https://ntrs.nasa.gov/search.jsp?R=20180001190 2020-05-09T21:08:49+00:00Z

What's New is What's Old: Use of Bode's Integral Theorem (circa 1945) to Provide Insight for 21st Century Spacecraft Attitude Control System Design Tuning

Mike Ruth[1]
Orbital Sciences Corporation, Launch Systems Group , Dulles Virginia 20166

Ken Lebsock[2]
Orbital Sciences Corporation, Technical Services Division, Greenbelt, MD 20770

and

Neil Dennehy[3]
NASA Engineering & Safety Center (NESC), NASA Goddard Space Flight Center, Greenbelt, Maryland 20711

This paper revisits the Bode integral theorem, first described in 1945 for feedback amplifier design, in the context of modern satellite Attitude Control System (ACS) design tasks. Use of Bode's Integral clarifies in an elegant way the connection between open-loop stability margins and closed-loop bandwidth. More importantly it shows that there is a very strong tradeoff between disturbance rejection below the satellite controller design bandwidth, and disturbance amplification in the 'penalty region' just above the design bandwidth. This information has been successfully used to re-tune the control designs for several NASA science-mission satellites. The Appendix of this paper contains a complete summary of the relevant integral conservation theorems for stable, unstable, and non-minimum-phase plants.

$C(s)$	=	transfer function of the controller
d	=	load disturbance
e	=	error signal to the controller
$G(s)$	=	transfer function of the plant
g_m	=	gain margin (dB)
H_o	=	pitch bias momentum (N·m·s)
j	=	imaginary number $\sqrt{-1}$
K	=	proportional gain (N·m/rad)
κ	=	gain constant
k_v	=	velocity constant
LHP	=	Left Half Plane
$L(s)$	=	open-loop transfer function
Ln	=	denotes the natural logarithm
$L_o(s)$	=	proper rational open-loop transfer function
Ms	=	maximum sensitivity
Mt	=	maximum complementary sensitivity
n	=	measurement noise
NMP	=	Non Minimum Phase

[1] Senior Scientist and Technical Fellow, , Member AIAA.
[2] Senior Scientist, Senior Member AIAA.
[3] NASA Technical Fellow for GN&C, Member AIAA.

n_{pz}	=	pole-zero excess
n_{up}	=	number of RHP poles
n_{uz}	=	number of RHP zeroes
p_i	=	the ith unstable pole
p_i^*	=	complex conjugate of p_i
RHP	=	Right Half Plane
r	=	command or reference signal
$S(s)$	=	sensitivity function
s	=	Laplace variable
s_m	=	stability margin
$T(s)$	=	complementary sensitivity function
u	=	controller output signal
x	=	process or system output
z_k	=	the kth unstable zero
z_k^*	=	complex conjugate of z_k
τ	=	time delay in the loop
τ_p	=	time constant of pole (sec)
τ_z	=	time constant of zero (sec)
ϕ_m	=	phase margin (deg)
ω	=	frequency (rad/sec)
ω_{3dB}	=	closed-loop bandwidth (rad/sec)
ω_{ms}	=	frequency at which maximum sensitivity occurs (rad/sec)

I. Introduction

HENDRIK Wade Bode, the American engineer, researcher and inventor, has long been recognized and respected as a pioneer of feedback control theory. He is probably best known to today's generation of controls engineering students for his development of the transfer function gain and phase plots that bear his name, the Bode Plot. However, one of his particularly insightful and important contributions to the field of controls engineering are the integrals of the sensitivity and the complementary sensitivity functions. These are commonly known as the Bode Integrals. In the opinion of the authors these are often overlooked/under emphasized in today's engineering classrooms and are not commonly applied in practice anymore. The Bode integral techniques can provide insights for a control system designer when developing controllers for plants that are inherently difficult to stabilize and control. The greatest limitations on stability and disturbance rejection are related to the presence of poles and zeroes in the Right Half Plane (RHP) and time delays. Control loop performance, in terms of stability and disturbance rejection, is commonly analyzed in the frequency domain. Loop compensation design can be viewed as tuning both the sensitivity and complementary sensitivity transfer functions to achieve stability and disturbance rejection goals. Stability is related to the peaks of these transfer functions as will be shown. The sensitivity transfer function is also known as the disturbance rejection transfer function. Likewise the complementary sensitivity transfer function is commonly known as the closed-loop transfer function. The roll-off characteristics of the sensitivity transfer function determine low-frequency disturbance rejection behavior. The complementary sensitivity transfer function determines the 3-dB controller bandwidth which both dictates closed-loop system performance and high frequency disturbance/noise attenuation. The frequency domain integrals of these transfer functions therefore determine the limitations of loop shaping.

The Bode Integral is a conservation theorem formula that quantifies some of the limitations of feedback control systems. The Bode Integral has two forms: one that applies to stable plants and the second to unstable plants. Simply put the integrals state that the log magnitude of the sensitivity function of a Single Input-Single Output (SISO) feedback system, integrated over frequency, is constant. The log magnitude integrals of the complementary sensitivity function, divided by frequency squared, is also equal to a constant. Per the conservation principle the

actual constant integrated value varies for stable, unstable, and non-minimum-phase plants. The implications of both these assertions will be examined in this paper.

Bode tells us these constants are zero for stable plants with all poles and zeros of the open-loop transfer function in the Left Half Plane (LHP). For unstable plants with poles in the RHP the constant associated with the Bode Sensitivity Integral is a positive value largely determined by the sum of the RHP poles. For the case of a Non-Minimum Phase (NMP) plant with zeros in the RHP the Bode Complementary Sensitivity Integral is a constant determined by the sum of the reciprocals of the NMP zeros and the time delay in the loop. The mathematical treatment of these ideas is addressed in detail in the Appendix of this paper as is the more complicated case where there are both poles and zeroes in the RHP.

Control researcher Gunter Stein, in his 1989 Bode Lecture, reminded the community that the Bode Integrals fundamentally capture and quantify, in the frequency domain, the challenge of designing stable controllers, especially for unstable plants (see References 1 and 2). This formula was often used in the 1950s and 1960s by earlier generations of controls engineers to synthesize single input-single output closed-loop control system transfer functions. In his lecture Stein goes on to comment *"every control theoretician and every control engineer should know these integrals and understand their meaning. Unfortunately, we have not always taught them well."* This paper, inspired by Stein's Bode Lecture, documents some recent applications of the Bode Integrals to modern control system design. One motivation for the NASA Engineering and Safety Center (NESC) sponsorship of this work was to capture and disseminate these classical design principles in a tutorial manner for the Guidance, Navigation & Control (GN&C) Community of Practice online learning site.

Use of Bode Integrals as a control system design tool clarifies in an elegant way the connection between open-loop stability margins, closed-loop bandwidth, and disturbance rejection. Its application to a satellite disturbance rejection design provides clear (but to some extent, unexpected!) insights for trading off spacecraft attitude pointing performance against control bandwidth. The result sometimes opposes the 'conventional wisdom' for maximizing closed-loop performance by maximizing the bandwidth subject to a constraint of preserving reasonable open-loop classical gain and phase stability margins. More importantly it shows that there is a strong tradeoff between disturbance rejection below the controller design bandwidth, and disturbance amplification in the 'penalty region' just above the design bandwidth. For many classes of satellite disturbance signals -- for example, star-tracker noise, uncompensated magnetic torquer inputs, etc. -- this tradeoff may dictate, counter-intuitively, that the best pointing performance is realized at a bandwidth appreciably below the maximum value obtained under standard gain-and-phase margin constraints. The primary objective of this paper is to provide real world design examples including NASA's NuSTAR and SORCE spacecraft, to realistically illustrate the Bode Integral theorem and its application to spacecraft Attitude Control System (ACS) tuning. A secondary objective is to bring the relevance of the Bode Integral to light for the current generation of controls engineers.

The paper is organized into several additional sections. Section II provides a brief refresher on SISO transfer functions and stability margins. Section III summarizes the use of the Bode Sensitivity Integral with an illustration of disturbance rejection for a simple satellite reaction wheel attitude control system. Section IV builds on these results, using a more complex example from the NuSTAR satellite science-mode control system. Section V provides a very brief summary of results from tuning the attitude control loop of the SORCE satellite using the Bode Sensitivity Integral. Section VI treats the effect of pronounced open-loop plant instability in a comparative way using a set of three different aerodynamic vehicle plant dynamics. Section VII provides an example of re-tuning both the sensitivity and complementary sensitivity functions of the INSAT satellite's Non-Minimum-Phase roll/yaw attitude controller. Section VIII provides a summary and lessons learned from the design examples. Finally, the Appendix of this paper provides a detailed mathematical description of the Bode and Poisson (for the case when both poles and zeros exist in the RHP) integrals. It also provides a comprehensive and more general development of frequency-domain integral constraints for stable, unstable and Non-Minimum-Phase plants. The Appendix is organized as a self-contained section for ease of reference.

II. Transfer Functions and Stability Margins

Consider the basic Single-Input Single-Output (SISO) feedback control system shown in the block diagram in Figure 1. The transfer function of the controller is $C(s)$ and that of the plant is $G(s)$. The inputs to the system are the

command or reference signal r, the load disturbance d, and the measurement noise n. The process output is x, the error signal to the controller is e, and the controller output command is u.

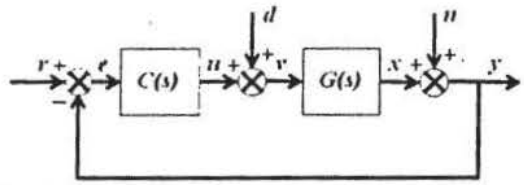


Figure 1. Single-Input Single-Output (SISO) feedback control system.

Define the open-loop transfer $L(s)$, the sensitivity function $S(s)$, and the complementary sensitivity function $T(s)$ as follows:

$$L(s) = G(s)\,C(s)$$

$$S(s) = \frac{1}{1+L(s)}$$

$$T(s) = \frac{L(s)}{1+L(s)}$$

The open-loop transfer $L(s)$ can be written as:

$$L(s) = L_o(s)\,e^{-s\tau}$$

where τ is the time delay in the loop and $L_o(s)$ is a proper rational transfer function. The open-loop transfer function describes how the system output x would respond to an input r if the feedback loop were not closed. Bode or Nyquist plots of the open-loop transfer function are commonly used to determine the stability margins of the closed loop system.

As mentioned above the complementary sensitivity function $T(s)$ is the closed-loop transfer function. It describes how the system output x will respond to an input r when the feedback loop is closed. The desired form of closed-loop transfer function is with unity gain and zero phase shift at low frequencies so that the reference input r is tracked perfectly. The desired amplitude of the closed-loop transfer function rolls off at frequencies that are higher than the system bandwidth in order to filter out sensor noise. The Bode and Poisson integrals are used to study the constraints on selecting the closed-loop system bandwidth.

The sensitivity function $S(j\omega)$ is simply the ratio of the output of the closed-loop system to that of the open-loop system. The sensitivity function tells how the variations in the output are influenced by feedback. At low frequency, a large magnitude of the open-loop transfer function causes the sensitivity to be small. This provides good load disturbance rejection and reference tracking. At high frequency, a small loop gain is used to avoid amplifying measurement noise.

