the asperities have a large deflection due to this contact.

(Fig. 2). Additional details on these lubrication regimes and the implications of these modes on bearings are found in [1], [2], [3]. In contrast to these references, boundary lubrication in this paper refers to the regime where boundary lubricant effects dominate.

The Stribeck Curve encompasses several modes of friction dependent on the steady state travel velocities. Beginning at zero velocity, static friction has no separation between the two surfaces. An illustration of static friction is found in Fig. 3. Any available lubricant is squeezed into the gaps or out of the contact region. Static friction has very high wear and friction forces. The effective spring constant during static friction is much higher.

At very low velocities, the friction force is very similar to static friction, but the effect of boundary lubricants becomes very important. Boundary lubricants often take the form of long molecular chains attached to the surfaces as shown in Fig. 4. Much of the load is still carried by the asperities, but the molecular chains distribute some of the load to a larger area. During these low velocity movements the shear stress of the molecular chains becomes an important factor of the friction force. Often this regime is referred to as pre-rolling because it occurs at the beginning of all movements from a stopped position.

At very high velocities the surfaces will completely separate due to the forces of the entrained lube, and the drag forces dominate. Hydrodynamic lubrication is able to use standard Newtonian assumptions on the surfaces and lube. Elasto-hydrodynamic lubrication is similar to hydrodynamic lubrication because of the full surface separation, but EHD

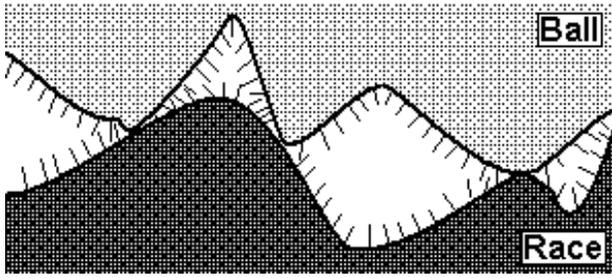


Fig. 4. Illustration of boundary lube effects on low velocities. Some boundary lubes can be diagrammed as long chain molecules attached to the surface. The boundary lube chain molecules share some of the force transmitted between surfaces, and friction forces show a dependence on the shear strength of the boundary lube molecules.

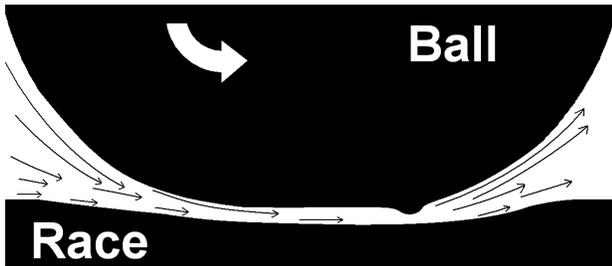


Fig. 5. Illustration of the system experiencing EHD lubrication with direction of spin labeled. Approximate lube velocity is shown.

has non-negligible elastic reaction of one or both surfaces paired with non-newtonian fluid flow. Figure 5 shows the ball and race in EHD lubrication. Hydrodynamic lubrication is similar but with negligible ball and race deformation.

Friction is lowest in the EHD regime. The effective damping of EHD is related to the damping of the fluid, and it is expected to be much higher damping than the BL mode. The friction mode between EHD and BL is often called partial EHD. It has properties of both EHD and BL, but the changes are dependent on the system. This work does not address partial EHD in detail.

The linear positioning stage transitions from zero to a peak velocity and back to zero for every movement. The position stage transitions through several velocity and friction regimes during a sinusoidal frequency response experiment. We assume that lower input amplitudes translate to lower peak velocities for which the corresponding lubrication regimes are BL and partial EHD lube. We also assume that EHD lubrication dominates during larger faster movements. This assumption is supported by [2]. Beyond these broad regime definitions, the effects of transient velocities are not fully addressed in the tribology literature.

