Closed-Loop Convex Formulation of Classical and Singular Value Loop Shaping

Craig Barratt Stephen Boyd*

Department of Electrical Engineering
Stanford University
Stanford CA 94305

September 1991

Abstract

We show that control system design via *classical loop shaping* and *singular value loop shaping* can be formulated as a *closed-loop convex* problem [4, 5, 22, 15]. Consequently, loop shaping problems can be solved by efficient numerical methods. In particular, these numerical methods can always determine whether or not there exists a compensator that satisfies a given set of loop shaping specifications. Problems such as maximizing bandwidth subject to given margin and cutoff specifications can be directly solved. Moreover, any other closed-loop convex specifications, such as limits on step-response overshoot, tracking errors, and disturbance rejection, can be simultaneously considered.

These observations have two practical ramifications. First, closed-loop convex design methods can be used to synthesize compensators in a framework that is familiar to many control engineers. Second, closed-loop convex design methods can be used to aid the designer using classical loop shaping by computing absolute performance limits against which a classical design can be compared.

To appear as a chapter in Advances in Control Systems, edited by C. T. Leondes, 1993.

*Research supported in part by NSF under ECS-85-52465 and AFOSR under 89-0228.

1 Introduction

We give a brief overview of classical and singular value loop shaping, which also serves to describe our notation.

1.1 Classical loop shaping

We first consider the standard classical one degree-of-freedom single-actuator, single-sensor (SASS) control system shown in figure 1. Here u is the actuator signal, y is the output signal, e is the (tracking) error signal, r is the reference or command signal, and d_{sensor} is a sensor noise. The plant and compensator are linear and time-invariant (LTI), with transfer functions given by P and C, respectively. The plant is given and the compensator is to be designed.

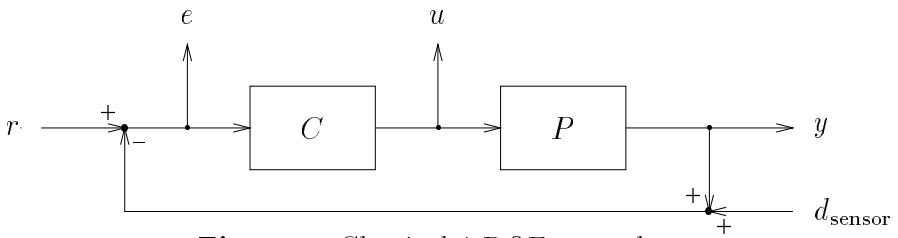


Figure 1 Classical 1-DOF control system.

In classical loop shaping, the designer focuses attention on the loop transfer function, given by

$$L \triangleq PC.$$

Many important aspects of closed-loop control system performance can be expressed in terms of L. For example, stability of the closed-loop system can be determined from L (provided there are no unstable pole-zero cancellations between P and C). Several important closed-loop transfer functions can be expressed in terms of L.

The *transmission* or input/output (I/O) transfer function

$$T \triangleq L/(1+L)$$

is the closed-loop transfer function from the reference input r to the output y. Its negative, $-T$, is the closed-loop transfer function from the sensor noise d_{sensor} to the output y.

The *sensitivity* transfer function is given by

$$S \triangleq 1/(1+L).$$

S is the transfer function from the reference input r to the tracking error e. The sensitivity derives its name from the important fact, observed by Bode [2], that to first order, the relative change in T is S times the relative change in P:

$$\frac{\delta T(s)}{T(s)} \simeq S(s)\frac{\delta P_0(s)}{P_0(s)},$$

or, equivalently, S is the first order percentage change in the I/O transfer function divided by the percentage change in the plant transfer function.

Classical loop shaping design is based on two important observations:

- the loop transfer function L has a very simple dependence on the compensator transfer function C, especially in a logarithmic (gain and phase) representation.
- many important requirements for the closed-loop system can be approximately reflected as requirements on the loop gain L.