The conservation theorems expressed with the Bode and Poisson integrals are useful for investigating the limitations on achievable closed-loop performance and stability. As detailed in the Appendix the integrals quantify the 'waterbed effect', i.e., improving the disturbance attenuation of controllers at some frequencies results in disturbance amplification at other frequencies. This 'waterbed effect' is notionally illustrated in Figure 2, taken from Stein's Bode Lecture. The Bode and Poisson integrals are conservation laws (Ref. 1). "The integrated value of the log of the magnitude of the sensitivity function is conserved under the action of feedback. The total amount of this quantity is always the same. It is equal to zero for stable plant/compensator pairs, and it is equal to some fixed positive amount for unstable ones." Similar comments apply to the complementary sensitivity function. The combination of RHP poles with non-minimum phase (NMP) zeroes, or time delays, increases the value of the

integrals. These integral constraints impose fundamental limitations on what can be achieved with feedback control (Ref. 5).

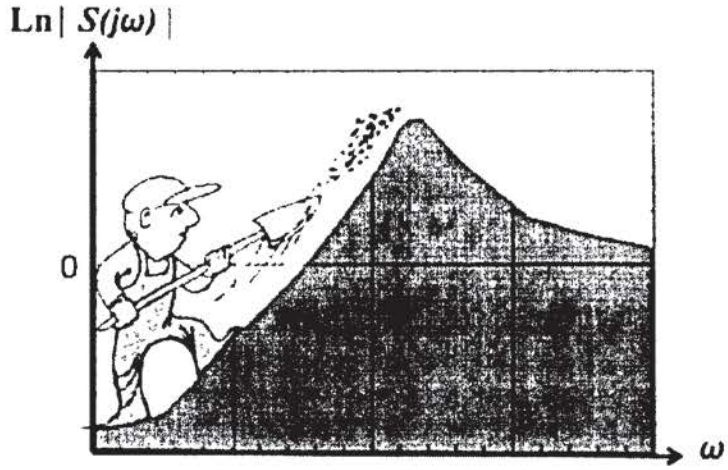


Figure 2: Notional Depiction of the 'Waterbed Effect' (from Ref. 1)

III. Summary of Bode Integral and Simple Application to Satellite Reaction Wheel Attitude Control

Bode's integral theorem, originally developed for feedback amplifier design in the 1930's, may be stated per below:

$$\int_0^\infty \log(S(j\omega)\, d\omega = \int_0^\infty \log(1/(1+L(j\omega))\, d\omega = \pi \sum \mathrm{Re}(p_k) \qquad \text{Eqn (1)}$$

where S is the sensitivity transfer function 1 / (1+L); and the summation is over the range of unstable poles of L(s). See Appendix A for more details. Note that for a stable plant-plus-compensator L(s), the integral evaluates to 0. The original statement of the integral theorem, published in 1945, assumed a stable plant-plus-compensator and the right-hand side was identically zero. Reference 1, the Bode Lecture by Gunter Stein, includes a discussion of additional contraints for unstable plants and this topic is extended in the Appendix of the current paper.

As stated, this integral defines a conservation law. We define the open-loop transfer function ("L" above) as the product C(s)*G(s), where the transfer function of the controller is **C(s)** and that of the plant is **G(s)**. The integral places constraints on our ability to modify the control system loop-shape by simply retuning the parameters in the controller transfer function or even by changing its basic form.

The conservation effect is best illustrated by example. In this section, a typical reaction-wheel control system is used to show the effects of the conservation constraints on controller performance. Note that this example (and the NuSTAR satellite one which follows) are stable plants for which Bode's integral of S(s) evaluates to zero. An example of an unstable plant (aerodynamic vehicle) is also provided for reference in Section V while the subsequent INSAT satellite controller of Section VI – which is strongly non-minimum phase – illustrates extensions to the conservation theorems needed for right-half-plane zeros. Figure 3shows a schematic of a generic control system for a reaction wheel. This controller is conceptually similar to speed-control loops used by both SORCE and NuSTAR. Here the basic torque command from the wheel is integrated to produce a scaled speed command; the speed command itself is realized and controlled via a simple proportional-plus-integral controller.

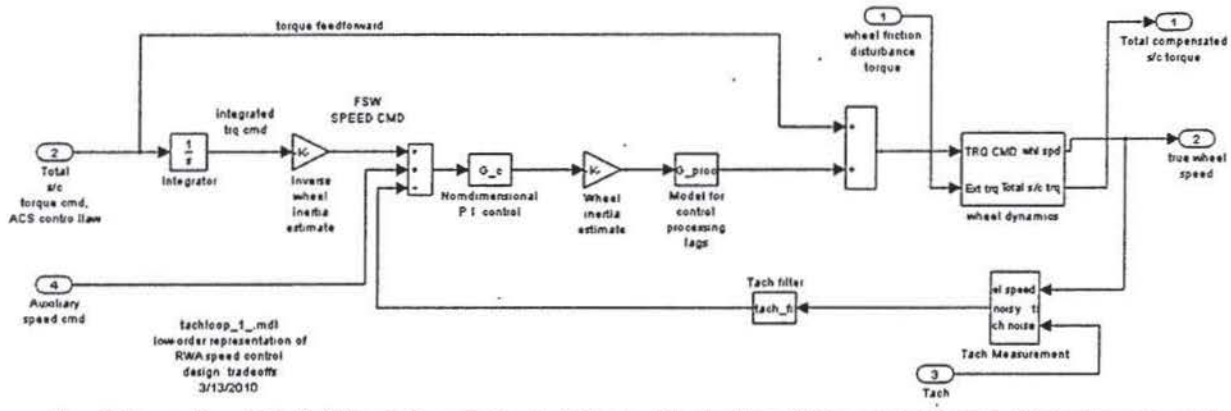


Figure 3: Schematic of Full Wheel-Speed Control Loop (Including Noise and Torque Disturbances and Generation of Speed Command from Torque Command)

Note that the integral of Equation 1 is a scalar function, where S(jω) or S(s) defines an input from a specific disturbance to a specific controlled variable. For this example - as well as the more complex example for NuSTAR in the following section – the focus will be on a fundamental torque disturbance input and a primary outer-loop output. In order to illustrate the integral it is useful first to graphically isolate the primary inputs and outputs, with a goal of defining a scalar loop gain which in turn can be de-composed into a plant (including actuation, measurement, lags etc) and the control law which provides the design degrees of freedom.

Figure 4 below shows the overall wheel control system simplified to represent just the torque rejection portion of the task, for a given reference speed command. At a constant reference speed command (or zero spacecraft wheel torque command) the remaining control task here is to miminize variations in the total wheel torque – sum of wheel motor torque and friction torque – which will be reacted to the spacecraft. For example a constant friction torque needs to be countered on a long time scale by an equal and opposite motor torque to null the net reaction torque to the spacecraft. Thus the relevant disturbance rejection transfer function here is from input 1 (the wheel friction torque) to output 1 (the total wheel torque, which will be reacted to the spacecraft). We would expect that at low frequencies this transfer function would have very small gain, and at high frequencies the transfer would approach unit. This is a standard characteristic of a sensitivity transfer function.

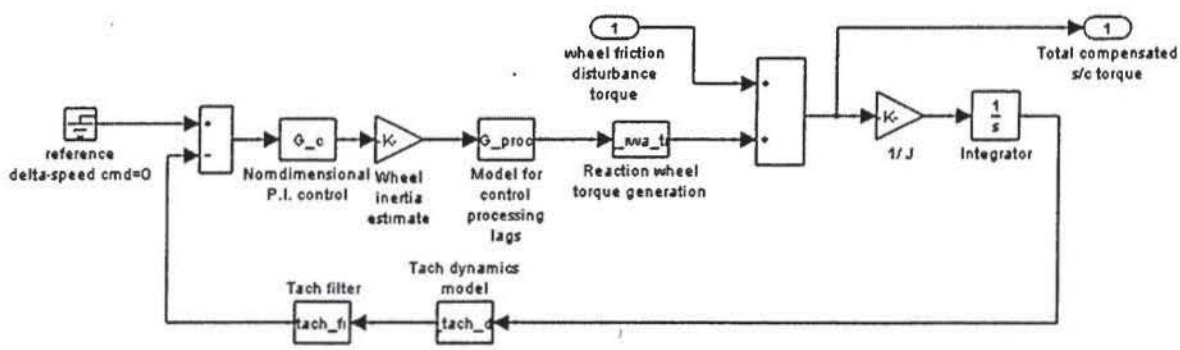


Figure 4: Reaction Wheel Controller Specialized to Torque-Disturbance Portion of Design

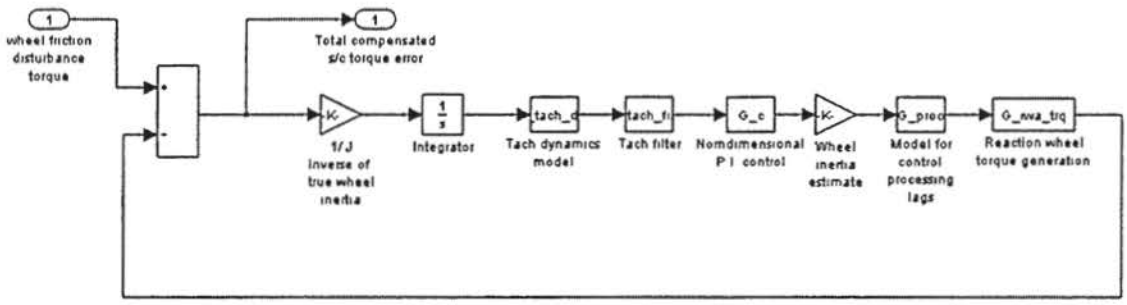
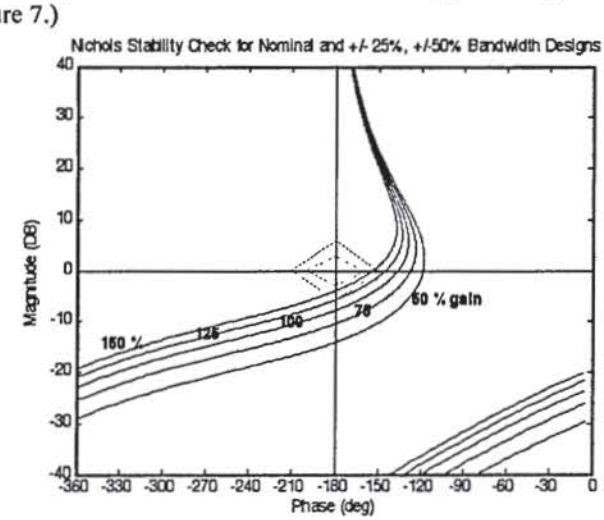


Figure 5: Reaction Wheel Disturbance Rejection Schematic, in Standard Form (Sensitivity Transfer Function S(s) Defined by Input 1 to Output 1)

Figure 5 above clarifies this problem setup further, showing that the needed loop gain to evaluate Bode's integral can be read off now as the product of all the transfer functions in series for the single loop system. By inspection the disturbance rejection sensitivity transfer function S(s) is simply 1/(1+L(s)) where L(s) is the total loop gain. Here, L(s) consists of the plant transfer function elements (wheel dynamics, integrations, measurement dynamics and processing lags) as well as the controller transfer function labeled G_c and the tachometer filter. The degrees of freedom available here are the controller transfer function and the tach. filter. In the rest of this section the effect of the controller transfer function will be assessed and related to the Bode Integral. (Note that with this same problem setup, the external input-1 and output-1 is commonly defined also as the reference command-input to physical-error-signal output, allowing the sensitivity transfer function to describe command-tracking performance rather than environmental disturbance rejection.)