IV. SYSTEM DESCRIPTION

The problem considered here is the control of one axis on a multiple stage positioning system built by a leading manufacturer of precision positioning equipment. The axis of interest consists of a stage supported by several recirculating ball bearings and driven by a linear actuator. The mechanical design is proprietary and therefore not shown here. During

point to point movements, which may be tens of centimeters long, the dynamics of the system change as the lubrication regime of the ball bearing changes from EHL toward BL when the stage approaches its target position.

Figure 6 shows two frequency responses obtained by swept sine measurements. The frequency response for the high amplitude input is representative of the dynamics of the stage during high speed movements and the EHL regime. Conversely, the frequency response for the low amplitude input is representative of the dynamics of the stage during the precision low speed movements at the end of a point to point movement and the BL regime.

The large amplitude swept sine measurements show significant variability at the lower frequencies, with both the frequency and the damping of the lowest resonance decreasing as the input amplitude increases. (These other responses are not shown at the request of the manufacturer.) However, all of the large amplitude responses are essentially identical above 300 rad/sec, which enables the servo engineer to design a single controller for the high speed operation of the stage using the high amplitude frequency response of Fig. 6.

The low amplitude swept sine measurements are also very similar. (Again, these other responses are not shown at the request of the manufacturer.) The similarity of the responses indicate that the bearing “locks-in” to the pre-rolling behavior characteristic of BL. Pre-rolling behavior resembles that of a lightly damped mass-spring system where the balls of the bearings are merely deforming rather than rolling. The important characteristics of this operating regime are the low DC gain due to high stiffness and low phase loss up to the first resonance at about 1200 rad/sec. The gain of the dynamics at low speed is significantly lower than the gain of the high speed movements up to 900 rad/sec.

For the purposes of loop shaping the variability at low frequencies of the frequency responses for high input amplitudes can be ignored because those frequency responses are essentially identical at 480 rad/sec where the open-loop 0 dB crossover will occur. For the same reason the variability of the frequency responses of the low input amplitudes can be ignored. The controller design can proceed with the two representative frequency responses of Fig. 6.

V. CONTROLLER DESIGN

There are many potential approaches for controlling systems subject to rapidly changing system behavior. Stiff controllers frequently resolve friction difficulties, but the actuator could not produce sufficient output to implement a stiff controller. Dithering is an approach used in many applications, but this linear stage application does not allow dithering because it can excite resonance modes adversely affecting the motion control and other machinery in the environment. Also, for this system the controller hardware limits the compensator to a PID controller and eight second order filters. Therefore, nonlinear and mode-switching controllers were not considered.

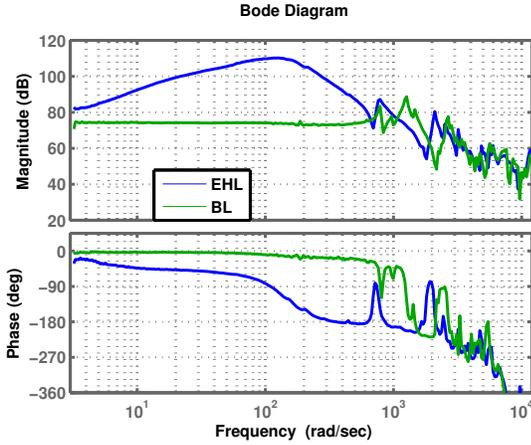


Fig. 6. The representative frequency responses for controller design. The solid line is representative of the high velocity/EHL determined from large amplitude swept sine measurements. The dotted line is representative of the low velocity/BL determined from low amplitude swept sine measurements.

Within these constraints, we chose to design a controller to simultaneously stabilize the two dynamic modes maximizing the open-loop crossover frequency in each mode. In each mode the minimum acceptable performance objectives are the following.

- Zero steady state error for a constant disturbance.
- Phase margin of at least 40 degrees.
- Gain margin of at least 6 dB.