Loop-shaping specifications constrain the magnitude and possibly the phase of the loop transfer function at each frequency. There are three basic types of loop-shaping specifications, which are imposed in different frequency bands:

- *In-band specifications.* At these frequencies we require $|L|$ to be large, so that S is small and $T \approx 1$. This ensures good command tracking, and low sensitivity to plant variations, two of the most important benefits of feedback.
- *Cutoff specifications.* At these frequencies we require $|L|$ to be small, so that T is small. This ensures that the output y will be relatively insensitive to the sensor noise d_{sensor}, and that the system will remain closed-loop stable in the face of plant variations at these frequencies, for example, excess phase from small delays and unpredictable (or unmodeled) resonances.
- *Crossover (margin) specifications.* Crossover or transition band specifications are imposed between the control bands (where L is large) and cutoff bands (where L is small). At these frequencies the main concern is to keep L a safe distance away from the critical point -1 (closed-loop stability depends on the winding number of L with respect to -1). Classical specifications include gain margin and phase margin. More natural "modern" specifications exclude L from some circle about -1. These modern specifications directly correspond to limiting the peaking of some closed-loop transfer function such as S or T.

The Nyquist criterion (which constrains the winding number of L about -1) is also included as an implicit specification that ensures closed-loop stability.

In many systems the in-band region is at low frequencies, from $\omega = 0$ to $\omega = \omega_B$, the cutoff region is at high frequencies, $\omega > \omega_C$, and the crossover region lies in

between, from $\omega = \omega_B$ to $\omega = \omega_C$. In some designs, however, there may be more than one crossover region and one or more in-band and cutoff regions.

A typical set of loop shaping specifications is:

$$\begin{array}{ll} |L(j\omega)| \geq l(\omega) & \text{for } \ 0 \leq \omega \leq \omega_B = 2, \\ |L(j\omega)| \leq u(\omega) & \text{for } \ \omega \geq \omega_C = 5, \\ -150^\circ \leq \angle L(j\omega) \leq 30^\circ & \text{for } \ \omega_B = 2 \leq \omega \leq \omega_C = 5 \end{array}$$

where l and u are the frequency dependent constraint functions shown in figure 2. The in-band and cutoff constraints, which consist of frequency dependent restrictions on the magnitude of L, are conveniently shown on a Bode magnitude plot, while the margin constraint, which is often independent of frequency, is conveniently shown on a Nyquist plot (see figure 3).