Figure 6 compares the stability properties of several different control-bandwidth designs. Note that the higher-bandwidth designs (100%, 125%) still preserve reasonable gain and phase margins. Intuitively we might expect that the higher- bandwidth design has the best performance, but the sensitivity transfer function of Figure 7 clearly shows that the higher bandwidth controller has better performance – defined as disturbance rejection -- only in the low-frequency range. (Note that the tabulated margin bounds in Figure 6 , columns 3 and 5, are actually drawn from the peak 'Ms' and the formulae in Appendix A, for the sensitivity transfer functions which are plotted in Figure 7.)



Bandwidth Ratio	Peak value Ms for Sens. TF	Gain Margin Bound from Ms (db)	Gain Margin (db)	Phase Margin Bound from Ms (deg)	Phase Margin (deg)
50%	1.34	12.0	14.1	43.9	61.3
75%	1.58	8.7	10.4	36.9	52.5
100%	1.92	6.4	7.8	30.2	43.6
125%	2.43	4.6	5.7	23.7	34.8
150%	3.29	3.2	4.0	17.5	26.0

Figure 6: Reaction-Wheel Controller Margins as Function of Gain (Bounds from S(s) Peak, per Appendix A)

Figure 7 shows the low-frequency performance improvement for higher bandwidth (about 5-10 db more disturbance attenuation. This suggests as expected that the extra bandwidth is more effective in nulling the effect of low-frequency or constant-offset wheel friction torques. The more important result is in the "penalty region" of the Figure. This shows that the high-bandwidth design will have appreciably worse disturbance rejection in that region between 1-2 rad/sec. Further bandwidth increases would make the peaking of the penalty region even worse.

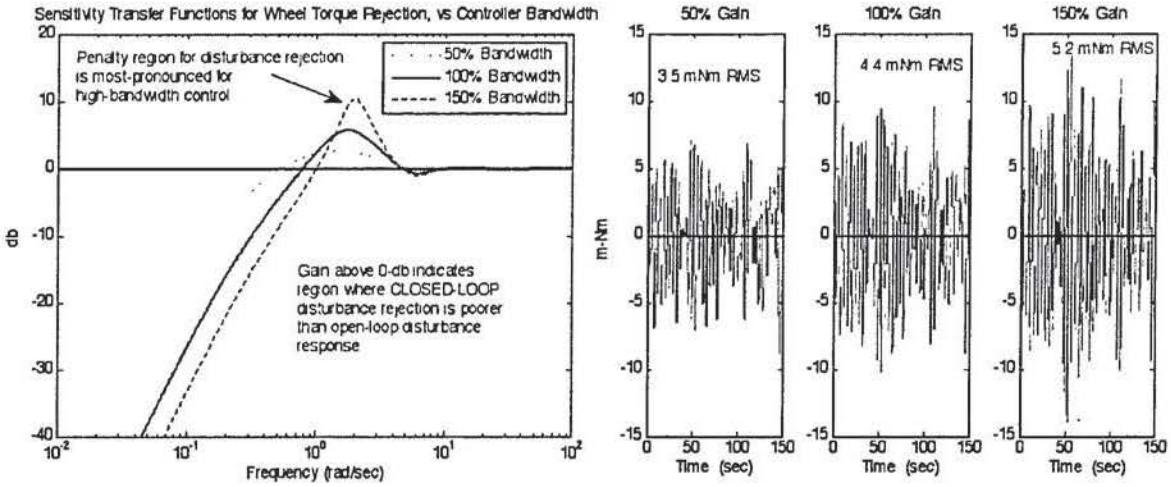


Figure 7: Comparison of Reaction Wheel Disturbance Rejection vs Control Bandwidth (Sensitivity Transfer Function and Disturbance Output for 50%, 100%, and 150% Design Bandwidth)

In many practical control systems the spectrum of disturbances may be appreciable in the 'penalty region' where S(s) is greater than unity, clearly indicating a point of diminishing or negative returns on performance even if classical margins are still acceptable. This rapidly-increasing peaking of S(s) – a closed-loop quantity -- is not evident from the more commonly-used margin-based tools (like Nichols charts) which focus on *open-loop* compensated response. In Figure 7, the time responses for the three different-gain controllers are computed from input 1 to output 1 as shown in Figure 5. The time-domain disturbance is a fairly notional one which still captures features common in satellite or reaction-wheel control – in this case, a three-component disturbance comprising DC or 'near-orbit-rate' disturbance at 5 mNm; 'high-frequency' disturbance near the 1.5 rad/sec penalty region at 2.5 mNm; and 'mid-frequency' disturbance at 0.15 rad/sec also at 2.5 mNm rms.

Note that all three of the controllers have successfully rejected most of the low-frequency and mid-frequency components. The higher-gain control, however, strongly amplifies the high-frequency disturbance. Note also that this amplification is present even for a moderate-gain design possessing 'reasonable' classical margins; here, near 8dB and 45 deg for the 100% bandwidth design. The best-performing controller here (in the sense of lowest-transmission feedthrough of external wheel torques) is actually the lowest-gain one – which also has the highest classical margins.

The clear message here is that the sensitivity transfer function may be used very profitably when directly coupled with information on the spectrum of disturbances. For example if the input was dominated by low-frequency and mid-frequency torques, the 100% bandwidth (rather than 50% bandwidth) controller might be more appropriate. When a range of controllers has acceptable margins, the most-suitable one for disturbance rejection will not depend solely on maximizing bandwidth, but in addition will depend on an assessment of the characteristics of the S(s) penalty region along with the disturbance spectrum. Use of the closed-loop characteristics for a range of designs makes these tradeoffs completely transparent. These results suggest that the common industry tuning practice of focusing mainly on open-loop classical margins and maximum bandwidth, may usefully be expanded to include the closed-loop sensitivity transfer function and the disturbance spectrum.

Finally, for reference Figure 8 shows the conservation of the sensitivity transfer function, and provides some mathematical justification for the shift in the disturbance rejection trend across frequency. Improvements in low-frequency performance will *always* be accompanied by worse performance near the design bandwidth. This is a fundamental limitation of control design. If the low-frequency performance of a lower-bandwidth design is already

acceptable, it may actually a better overall design than a higher-gain controller – even if both preserve reasonable classical stability margins.

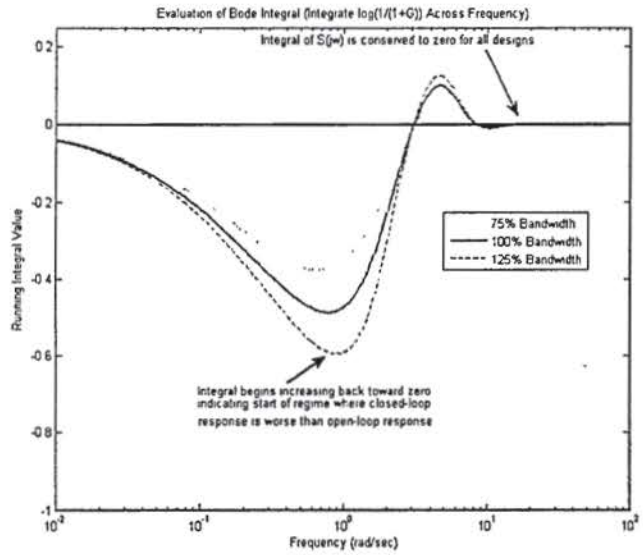


Figure 8: Confirmation of Conserved Value for Bode Integral, for Range of Control-Gain Designs

IV. Closed-Loop Transfer Function Tradeoffs for the NuSTAR Science-Mode Control System

The Nuclear Spectroscopic Telescope Array (NuSTAR) scientific observatory will use high-energy X-rays to detect black holes and other energetic phenomena in the universe. Scheduled for a February 2012 launch into Low Earth Orbit (LEO), the NuSTAR satellite (see artist's depiction shown in Figure 9) is being built by Orbital Sciences under a contract from the California Institute of Technology and the Jet Propulsion Laboratory. NuSTAR has a highly-directional inertia tensor, with the slender axis along the its deployable mast exhibiting a moment-of-inertia two orders of magnitude lower than the other principal axes. Such systems are often designed with a higher-bandwidth for the lower-inertia axis, in order to reject coupling torques from commands to the other axes. Initial ACS design work for NuSTAR proceeded in this direction, which of maximizing the bandwidth subject to maintaining minimum gain and phase stability margins in the range of 6-8 dB and 30-45 degrees.

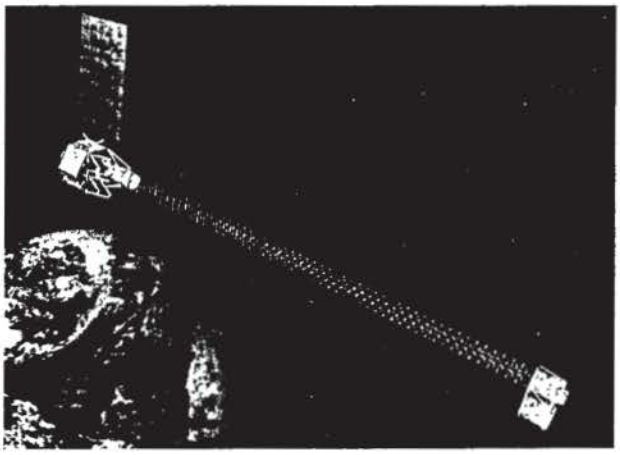


Figure 9: NuSTAR Satellite

Figure 10 depicts the typical disturbance spectrum ranges for NuSTAR (which are typical for science-mode LEO satellites) as well as sensitivity transfer functions for a range of controllers. These controllers span the practical range of useful bandwidth, above which the classical margins would be too low. This figure parallels the sensitivity transfer function sets described for the reaction wheel-controller of Figure 7, but it should be noted that the analysis model here is high-order, including modes and a 16-state torque filter. More importantly, the notional disturbance spectrum ranges are also included here, and the value of plotting S(s) along with the input spectrum is clear. In the area of typical low-frequency torques, all of the controller designs exhibit good roll-off – 20-40 db or more of attenuation. In the S(s) "penalty region" near 0.1 Hz, there are appreciable unmeasured disturbances; this immediately suggests that lower-gain controllers may be most-appropriate.