The design approach is to design a nominal controller for achieving these objectives for the high velocity/EHL mode with phase margin to spare. Then a lag controller is cascaded with the nominal controller to increase the low frequency gain for the low velocity/BL mode with some loss of phase margin in the high velocity/EHL mode. The loop shaping is done in continuous-time and then converted to discrete-time with a sampling time of 125 microseconds.

The nominal controller consists of a PID compensator, a cascade of five notch filters, a complex lag, and a second order low pass filter. The controller transfer function is

$$C_{nominal}(s) = C_{PID}(s) \prod_{i=1}^7 C_i(s) \quad (4)$$

$$C_{PID}(s) = \frac{K_D s^2 + K_P s + K_I}{s} \quad (5)$$

$$C_i(s) = \frac{s^2 + 2\zeta_{ni}\omega_{ni} + \omega_{ni}^2}{s^2 + 2\zeta_{di}\omega_{di} + \omega_{di}^2} \quad (6)$$

The values for $C_{PID}(s)$ are $K_D = 3.63e-8$, $K_P = 4.36e-6$, and $K_I = 1.31e-4$. The other constants are defined in Table I.

The complex lag of the nominal controller provides 50 degrees of phase lag at 730 rad/sec with the low damping ratio $\zeta = 0.05$. It is used to reduce the amplitude of the peak at 730 rad/sec without the significant phase loss at lower frequencies that a notch would cause.

Figure 7 shows the frequency response of compensated open-loop for both the EHL (high velocity) mode and the

i	Type	ζ_{ni}	ω_{ni}	ζ_{di}	ω_{di}
1	Complex Lag	0.05	747	0.05	713
2	Notch	0.05	1300	0.3	1300
3	Notch	0.06	2050	.15	2050
4	Notch	0.022	2670	0.1	2670
5	Notch	0.056	3280	0.1	3280
6	Notch	0.39	6000	0.7	6000
7	Low Pass	0.476	1.92e4	0.356	7120

TABLE I
NOMINAL CONTROLLER PARAMETERS

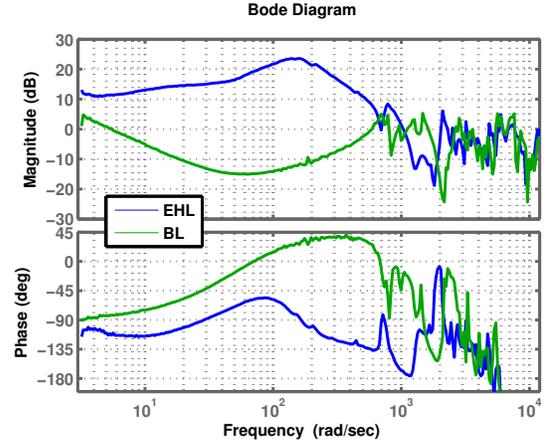


Fig. 7. Bode plot of the open-loop systems with the nominal compensator. Both the EHL mode and BL mode are stable.

BL (low velocity) mode. The phase margin is 50 degrees at 480 rad/sec for the EHL mode with a gain margin of 6.1 dB. The BL mode is also stable, because its magnitude remains less than unity.

The phase margin at the 480 rad/sec crossover in the EHL mode is larger than required. Thus there is an opportunity to increase the low frequency gain of both open-loop responses with a lag compensator at the expense of a reduced phase margin in the EHL mode. The BL mode will benefit the most, because increasing gain increases the 0 dB crossover. There are three constraints on the lag compensator.

- 1) It can reduce the phase by no more than 9 degrees at 480 rad/sec to retain the desired phase margin and robust performance.
- 2) The phase must be greater than -180 degrees for the EHL mode from DC to 4000 rad/sec where the gain margin is measured for robust performance.
- 3) The effects of the lag on magnitude cannot be too wide.