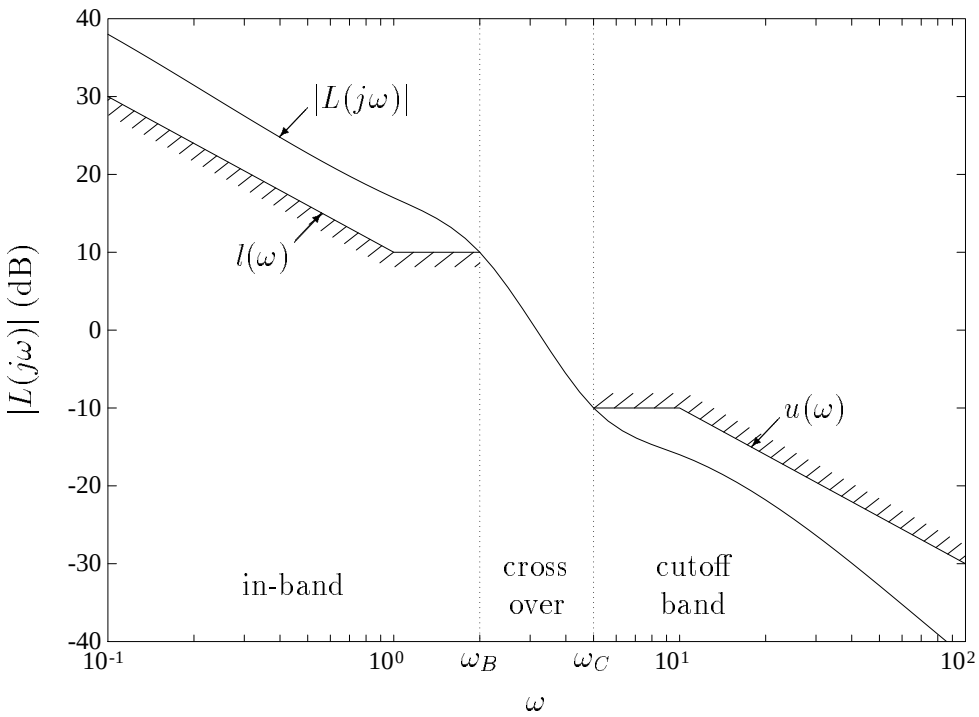


Figure 2 A typical set of in-band and cutoff specifications. In the in-band region, $\omega \leq \omega_B$, the loop gain magnitude $|L|$ is required to exceed the frequency dependent lower bound $l(\omega)$. In the cutoff region, $\omega \geq \omega_C$, the loop gain magnitude $|L|$ is required to be below the upper bound $u(\omega)$. In the crossover region, $\omega_B < \omega < \omega_C$, the loop gain crosses $|L| = 0$dB.

In this example, the in-band region is $\omega \leq \omega_B$. Over this region, the large loop gain will ensure good command tracking ($T \approx 1$), and low sensitivity ($|S| \ll 1$). In the cutoff region, $\omega \geq \omega_C$, the small loop gain ensures that sensor noise will not affect the output, and small time-delays and variations in P will not destabilize the

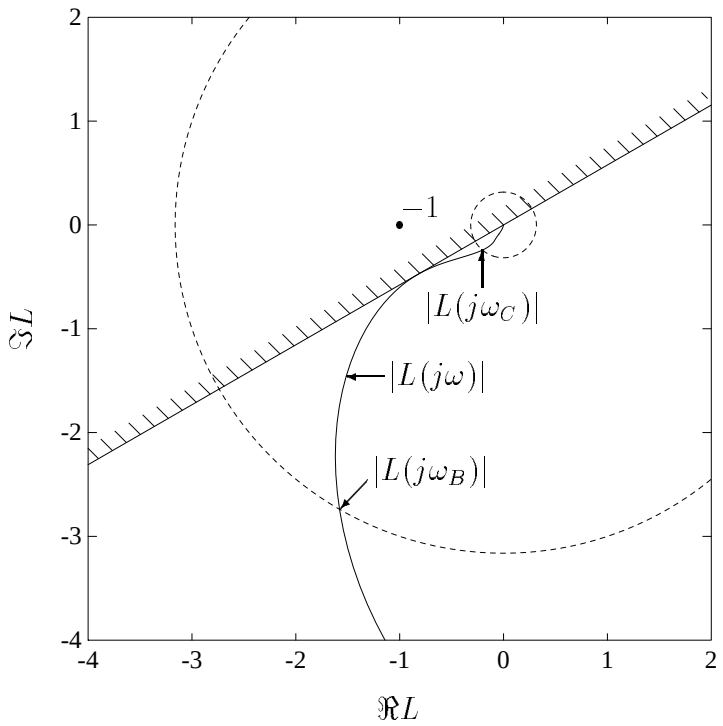


Figure 3 A typical margin specification requires the phase of the loop transfer function to lie between $-150°$ and $+30°$ over the frequency band $\omega_B \leq \omega \leq \omega_C$. The dotted circles shown correspond to the magnitude constraints $|L(j\omega)| > +10\text{dB}$ and $|L(j\omega)| < -10\text{dB}$ which must be satisfied for $\omega \leq \omega_B$ and $\omega \geq \omega_C$, respectively.

closed-loop system. In the in-band region, L cannot be close to the critical point -1 since $|L|$ exceeds $+10$dB there; similarly, in the cutoff region, $|L|$ is less than -10dB and so cannot be close to -1. The margin specification ensures that L cannot be too close to -1 in the transition region $\omega_B \leq \omega \leq \omega_C$ by constraining $\angle L$. Of course, the phase bounds in the margin constraint can be frequency dependent.