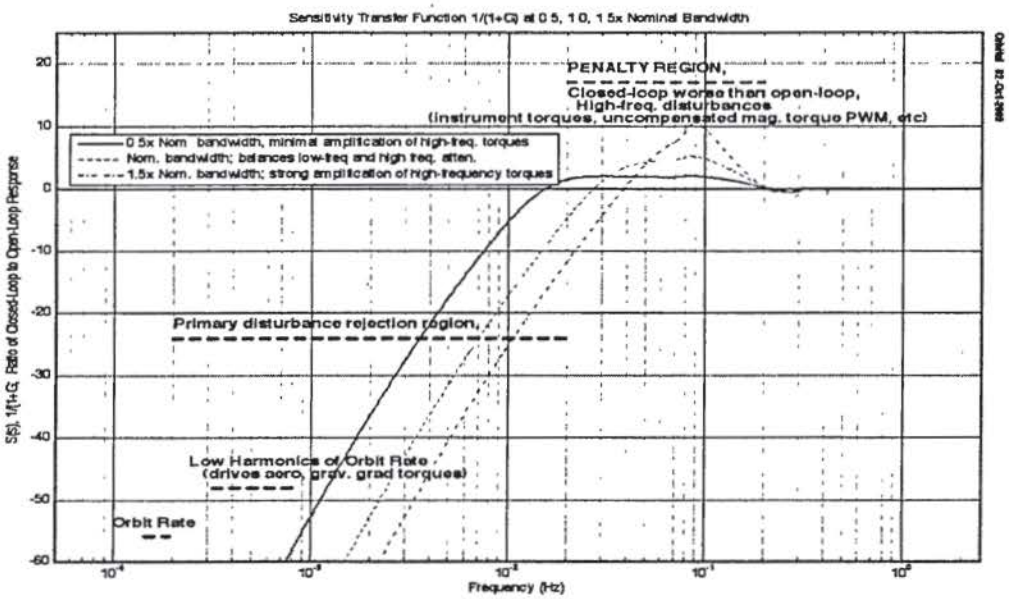


Figure 10: Sensitivity Transfer Function and Disturbance Spectrum Range for NuSTAR

Figure 11 below illustrates how the increased gain in the penalty region affects the output spectrum. In Figure 11, two detailed control tunings are checked, which cover the general range from the '50%' to '100%' tuning of Figure 10. For NuSTAR there is appreciable and broadband disturbance energy from near orbit-rate up to about 0.05 Hz, and then very strong tonal disturbance energy near 0.1 Hz. From this information, along with the knowledge of the peaking in the S(s), we would expect the higher-gain controller to have somewhat better low-frequency response but much worse response near 0.1 Hz. The right-hand plot of Figure 11, containing the output disturbance response, confirms this. The higher-gain control has pronounced narrowband response near 0.1 Hz, which is an unavoidable consequence of the combination of peaking in S(s), and finite input energy near this frequency. In fact, further bandwidth increases would worsen this tonal response near 0.1 Hz. These results show that the higher-gain tuning is already at the point of diminishing returns with respect to disturbance rejection. In addition, the response at 0.1Hz in a sense is worse than a tonal response of similar magnitude in the range of 0.01 to 0.05 Hz, because the corresponding jitter or rate response increases with the first-power of frequency. Hence it was decided to adopt the lower-gain controller, which had better linear margins as well as much-reduced high-frequency disturbance response. Again the use of the sensitivity transfer function makes this tradeoff clear. Finally, it should be noted that the tuning does not use the Bode Integral itself, but rather focuses on the sensitivity transfer function and the input and output spectra. The Bode Integral is 'lurking in the background,' though, because this integral places a hard constraint on the allowable shape of S(s). For this application, the integral dictates that the improvement in disturbance rejection at lower frequencies comes at a strong cost in undesired tonal response near 0.1 Hz.

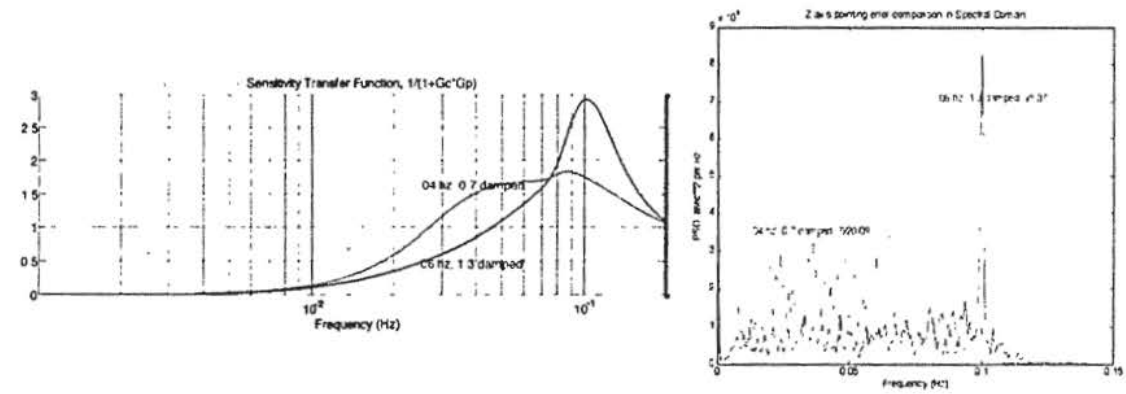
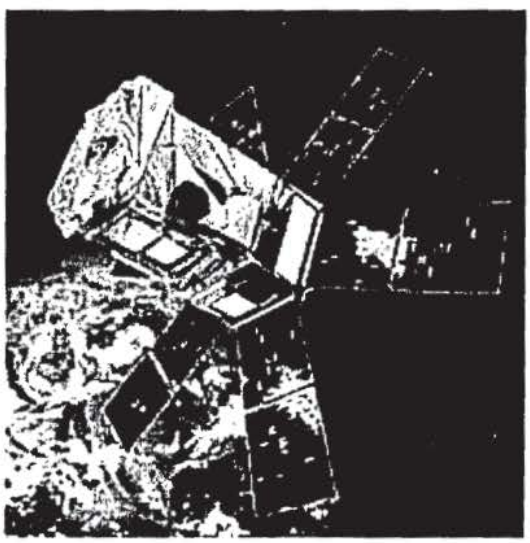


Figure 11: Sensitivity Transfer Function and Spectrum of Pointing Error

V. Designing to Minimize Closed-Loop Disturbance Amplification for the SORCE Satellite



The SORCE (SOlar Radiation and Climate Experiment) spacecraft was launched in 2003 and is currently operating 3 years past its design lifetime. SORCE is a NASA-sponsored mission that is operated by the Laboratory for Atmospheric and Space Physics (LASP) at the University of Colorado (CU) in Boulder, Colorado. Orbital Sciences developed the vehicle bus (artist's depiction shown in Figure 12) for this mission, and the science-mode ACS design proceeded in several phases. The performance changes improvements resulting from several ACS re-turnings are illustrated in Figure 13.

Initial work consisted of adjusting the PID gains from a preliminary baseline value to maximize margins; the performance benefit is appreciable here, and is shown by the difference in the pointing error between the first subplot and the second subplot. The goal of this re-tuning was a fairly typical one: maximize the control bandwidth while preserving generous stability margins of 8 db and 45 deg. The horizontal boundaries of each subplot define the notional pointing error requirement, which was met but not with much margin.

Figure 12: SORCE Spacecraft

Examination of the results from Figure 13, which were developed using a high-fidelity flight simulation, reveals that the pointing errors for the middle sub-plot were actually being driven by external noise. SORCE employs a gyro-less science mode control system using data from two star trackers, and the attitude rate reference is formed from an averaged, divided-difference of tracker measurements. The spectrum of star tracker noise lies in the penalty region for the sensitivity transfer function (the region of S(s) >1), and increases in the control gain to increase bandwidth and move the penalty region well-above this spectrum were not feasible due to stability constraints. This suggested that further performance improvement might be possible by a significant decrease in control bandwidth. The third subplot shows this effect, revealing that a substantial bandwidth reduction of 50% provided much better pointing performance. SORCE has been flying with a reduced bandwidth (relative to original tuning) for the last 7 years and is meeting all its science requirements.

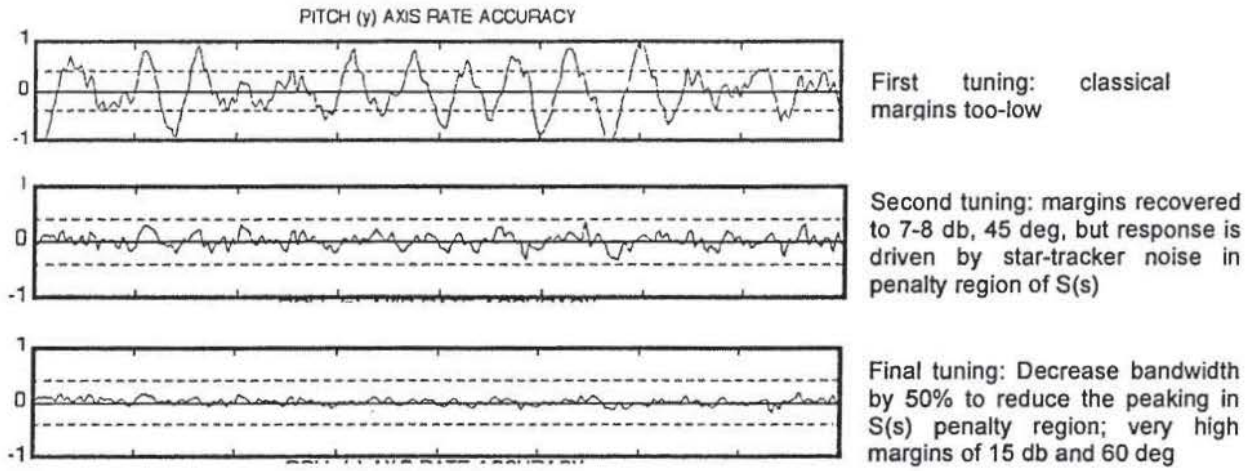


Figure 13: SORCE Pointing Error Improvement for Successive Retuning (Bandwidth decreased by 50% between Second and Third Tuning)

VI. Effect of Unstable Aerodynamic Vehicle Plant Dynamics

As noted earlier, increasing plant instability makes the control design task fundamentally more difficult; this point is discussed at length in the Bode Lecture (Ref. 1) as well. Loosely speaking, a controller for a plant with a high degree of open-loop instability will tend to have poorer disturbance rejection and poorer stability margins than a stable plant at similar bandwidth. The constraint implied by the Bode Sensitivity Integral for unstable plants will dictate that the 'penalty region' where the sensitivity exceeds unity is greater than it is for a stable system with the same bandwidth.

Plants which are highly unstable will tend to exhaust all the design degrees of freedom just in the effort to stabilize the plant dynamics, with no leverage left over for other useful goals such as improving disturbance rejection. This effect can be seen in Figure 12. Compensated open-loop Bode magnitude and phase responses for L(s) are provided Figure 14 for a plant-plus-compensator set representative of booster or missile dynamics. Low-order airframe modeling is discussed in detail in Reference 10. Three different vehicle dynamics cases – ranging from stable to very unstable – are illustrated in the Bode plots.

A linear model for the vehicle dynamics typically includes a lightly damped complex pole pair (for the stable airframe) or a pair of real-axis poles of similar magnitude, one of which is in the RHP (for the unstable airframe). For a severe enough degree of instability the plant may not be stabilized in the sense of preserving classical margins. The different degrees of instability spanned by these three airframe cases show the progression of increasing design difficulty. At similar control bandwidth the margins decrease progressively going from Case1 to Case 3 (see Figure 14) as the degree of instability increases. The gain margins for Plant 3 (at near 4 db at both high and low frequency) are already constraining the design and cannot be improved much -- and there is no real freedom left to address tracking or disturbance rejection. The corresponding sensitivity transfer functions (see Figure 15) provide a similar message and show the integrated constraint imposed by instability. The 'penalty region' and area of worsened margin is broader and higher for the most-unstable plant; simultaneously the low-frequency disturbance rejection is also the worst. In other words, the increased area of the penalty region is not available to improve low-frequency tracking or disturbance rejection in the unstable case.