To further quantify the second constraint there can be no more than 100 degrees of phase loss at 60 rad/sec. The third constraint may be clarified with respect to this system by stating that the transition from the low frequency gain asymptote to the high frequency gain asymptote of the lag compensator should occur between the 0 dB crossover frequency of the BL mode loop and the 0 dB crossover frequency of the EHL mode.

Figure 8 shows the Bode plots of the best double lag and the best complex lag found by experimentation that satisfy

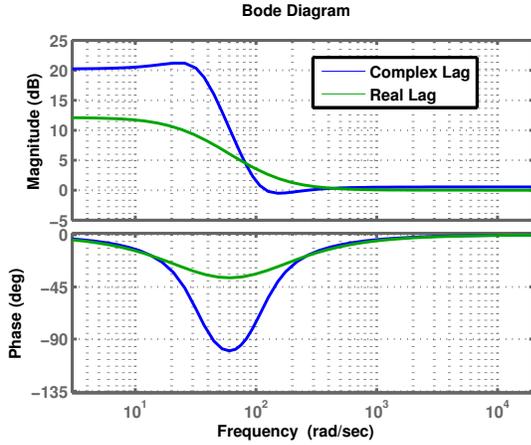


Fig. 8. Comparison of lags satisfying the phase loss constraints. The complex lag has a damping ratio of 0.5 and provides 100 degrees of lag at 60 rad/sec. Both lags provide only 9 degrees of phase loss at 480 rad/sec. The much narrower phase notch of the complex lag allows a larger ratio of low frequency gain to high frequency gain.

	Real Lag	Complex Lag
Gain	0.964	1.033
ζ	1	0.5
ω_m	81.8	107.3
$2\phi_m$	37°	100°
Low Freq. to High Freq Ratio	10.8dB	19.6dB

TABLE II
PARAMETERS OF THE TWO LAG COMPENSATORS

the phase constraints. Both provide 9 degrees of phase loss with a gain of 0 dB at 480 rad/sec. The real lag provides a phase notch of 37 degrees at 60 rad/sec.

$$C_{rlag}(s) = 1.86 \left(\frac{0.5183s^2 + 86.39s + 3600}{s^2 + 86.39s + 1866} \right). \quad (7)$$

The parameters can be found in Table II. The ratio of low frequency to high frequency gain is 10.8 dB. Figure 9 shows that the 0 dB crossover of the BL loop is 5.5 rad/sec when $C_{total}(s) = C_{nominal}(s)C_{rlag}(s)$ where $C_{nominal}$ is given by Eqn. 4. The EHL mode phase margin is 41 degrees and the EHL gain margin is at least 6.1 dB in the region of the -180 degree crossover when using the real lag.

Implementing double real lag compensators at a lower ω_m frequency and a larger maximum phase ϕ_m has undesirable effects on the sensitivity function. The sensitivity function of the BL mode exhibits a large peak after the open-loop 0 dB crossover frequency if the phase lag increases significantly. Additionally, disturbance rejection in the EHL mode is not as good above the reduced ω_m .

The complex lag of Fig. 8 has a damping ratio of 0.5 and provides a phase notch of 100 degrees at 60 rad/sec. It has the transfer function

$$C_{clag2}(s) = 3.20 \left(\frac{0.3228s^2 + 34.09s + 3600}{s^2 + 34.09s + 1162} \right). \quad (8)$$

Both the complex lag and the double real lag have essentially the same phase lag below 15 rad/sec and above 480 rad/sec.

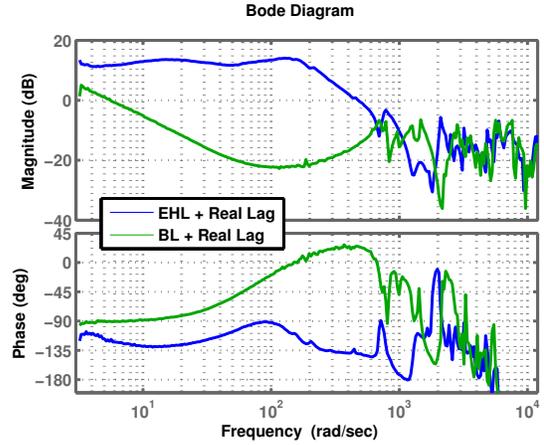


Fig. 9. Open loop frequency responses with the real lag compensator applied. The blue line is the high velocity/EHL mode response and the green line is the low velocity/BL mode response.