Table 1: Plant Dynamics & Controller Dynamics (Three Cases of Different Degrees of Airframe Instability)

Case	Dynamics	Controller
1) Stable	$\frac{150(s+0.936)}{(s^2+s+60)}$	$\frac{0.11263(s+6)}{s}$
2) Unstable but Acceptable	$\frac{150(s+1.064)}{(s+8.262)(s-7.262)}$	$\frac{0.15222(s+6)}{s}$
3) Unstable and Unacceptable	$\frac{150(s+1.128)}{(s+11.47)(s-10.47)}$	$\frac{0.17201(s+6)}{s}$

The transfer functions for the plant dynamics and the controllers for the three cases are provided in Table 1. The two elements of the control system that are common to all three cases include an actuator for which the transfer function is:

$$\frac{15790}{(s^2+150.8s+1.579\times10^4)} \qquad \text{Eqn (2)}$$

as well as the lumped system lag transfer given by:

$$\frac{(s^2-300s+3\times10^4)}{(s^2+300s+3\times10^4)} \qquad \text{Eqn (3)}$$

The lag transfer function in Eqn.(3) is essentially a simple model for a number of system elements lumped together: the gyro dynamics, the gyro internal filtering, messaging on the vehicle's data bus to receive and unpack the gyro data and a continuous model of the zero-order hold. For a typical of a medium-performance system these individual lag effects can be rolled up into a single lag model of approximately 20 milli-seconds. A higher-order control design and analysis process would include individual transfer function models for gyro dynamics, filtering, computational delays, sampler models, etc.

This result provides strong mathematical justification for the long-established design dilemma for flight control of unstable vehicles: often the plant instability and lags are so constraining that it is all the control designer can do just to stabilize the airframe. Once the airframe is stabilized there may be little design degrees of freedom left for other goals such as load relief or increasing response speed. If better tracking or disturbance rejection is required, we should not expect any further 'magic' from control design by itself. The only recourse for the designer to truly improve closed-loop response may often be to improve the plant itself, by use of higher sample rates, higher-bandwidth actuators, tighter management of the center-of-gravity envelope, etc. The process of 'improving or managing the plant dynamics' is really a process of requirements management. This is a Systems Engineering

function which is important to establish very early in a vehicle development program. Plotting of the sensitivity transfer functions and associated integral constraints, for a range of vehicle dynamics options, is a tool which the GN&C engineer can profitably use to develop and communicate subsystem requirements to the rest of a multidisciplinary design team.

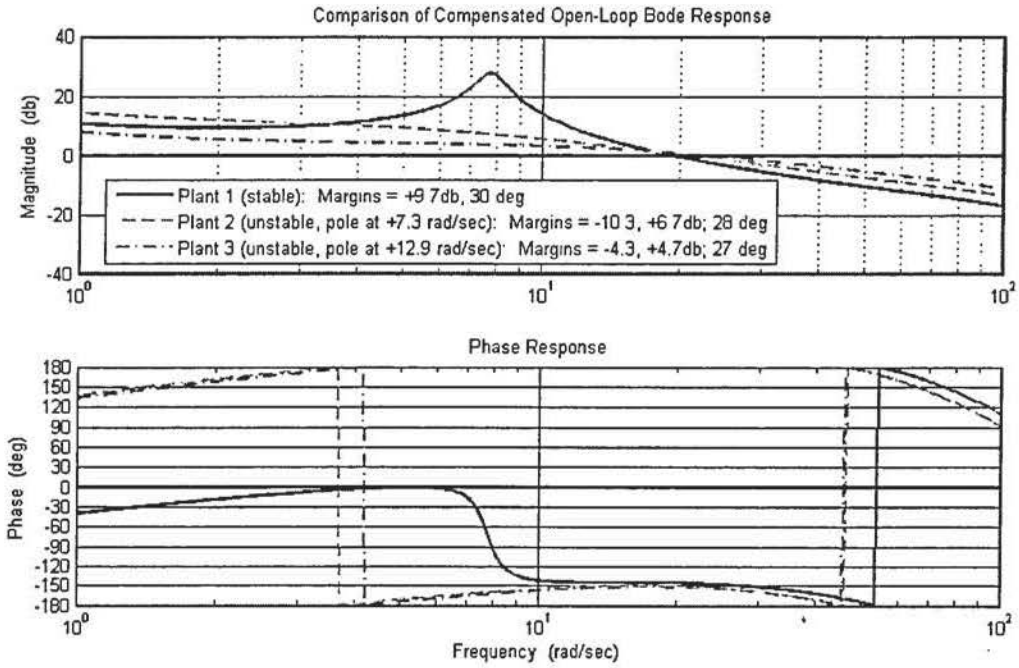


Figure 14: Open-Loop Compensated Gain-Phase, Margins, for Different Degrees of Airframe Instability

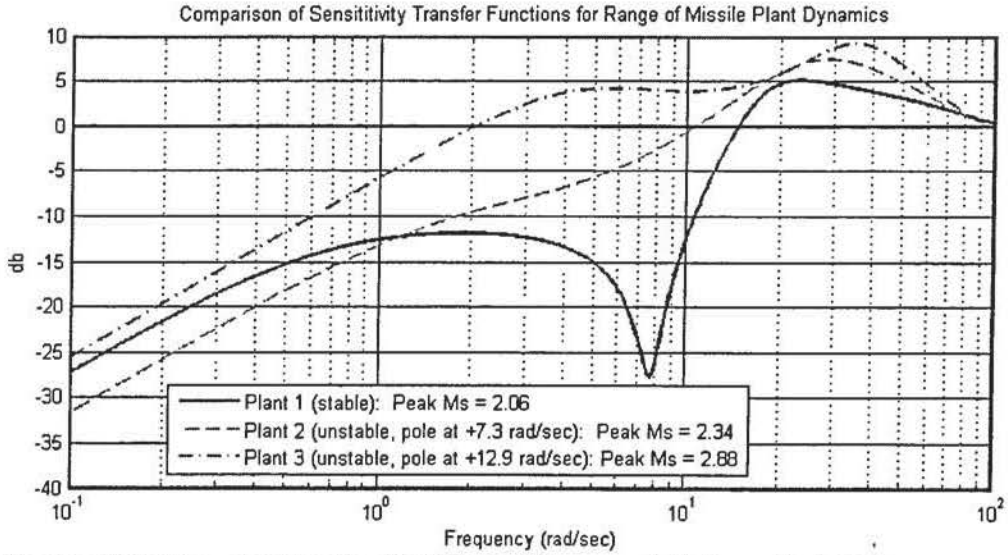


Figure 15: Sensitivity Transfer Function for Different Degrees of Airframe Instability

VII. Illustration of Integral Constraints for Non-Minimum-Phase Spacecraft System: INSAT

The Insat-1 satellite system provided TV broadcast, telecommunications, and meteorological services to India from Geosynchronous orbit starting in the early 1980s. Star trackers and long life gyros were not available for commercial satellite attitude control systems during that time period. Horizon sensors were used to provide pitch and roll attitude sensing during normal on-orbit operation. The example considered in this section is concerned only with the normal mode roll/yaw control system; it is an unconventional design because it includes a non-minimum-phase zero.

Insat-1A was launched into geosynchronous orbit (GEO) by a Delta expendable launch vehicle in April 1982. The on-orbit attitude control system consisted of a small variable speed wheel aligned with the yaw axis, two larger variable speed momentum wheels canted off slightly from the pitch axis in the pitch-yaw plane, redundant horizon sensors to measure pitch and roll errors, and an assembly of hydrazine thrusters to provide momentum desaturation torques and stationkeeping thrust. Varying the pitch momentum wheel speed ±5% about a nominal bias value controlled the pitch attitude. Roll attitude was actively controlled and roll/yaw nutation was damped by varying the angular momentum along the yaw axis. The pitch momentum bias of the wheels provided yaw restraint and, together with the orbital kinematics, transferred yaw errors into roll for correction by the yaw momentum control system. Excessive accumulation of angular momentum on the wheels due to secular environmental torques was prevented by firing a short pulse from a hydrazine thruster. The pitch axis momentum bias and the roll/yaw desaturation deadband were determined by pointing accuracy requirements and the disturbance torques acting on the spacecraft.

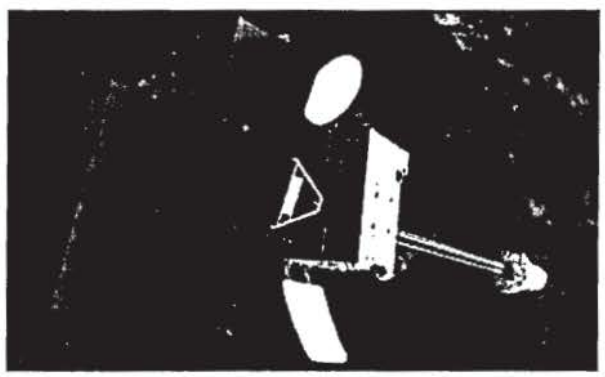


Figure 16: Artist's View of INSAT-1A Satellite

The non-minimum-phase control law was originally proposed for LEO spacecraft in 1967 (Ref. 8). The controller transfer function from the measured roll attitude to the commanded yaw axis wheel momentum is:

$$C(s) = \frac{K(1 - s\tau_z)}{s(1 + s\tau_p)} \qquad \text{Eqn (4)}$$

This control system design has been implemented on many geosynchronous communications satellites, including INSAT, ARABSAT, and GOES spacecraft. Representative spacecraft and control system parameters are taken from Reference 9, Chapter 9.2, entitled "Attitude Control Design for a Bias-Momentum Stabilized Spacecraft". For this example the moments of inertia are $\{I_{xx} = 3026, I_{yy} = 440, I_{zz} = 3164\}$ kg·m^2. The pitch axis bias momentum is H_o= 91.4 N·m·s which gives a nutation frequency of 2.954 E-02 rad/sec. The orbit rate roll/yaw coupling frequency is 7.292 E-05 rad/sec. Ignoring the high frequency flexible body dynamics, the numerical plant transfer function from yaw axis wheel momentum to roll attitude is:

$$G(s) = \frac{9.5267\left(s - \frac{7.292j}{10^5}\right)\left(s + \frac{7.292j}{10^5}\right)}{10^6\left(s - \frac{2.948j}{10^2}\right)\left(s + \frac{2.948j}{10^2}\right)\left(s - \frac{7.268j}{10^5}\right)\left(s + \frac{7.268j}{10^5}\right)} \qquad \text{Eqn (5)}$$

The values for the controller parameters used in Reference 9 are K=0.236 N·m/rad, τ_z = 193.45 sec, and τ_p = 48.36 sec. The root loci using these baseline parameters are shown in the following Figure 17.

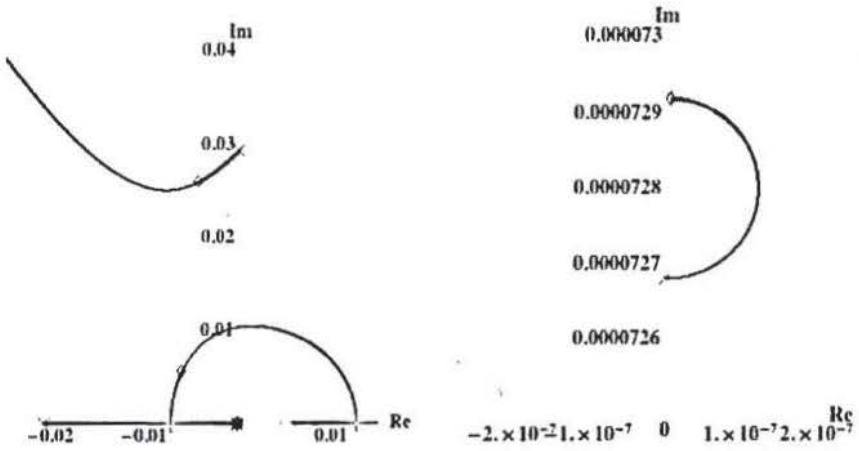


Figure 17: Root Loci for the Non-Minimum-Phase Zero Control System.

The small-scale figure on the left hand side shows that the roll motion is well controlled with a damping ratio of approximately 0.7 and that the higher frequency nutation dynamics have approximately the same settling time but with a damping ratio of about 0.17. The non-minimum-phase zero causes the roll dynamics locus to cross into the right half plane as the gain increases; this would ultimately cause the roll root to become unstable.

The large-scale locus on the right shows an expanded region around orbit frequency. Even though the imaginary pole is smaller than the imaginary zero, the locus departs into the right half plane and never comes back into the left half plane. This occurs because the loop gain is negative. The closed loop unstable orbit rate roots are located at 6.69 E-09 +/- 7.29 E-05 j which corresponds to a time constant of 4 ¾ years. Although the real part is very small, the orbit rate root must be stabilized with an outer loop to provide momentum desaturation. INSAT 1A flew with a mass expulsion desaturation system but the mission was abandoned after a year and a half when its attitude control propellant was exhausted. An alternate magnetic system using an environmental torque estimator was subsequently proposed to provide momentum desaturation (Ref. 7).

The sensitivity and complementary sensitivity transfer functions for the controller using these baseline parameters are shown with dashed lines in the Figure 18. The sensitivity function displays good low-frequency disturbance rejection except at orbit rate where there is no attenuation of disturbances. This agrees with the root locus result that showed a small positive root at orbit rate. The performance of the closed-loop system is worse than that of the open-loop system in the decade from 0.0026 to 0.0262 rad/sec. Disturbances in this frequency range would be amplified because the sensitivity is greater than unity. The maximum value of the sensitivity is Ms = 2.36. Using the relationships between the maximum sensitivity and the lower bounds of the gain and phase margins given in the appendix:

$$g_m \geq 20\mathrm{Log}_{10}\left[\frac{Ms}{Ms-1}\right] = 4.79\,\mathrm{dB} \quad \text{and} \quad \phi_m \geq 2\mathrm{ArcSin}\left[\frac{1}{2Ms}\right] = 24.5\,\mathrm{deg} \qquad \text{Eqn (6)}$$

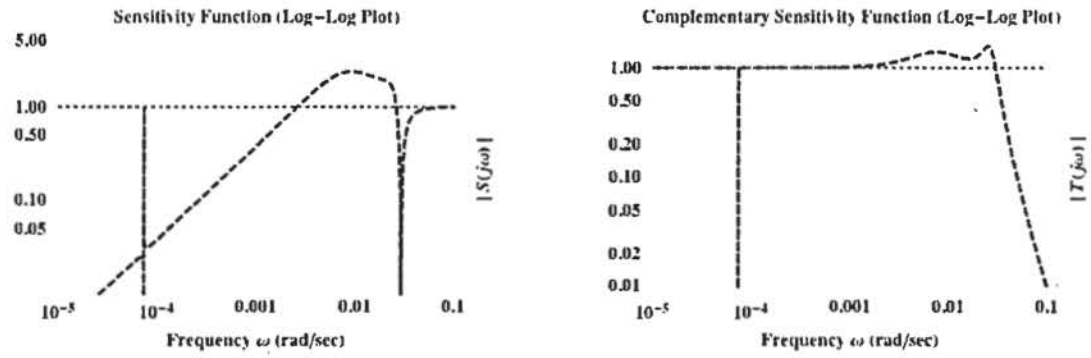


Figure 18: Baseline Sensitivity and Complementary Sensitivity Functions

The actual margin, $g_m = 4.86$ dB, is only slightly larger than the lower bound estimate of 4.79 dB. However the actual phase margin, $\phi_m = 51.7$ deg, is more than twice as large as the lower bound estimate of 24.5 deg. The stability of the control loop could be improved to have a gain margin closer to 6 dB if the sensitivity function were retuned to decrease its peak value.

The closed-loop bandwidth is defined as the frequency where the complementary sensitivity function crosses through 3-dB from above. The bandwidth of 0.0316 rad/sec is sufficiently large to provide nutation damping. At higher frequencies the complementary sensitivity rolls off steeply at 60 dB per decade. This provides excellent high frequency horizon sensor noise attenuation. However the complementary sensitivity is greater than unity for more than a decade at lower frequencies. In fact it has 2 peaks: $|T(.0075)| = 1.40$ and $|T(.025)| = 1.59$. Retuning the complementary sensitivity to reduce these peaks would improve control system performance within the bandwidth.

The Bode integrals impose constraints on reshaping the sensitivity and complementary sensitivity functions to improve stability and performance. There are no RHP poles but there is a non-minimum-phase zero located at $1/\tau_z$. The Bode Integral for Sensitivity is:

$$\int_0^\infty \mathrm{Ln}\left|S(j\omega)\right| d\omega = 0 \qquad \text{Eqn (7)}$$

This is simply a restatement of the waterbed effect; reducing the Ms peak value will require broadening the range over which sensitivity is greater than unity. The control loop left hand plane pole at $1/\tau_p$ determines the low frequency end of this range. Increasing τ_p will stretch out the region where the function $|S(j\omega)|>1$ and thereby reduce its peak value due to the waterbed effect. The reduction in Ms will result in an improvement in stability.

The open-loop transfer function has no time delays and has only a single pole at the origin of the complex plane. Consequently the Bode Complementary Sensitivity Integral is:

$$\int_0^\infty \frac{1}{\omega^2}\mathrm{Ln}\left|T(j\omega)\right| d\omega = \pi\tau_z - \frac{\pi}{2k_v} \qquad \text{where} \qquad k_v = 0.011\,K \qquad \text{Eqn (8)}$$

Decreasing the integrated area under the curve can reduce the peaks in the complementary sensitivity function. This can be accomplished by decreasing the control loop's proportional gain, K, which in turn will reduce the velocity constant, k_v. The value of the Bode integral will thus be reduced because the velocity constant appears in the denominator of a negative term.

For example, the transfer functions can be retuned to achieve better stability and performance by increasing τ_p by 20% and decreasing K by 10%. The sensitivity and complementary sensitivity transfer functions for the controller

using these revised parameters (solid lines) are compared to those with the baseline parameters (dashed lines) in the following Figure 19.

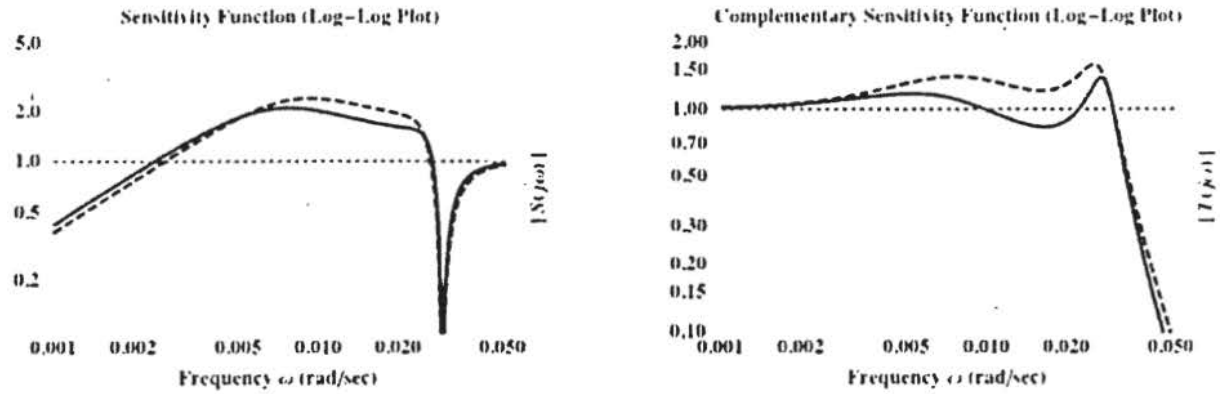


Figure 19: Comparison of Sensitivity and Complementary Sensitivity Functions for the Baseline and the Retuned Designs

The peak sensitivity has been reduced from 2.36 to Ms = 2.07 primarily due to the increase in τ_p. This caused the gain margin to increase by 1.12 dB to g_m = 5.98 dB and the phase margin to increase by 3.0 deg to ϕ_m = 54.7 deg. The penalty associated with this improvement in stability is that the low frequency disturbance attenuation has been slightly degraded. Decreasing the loop gain reduced the area under the complementary sensitivity and caused the curve to be more balanced about unity within the control system's bandwidth. The complementary sensitivity peaks were reduced to |T(0.0052)|=1.18 and |T(0.026)|=1.38. The closed-loop bandwidth was slightly reduced to 0.0312 rad/sec which has a negligible effect on performance

VIII. Conclusions

This paper has addressed frequency-domain constraints for control design, from an applications perspective. Significant additional material on the theory for these constraints is provided in the Appendix. The applications results suggested the following trends:

- There is a strong payoff for using the closed-loop transfer functions (sensitivity and complementary sensitivity) *together* with the input disturbance spectrum. For several of the satellite examples considered, improved performance was obtained by a gain reduction even though the classical margins for higher-gain designs were already acceptable. The shape of the sensitivity transfer function in the 'penalty region' above unity (zero dB) provides clarification of cases where better overall disturbance rejection may be achieved with lower gain. In general, these cases may occur when there is significant disturbance energy near the penalty region.

- As noted also in the Bode Lecture by Stein, unstable plants place appreciably more severe restrictions on the shape of the closed-loop transfer functions, than do stable ones. The flight control example Section VI showed that for increased degree-of-instability the design degrees of freedom tend to get completely exhausted by the stabilization task, with minimal leverage available for other tasks such as load relief or improved tracking.

- Extensions to the Bode integral constraints are available (see the Appendix) which address additional issues such as Non-minimum-Phase plants with the additional complication of unstable poles. The Poisson integral constraints show why neighboring right-half-plane poles and zeros are exceptionally constraining, and should be 'designed-out' or avoided.

- The Bode Integral theorems 'lurk in the background' and control the shape of the achievable open-loop and closed-loop transfer functions which the designer may work with more directly. Constraints from the integrals place fundamental bounds on the shape of the sensitivity and complementary sensitivity functions and these constraints cannot be evaded by high-end mathematical control design strategies.

- The integral constraints (for unstable plants in particular) may provide guidance early in the systems design process for areas where the plant dynamics may profitably be improved. The constraints will be useful early in the Systems Engineering process to levy balanced requirements on areas such as latency, bandwidth, and the allowable degree open-loop instability.

Acknowledgments

The authors first wish to acknowledge Bode's foundational work in network analysis and feedback amplifier design that paved the way for advances in the field of control system theory and practice. He was awarded the Presidential Certificate of Merit for his work in this area in 1948. Gunther Stein's excellent 1989 Bode Lecture (and its publication in full in the IEEE Control Systems Magazine in 2003) provided the motivation for this paper. The authors would also like to thank Scott Glubke (GSFC Code 590 Chief Engineer) and Eric Stoneking (also of GSFC Code 590) for their thoughtful review and commentary of this paper.