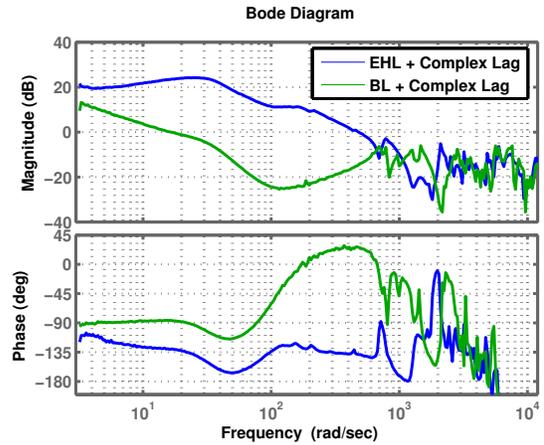


Fig. 10. Open loop frequency responses with the complex lag compensator applied. The blue line is the high velocity/EHL mode response and the green line is the low velocity/BL mode response.

The big difference in phase occurs between those frequencies. The ratio of low frequency to high frequency gain is 19.6 dB, which is 8.8 dB more than provided by the real lag. Figure 10 shows the 0 dB crossover of the BL loop is 20.8 rad/sec when $C_{total}(s) = C_{nominal}(s)C_{clag2}(s)$. The EHL mode phase margin is 41 degrees and the EHL gain margin is at least 6.1 dB in the region of the -180 degree crossover when using the complex lag. Table III shows a comparison of the open-loop and closed-loop results of using the two different lag compensators.

Sensitivity Functions

Table III summarizes the simulated performance of the two controllers based on experimental frequency response data acquired prior to the application of the controls. Improvements in the time response of the system were noted. However, permission was not given to show this data. Figures 11 and 12 show frequency responses of the closed-loop sensitivity functions. The system employing the nominal

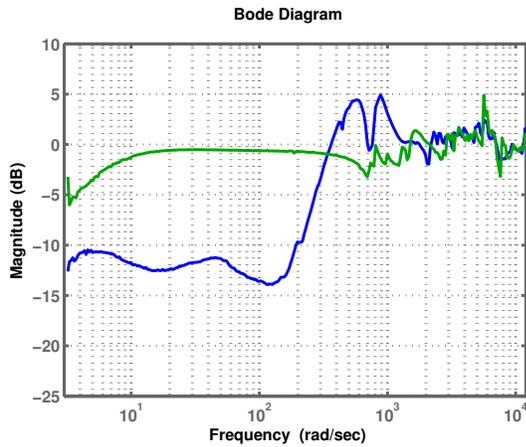


Fig. 11. Sensitivity function of the closed-loop systems with the complex lag applied. The blue curve is the sensitivity function for the EHL mode. The green line are the sensitivity functions for the BL mode.

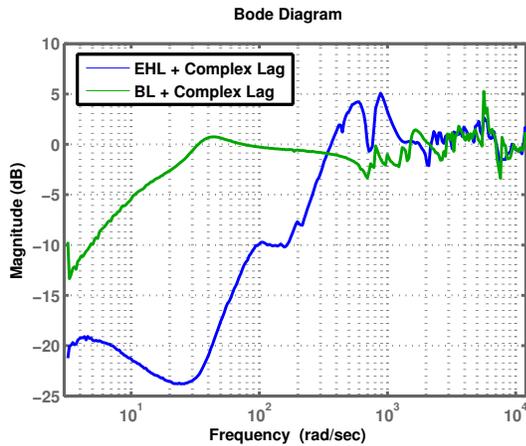


Fig. 12. Sensitivity function of the closed-loop systems with the complex lag applied. The blue curve is the sensitivity function for the EHL mode. The green line is the sensitivity function for the BL mode.