References

[1]Stein, G , "Respect the Unstable," *AIAA Journal*, Vol. 24, No. 11, 1986, pp. 1872, 1873. Control Systems Magazine, IEEE Volume 23, Issue 4, Aug. 2003 Page(s): 12 – 25.

[2] Bode, H. W., *Network Analysis and Feedback Design*, Van Nostrand, New York, 1945.

[3]Boyd, S. and Pearson, B. *Linear Controller Design: Limits of Performance,* Prentice-Hall, New York, 1991.

[4]Freudenberg, J. S.; and D. P. Looze: Right Half Plane Poles and Zeros and Design Tradeoffs in Feedback System, *IEEE Transactions on Automatic Control*, vol. AC-30. no. *6,* June 1985.

[5]Goodwin, G. C.; S. F. Graebe; and M. E. Salgado: *Control System Design*, Prentice Hall, 2000.

[6]Skogestad, S. and I. Postlethwaite: *Multi-Variable Feedback Control*, John Wiley and Sons, 2004.
Bode, H.: *Network Analysis and Feedback Amplifier Design*, Van Nostrand, New York, 1945.

[7]Lebsock, K. L., Magnetic Desaturation of a Momentum Bias System, *Journal of Guidance and Control*, vol. 6, no. 6, Nov.-Dec. 1983

[8]Terasaki, R.M., Dual Reaction Wheel Control of Spacecraft Pointing, *Symposium of Attitude Stabilization and Control of Dual Spin Spacecraft*, SAMSO and Aerospace Corporation, El Segundo, Calif., Aug. 1967.

[9]Wie, B., *Space Vehicle Dynamics and Control*, ch 9.2, American Institute of Aeronautics and Astronautics, Inc., Reston, VA, 2008

[10]Curtis P. Mracek and D. Brett Ridgely, *Missile Longitudinal Autopilots: Comparison of Multiple Three Loop Topologies* AIAA paper 2005-6380 , 2005 AIAA Guidance, Navigation and Control Conference, San Francisco CA.

Appendix A - Bode and Poisson Integrals

I-A Conservation Laws in the Frequency Domain

SOME linear systems are inherently difficult to control. The greatest limitations on stability and disturbance rejection are related to the presence of poles and zeroes in the right half-plane (RHP) and time delays. Control loop performance, in terms of stability and disturbance rejection, is commonly analyzed in the frequency domain. Loop design can be viewed as tuning the sensitivity and complementary sensitivity transfer functions to achieve stability and disturbance rejection goals. Stability is related to the peaks of these transfer functions. Rolling off the transfer functions in certain frequency ranges determines disturbance rejection. The Bode and Poisson integrals of these transfer functions determine the limitations of loop shaping.

The integrals quantify the 'waterbed effect', i.e., improving the disturbance attenuation of controllers at some frequencies results in disturbance amplification at other frequencies. The Bode and Poisson integrals are conservation laws (Ref. 1). "The integrated value of the log of the magnitude of the sensitivity function is conserved under the action of feedback. The total amount of this quantity is always the same. It is equal to zero for stable plant/compensator pairs, and it is equal to some fixed positive amount for unstable ones." Similar comments apply to the complementary sensitivity function. The combination of RHP poles with non-minimum phase (NMP) zeroes, or time delays, increases the value of the integrals. These integral constraints impose fundamental limitations on what can be achieved with feedback control (Ref. 5).

II-A Transfer Functions and Stability Margins:

Consider the basic single-input single-output (SISO) feedback control system shown in the block diagram below. The transfer function of the controller is ***C(s)*** and that of the plant is ***G(s)***. The inputs to the system are the command or reference signal ***r***, the load disturbance ***d***, and the measurement noise ***n***. The process output is ***x***, the error signal to the controller is ***e***, and the controller output signal is ***u***.

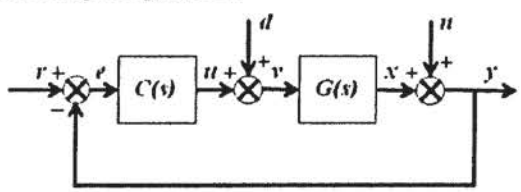


Figure 1. Single-input single-output (SISO) feedback control system.

Define the open-loop transfer ***L(s)***, the sensitivity function ***S(s)***, and the complementary sensitivity function ***T(s)*** as follows:

$$L(s) = G(s)\,C(s)$$

$$S(s) = \frac{1}{1+L(s)}$$

$$T(s) = \frac{L(s)}{1+L(s)}$$

The open-loop transfer ***L(s)*** can be written as:

$$L(s) = L_o(s)\,e^{-s\tau}$$

where τ is the time delay in the loop and $L_o(s)$ is a proper rational transfer function. The open-loop transfer function describes how the system output x would respond to an input r if the feedback loop were not closed. Bode or Nyquist plots of the open-loop transfer function are commonly used to determine the stability margins of the closed loop system.

The complementary sensitivity function $T(s)$ is the closed-loop transfer function. It describes how the system output x will respond to an input r when the feedback loop is closed. The desired form of closed-loop transfer function is with unity gain and zero phase shift at low frequencies so that the reference input r is tracked perfectly. The desired amplitude of the closed-loop transfer function rolls off at frequencies that are higher than the system bandwidth in order to filter out sensor noise. The Bode and Poisson integrals are used to study the constraints on selecting the system bandwidth.

The sensitivity function $S(j\omega)$ is simply the ratio of the output of the closed-loop system to that of the open-loop system. The sensitivity function tells how the variations in the output are influenced by feedback. At low frequency, a large magnitude of the open-loop transfer function causes the sensitivity to be small. This provides good load disturbance rejection and reference tracking. At high frequency, a small loop gain is used to avoid amplifying measurement noise. Load disturbances with frequencies such that $|S(j\omega)| < 1$ are attenuated, but disturbances at frequencies where $|S(j\omega)| > 1$ are amplified by feedback. The maximum sensitivity $Ms = \left\| S(j\omega) \right\|_\infty$ occurring at the frequency ω_{ms}, is thus a measure of the largest amplification of the low-frequency load disturbances.

The conservation theorems expressed with the Bode and Poisson integrals are useful for investigating the limitations on achievable closed-loop stability. The maximum magnitude of $S(j\omega)$ is also the minimum of $|1+L(j\omega)|$. An alternative way to express margins is by a single number, the stability margin s_m, which is the shortest distance from the Nyquist curve (shown below) to the critical point, i.e., $|1+L(j\omega)|$. Therefore the maximum sensitivity is also a measure of robustness since $Ms = 1/s_m$. The sensitivity peak is a more compact indicator of stability than a pair of gain and phase margins.

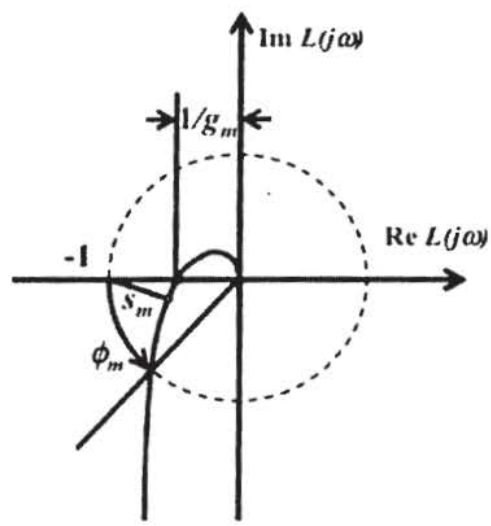


Figure 2. Nyquist plot.

The relationships between the maximum sensitivity and the lower bounds of the gain and phase margins are given by the following inequalities:

$$g_m \geq 20\mathrm{Log}_{10}\left[\frac{Ms}{Ms-1}\right]\text{(db)} \quad \text{and} \quad \phi_m \geq 2\mathrm{ArcSin}\left[\frac{1}{2Ms}\right] \geq \frac{1}{Ms}\text{(rad)}$$

Typical specifications for the peak sensitivity magnitude are in the range of 1.33 to 2 which corresponds approximately to gain and phase margins (12 dB and 45°) to (6 dB and 30°).

The complementary sensitivity function $T(j\omega)$ describes how closely the output x of the closed-loop system tracks the reference input r and the measurement noise n. At low frequency, a large magnitude of the open-loop transfer function, $L(j\omega)$, causes the complementary sensitivity to be nearly unity, thereby providing good command following. At high frequency the loop gain is decreased to roll off the complementary sensitivity so as to avoid amplifying high-frequency measurement noise. A requirement on the closed-loop bandwidth is often specified for the frequency that divides these two regimes. The closed-loop bandwidth is defined as the highest frequency at which $|T(j\omega)|$ crosses $1/\sqrt{2}$ (approximately −3dB) from above.

Note that since $S + T = 1$ at any frequency, it follows that S and T differ by no more than 1. A large value of Ms occurs if and only if $Mt = \left\| T(j\omega) \right\|_\infty$ also is large. Classical control requirements often specify an upper bound on the magnitude of the closed-loop resonant peak Mt for high frequency noise rejection. A typical requirement of $Mt <$ 1.5 also implies that $g_m > 4.4$db since

$$g_m \geq 20\mathrm{Log}_{10}\left[1+\frac{1}{Mt}\right](\mathrm{db})$$

In the sequel, all of the frequency-domain integral constraints are expressed as weighted areas under a logarithmic sensitivity or complementary sensitivity function. The weighting factors for the Bode integrals are unity. If both NMP zeroes and RHP poles are involved, then Poisson integrals are employed with weighting factors that are functions of the unstable roots and frequency. In all cases, the integral constraints can be visualized as rules requiring conservation of area as depicted in a plot of a logarithmic quantity vs. linear frequency. Note that this representation is not the same as the conventional Log-Log Bode magnitude plots where Log magnitude is plotted against Log frequency.

III-A Sensitivity Integral Limitations due to Pole-Zero Excess, Time Delay and RHP poles:

The Bode Sensitivity Integral is defined in terms of the natural logarithm of the sensitivity transfer function. The original formulation of the equation for stable systems is due to Bode in 1945 (Ref. 2). The integral constraint is called the waterbed effect because it implies that if the sensitivity is pushed down in a certain frequency range, it increases by an equal amount at other frequencies. The Bode integral is a conservation principle because the net area under the sensitivity function is conserved. If $L(s)$ is proper (i.e., number of poles is greater than the number of zeroes), has no unstable roots, and there is no time delay, then the value of the integral is zero.

$$\int_0^\infty \mathrm{Ln}\left|S(j\omega)\right| d\omega = 0 \quad \text{for} \quad n_{pz} \geq 1 \quad \& \quad n_{up} = 0 \quad \& \quad \tau = 0$$

where n_{pz} is the relative order or pole-zero excess (number of poles minus number of zeroes), n_{up} is the number of unstable poles in the RHP, τ is the time delay in the loop, and **Ln** denotes the natural logarithm. This result is independent of whether or not there are NMP zeroes.

However, if there are unstable poles, the value of the integral is no longer automatically equal to zero. Then the integral of the Log Magnitude of the sensitivity function $\mathrm{Ln}|S(j\omega)|$ is greater than zero by π times the sum of the unstable poles. Note that in the second equation below when the pole-zero excess is unity (the degree of the denominator is exactly one greater than the degree of the numerator) the integrated area penalty is decreased by $\kappa\frac{\pi}{2}$ where $\kappa = \lim_{s \to \infty}\left[s\,L(s)\right]$. The Bode integrals for unstable systems are:

$$\int_0^\infty \mathrm{Ln}\left|S(j\omega)\right| d\omega = \pi \sum_i^{n_{up}} p_i \quad \text{for} \quad \tau > 0 \quad \& \quad n_{pz} \geq 0$$

$$\int_0^\infty \mathrm{Ln}\left|S(j\omega)\right| d\omega = \pi \sum_i^{n_{up}} p_i - \kappa \frac{\pi}{2} \quad \text{for} \quad \tau = 0 \quad \& \quad n_{pz} = 1$$

$$\int_0^\infty \mathrm{Ln}\left|S(j\omega)\right| d\omega = \pi \sum_i^{n_{up}} p_i \quad \text{for} \quad \tau = 0 \quad \& \quad n_{pz} > 1$$

Observations:

- Decreasing $|S(j\omega)|$ at some frequencies will increase it at other frequencies. Hence, the area under the Log Magnitude curve is conserved.
- The total amount of area under the $\mathrm{Ln}|S(j\omega)|$ vs. ω curve is increased if the open-loop system $L(s)$ is unstable. This places a lower bound on the bandwidth.
- The closed-loop bandwidth should be set at least twice as large as the real part of any unstable pole.
- When $n_{pz} = 1$, the total amount of area under the curve is decreased by increasing the high-frequency gain (which increases the value of κ). However, the peak magnitude of the sensitivity $|S(j\omega)|$, which is inversely related to robustness, may also be increased.
- $L(s)$ has relative degree $n_{pz} > 1$ if both the plant $G(s)$ and the controller $C(s)$ are strictly proper.
- Note that $\kappa = 0$ if $n_{pz} > 1$.
- There is no bound on the integral.of the logarithm of sensitivity when $\tau = 0$ and $n_{pz} = 0$ (i.e., when there is no time delay and the open-loop transfer function is improper).

The extension of the original Bode integral to open-loop unstable systems is due to Freudenberg in 1985 (Ref. 4). The presence of unstable poles modifies the waterbed effect. Usually this means that the peak of the sensitivity is larger for unstable systems than for stable systems. Note that in any frequency band where the sensitivity $|S(j\omega)|>1$, the closed-loop system performs worse than the open-loop system.

IV-A Additional Sensitivity Integral Limitations due to NMP Zeroes:

For open-loop transfer functions with non-minimum phase zeroes the sensitivity function must also satisfy a second type of waterbed formulae. These are Poisson integral constraints and they can have a very strong influence on the peak sensitivity. The Poisson integrals include the effects of n_{uz} NMP zeroes in addition to the n_{up} unstable poles $p_1 \ldots\ldots p_{n_{up}}$ of the open-loop transfer function $L(s)$. These integrals must be evaluated individually for each NMP zero. Each integrand is multiplied by a weighting factor that is a function of the real and imaginary parts of the NMP zero. The value of the resulting integral can be expressed as a Blaschke product of the NMP zero and all of the unstable poles. For every NMP zero, $z_k = \gamma_k + j\delta_k$ (with $\gamma_k > 0$), the sensitivity must satisfy:

$$\int_0^\infty \mathrm{Ln}\left|S(j\omega)\right| \left(\frac{\gamma_k}{\gamma_k^2 + (\delta_k - \omega)^2} + \frac{\gamma_k}{\gamma_k^2 + (\delta_k + \omega)^2} \right) d\omega = \pi\, \mathrm{Ln} \left| \prod_{i=1}^{n_{up}} \frac{z_k + p_i^*}{z_k - p_i} \right| \quad \text{for} \quad k = 1, \ldots\ldots n_{uz}$$

The notation p_i^* denotes complex conjugate of p_i. For a real NMP-zero $(\gamma_k = z_k, \delta_k = 0)$, this integral above simplifies to:

$$\int_0^{\infty} \mathrm{Ln}\left|S(j\omega)\right| \frac{2\,z_k}{z_k^2+\omega^2}\,d\omega = \pi\,\mathrm{Ln}\left|\prod_{i=1}^{n_{up}} \frac{z_k+p_i^*}{z_k-p_i}\right| \quad \text{for} \quad k=1,\ldots\ldots n_{uz}$$

Observations:

- For stable systems $(n_{up}=0)$, the right-hand side is zero so the areas of sensitivity increase and decrease are evenly balanced in a waterbed effect.
- The weighting function in the integrand is positive and the product on the right-hand side is greater than or equal to one. Consequently, the weighted area integral of the sensitivity function will be positive if there are both NMP zeroes and RHP poles.
- The excess positive area may appear in either a large sensitivity peak (decreased stability margin) or sensitivity that is greater than unity over a broader frequency range (decreased closed-loop performance) or both.
- The right-hand side of the equation becomes very large when there is an unstable pole close to the unstable zero. This condition makes the system very hard to control.
- The closed-loop bandwidth should not exceed the magnitude of the smallest NMP open-loop zero. Otherwise a very large sensitivity peak will occur, leading to fragile loops (non-robust) and large undershoots and overshoots.
- The maximum sensitivity peak should be limited in order to reduce disturbance amplification and maintain a satisfactory stability margin.

V-A Complementary Sensitivity Integral Limitations due to NMP Zeroes and Time Delays:

In contrast with the area conservation integrals for the sensitivity function, the Bode and Poisson integral constraints for the complementary sensitivity transfer function are influenced by time delay. Recall that the open-loop transfer $L(s)$ can be written as:

$$L(s) = L_o(s)\,e^{-s\tau}$$

where τ is the time delay in the loop and $L_o(s)$ is a proper rational transfer function. The integral constraints for the complementary sensitivity transfer function are also influenced by NMP zeroes whether or not there are unstable poles. Assume that $L(s)$ has $n_{uz} \geq 0$ non-minimum phase zeroes (counting multiplicity) $z_1 \ldots\ldots z_{n_{uz}}$.

Then, the sensitivity function $T(j\omega)$ satisfies

$$\int_0^{\infty} \frac{1}{\omega^2}\mathrm{Ln}\left|T(j\omega)\right| d\omega = \pi \sum_i^{n_{uz}} \frac{1}{z_i} + \frac{\pi\tau}{2} \qquad n_{po} \geq 2$$

$$\int_0^{\infty} \frac{1}{\omega^2}\mathrm{Ln}\left|T(j\omega)\right| d\omega = \pi \sum_i^{n_{uz}} \frac{1}{z_i} + \frac{\pi\tau}{2} - \frac{\pi}{2k_v} \qquad n_{po} = 1$$

$$\int_0^{\infty} \frac{1}{\omega^2}\mathrm{Ln}\left|T(j\omega)\right| d\omega = \pi \sum_i^{n_{uz}} \frac{1}{z_i} + \frac{\pi\tau}{2} \qquad n_{po} = 0$$

where the velocity constant for the loop is defined as:

$$k_v = -\left(\lim_{s \to 0} \frac{dT(s)}{ds} \right)^{-1} = + \lim_{s \to 0} s\,L(s)$$

Observations:

- Decreasing $|T(j\omega)|$ at some frequencies, ω, will increase it at other frequencies. The area under the $\mathrm{Ln}\,|T(j\omega)|/\omega^2$ vs. ω curve is conserved.
- The total amount of area under the curve is increased if the open-loop system $L(s)$ has non-minimum phase zeroes or time delays, since the right-hand side of the equations above is increased.
- Slow unstable zeroes and long time delays limit the upper bound on bandwidth. The closed-loop bandwidth should be half the size of the slowest NMP zero.
- If $L(s)$ does not have free integrators (i.e. $T(0) - 1$), then a similar relation exists, except that the integrand is now $\frac{1}{\omega^2}\mathrm{Ln}\left|\frac{T(j\omega)}{T(0)}\right|$.
- If $L(s)$ has at least one pole at the origin of the complex plane, $n_{po} > 0$, then $1/L(0) = 0$ and $T(0) = 1$.
- If $L(s)$ has only one free integrator, one can decrease the total area under the $\frac{1}{\omega^2}\mathrm{Ln}\left|T(j\omega)\right|$ curve by tolerating steady-state error due to a ramp input.
- If $L(s)$ has at least $n_{po} \geq 2$ free integrators (poles at 0), then the velocity constant is $k_v = \infty$. This ensures that the steady-state error is 0 for a ramp input.

VI-A Complementary Sensitivity Integral Limitations due to Unstable Poles in Addition to Time Delay and NMP Zeroes:

The constraints on complementary sensitivity due to both non-minimum-phase zeroes and unstable right-half-plane poles are expressed as Poisson integrals. This case also includes the effect of a feedback loop having a time delay $\tau \geq 0$ in addition to the n_{up} open-loop RHP poles located at $p_1 \ldots\ldots p_{n_{up}}$ where $p_i = \alpha_i + j\,\beta_i$ and the n_{uz} open-loop zeroes are located at $z_1 \ldots\ldots z_{n_{uz}}$.

For each NMP zero the complementary sensitivity satisfies:

$$\int_0^\infty \mathrm{Ln}\left|T_o(j\omega)\right| \left(\frac{\alpha_i}{\alpha_i^2 + (\beta_i - \omega)^2} + \frac{\alpha_i}{\alpha_i^2 + (\beta_i + \omega)^2} \right) d\omega = \pi\,\mathrm{Ln}\left| \prod_{k=1}^{n_{uz}} \frac{p_i + z_k^*}{p_i - z_k} \right| + \pi\tau\alpha_i \quad \text{for} \quad i = 1, \ldots\ldots n_{up}$$

For a real RHP pole ($p_i = \alpha_i$, $\beta_i = 0$) this equation simplifies to:

$$\int_0^\infty \mathrm{Ln}\left|T_o(j\omega)\right| \frac{2\alpha_i}{\alpha_i^2 + \omega^2}\, d\omega = \pi\,\mathrm{Ln}\left| \prod_{k=1}^{n_{uz}} \frac{p_i + z_k^*}{p_i - z_k} \right| + \pi\tau\alpha_i \quad \text{for} \quad i = 1, \ldots\ldots n_{up}$$

Observations:

- The weighting function in the integrand is positive and the product on the right-hand side is greater than or equal to one. Consequently, the weighted area integral of the sensitivity function will be positive if there are both NMP zeroes and RHP poles.
- The right-hand side of the equation becomes very large when there is an unstable pole close to the unstable zero. This will result in poor noise rejection and/or poor reference command tracking.
- A large peak in the complementary sensitivity will result if the closed-loop bandwidth is smaller than the magnitude of a RHP pole.
- Time delays place an additional limit on the achievable bandwidth. The effect of time delays is more pronounced for fast unstable poles (i.e., large α_i).
- The closed-loop bandwidth should be less than the slowest NMP zero and larger than the fastest RHP pole. Skogestad and Postlethwaite give detailed guidelines for complex roots in their book (Ref. 6).
- In the case of real RHP poles $p_1 < p_1 < \ldots\ldots < p_{n_{up}}$, real NMP zeroes $z_1 < z_2 < \ldots\ldots < z_{n_{uz}}$, and with a time delay τ, the 3dB bandwidth, ω_{3dB}, of the closed-loop system should be:

$$2p_{n_{up}} < \omega_{3dB} < \frac{z_1}{2} \qquad \text{and} \qquad \omega_{3dB} < \frac{1}{\tau}$$