	Real Lag	Complex Lag
0dB crossover of BL	5.5 $\frac{rad}{sec}$	20.8 $\frac{rad}{sec}$
Minimum EHL Phase Margin	6.1 dB	6.1 dB
EHL Sensitivity Peak	4.9 dB	5.1 dB
EHL Rejection of 100 $\frac{rad}{sec}$	-13.6 dB	-9.7 dB
EHL Rejection of 10 $\frac{rad}{sec}$	-11.8 dB	-21.3 dB
BL Sensitivity -3 dB Crossing	5.5 $\frac{rad}{sec}$	18 $\frac{rad}{sec}$

TABLE III

SUMMARY OF OPEN-LOOP AND CLOSED-LOOP RESULTS

controller and real double lag compensator in Fig. 11 has less disturbance rejection overall. There is more disturbance rejection in the frequency range of 3 to 70 rad/sec in the EHL mode when employing the complex lag compensator, although the 0 dB crossover frequency of the sensitivity function is essentially the same as with the real double lag compensator. This is expected because the lag compensator designs have the same open-loop 0 dB crossover frequency and the same phase margin in the EHL mode.

The sensitivity functions for the BL mode show a large difference in the 0 dB crossover frequency, however. Employing the complex lag compensator results in significantly more disturbance rejection below 20 rad/sec. At 3 rad/sec disturbance rejection provided by the complex lag compensator exceeds the disturbance rejection provided by the real double lag compensator by more than 6 dB. The -3 dB point is about 20 rad/sec when using the complex lag compensator and only 6 rad/sec when using the real double lag compensator.

VI. CONCLUSIONS

The complex lag compensator in this work has several desirable characteristics for trading off phase margin for performance. This paper showed the application of the complex lag for a simultaneous stabilization and control problem for a precision positioning system operating in a low velocity/boundary lubrication mode and a high velocity/elastohydrodynamic lubrication mode. The use of this compensator provided more than a factor of three improvement in bandwidth for the low velocity mode compared to a real lag compensator. The disturbance rejection at lower frequencies was more than twice as large when employing the complex lag compensator for both the low velocity and the high velocity modes with good robustness. Other precision positioning and pointing systems with bearings and similar constraints could benefit from the simultaneous stabilization application of the complex lag.

REFERENCES

- [1] J. Williams, "Engineering Tribology," Cambridge University Press, New York, NY, 2005.
- [2] T.A. Harris and M.N. Kotzalas, "Essential Concepts of Bearing Technology," CRC Press, Boca Raton, FL, 2007.
- [3] T.A. Harris and M.N. Kotzalas, "Advanced Concepts of Bearing Technology," CRC Press, Boca Raton, FL, 2007.
- [4] W. Messner, "The development, properties, and application of the complex phase lead compensator," in Proceedings of the 2000 American Control Conference, Chicago, IL, July 2000, pp. 2621-2626.
- [5] W. Messner, "Some advances in loop shaping with applications to disk drives," IEEE Transactions on Magnetics, vol. 37, no. 2, Mar 2001, pp. 651-656.
- [6] W. Messner and R. Oboe, "Phase stabilized design of a hard disk drive servo using the complex lag compensator," in Proceedings of the 2004 American Controls Conference, Boston, MA, June 30-July 2, 2004, pp. 1165-1170.
- [7] W. Messner, L. Xia, M. Bedillion, D. Karns "Lead and Lag Compensators with Complex Poles and Zeros," IEEE Control Systems, pp. 44-54, n. 1, vol. 27, 2007.
- [8] N.S. Nise, "Control Systems Engineering," John Wiley & Sons, Inc, New York, NY, Third Edition, 1994.