While many important closed-loop properties can be specified via L, some cannot. For example, loop shaping does not explicitly include specifications on $C/(1+PC)$ (actuator effort) and $P/(1+PC)$ (effect of input-referred process noise on y). A design will clearly be unsatisfactory if either of these transfer functions is too large. The specification that these transfer functions should not be too large is usually included as implicit "side information" in a classical loop shape design. Specifications that limit the size of these transfer functions are closed-loop convex, however, and so are readily incorporated in a closed-loop convex formulation.

Given a desired set of loop shaping specifications, the compensator C is typically synthesized by adding dynamics until the various requirements on the loop transfer function L are satisfied (or until the designer suspects that the loop shaping specifi-

cations cannot be met).

Classical loop shaping is described in many texts; see, for example, [2, 16, 21, 8, 17, 13]. The discussions found in these references emphasize techniques that help the engineer "do" loop-shaping design. With the exception of Bode's work on optimal cutoff characteristics and integral constraints, these references do not consider questions such as:

- Is there a compensator that meets a given set of loop-shaping specifications?
- For a given set of in-band and margin specifications and shape of the cutoff specification, what is the smallest cutoff frequency that can be achieved?
- For a given set of cutoff and margin specifications, how large can the loop gain be made in the in-band region?

The main point of this paper is that such questions are readily answered.

1.2 Singular value loop shaping

We now consider the case in which there are multiple actuators and multiple sensors (MAMS) in the control system shown in figure 1. The plant P and compensator C are given by transfer matrices: P is n_{sens} by n_{act} and C is n_{act} by n_{sens}, where n_{sens} is the number of sensors, and n_{act} is the number of actuators.

Unlike the SASS case, there is no longer a unique choice for the "loop transfer function." A common choice is the loop transfer matrix cut at the sensors:

$$L \triangleq PC.$$

The transmission or input/output (I/O) transfer matrix is

$$T \triangleq (I + L)^{-1} L,$$

and the sensitivity transfer matrix is given by

$$S \triangleq (I + L)^{-1}.$$

These transfer matrices have interpretations that are are similar to those in SASS case. For example, if the plant transfer matrix P changes to $(I + \Delta)P$, then the I/O transfer matrix T, to first order, changes to $(I + S\Delta)T$. (Note that Δ can be interpreted as the output-referred the output-referred fractional change in the I/O transfer matrix T is then given b $S\Delta$ [6, 4].)

In contrast, the loop transfer matrix cut at the actuators is denoted $\tilde{L}$, the complementary loop transfer matrix:

$$\tilde{L} \triangleq CP.$$

Note that the loop transfer matrix and the complementary loop transfer matrix may have different dimensions: L is n_{sens} by n_{sens}, while $\tilde{L}$ is n_{act} by n_{act}. Moreover, loop specifications on L and $\tilde{L}$ are in general different and inequivalent. For example, it is possible for L to be "large" (in the sense to be described below), while $\tilde{L}$ is not "large."

A second difficulty with the extension of SASS loop shaping is choosing a measure for the "size" of the loop transfer matrix. Provided the individual sensor signals are scaled appropriately, a natural (and widely used) measure of the size is based on the singular values of the loop transfer matrix. (The singular values of a matrix M are the square roots of the eigenvalues of the Hermitian matrix M^*M.) Specifically, if all the singular values of the loop transfer matrix are large, then the loop transfer matrix is "large in all directions," and it follows that the sensitivity transfer matrix S is small and $T \approx I$. Similarly, if all the singular values of the loop transfer matrix are small, then the loop transfer matrix is "small in all directions," and it follows that T is small and $S \approx I$. These important ideas are discussed in, for example, [11, 6, 14, 19, 18, 4].

At in-band frequencies, singular value loop shaping specifications have the form

$$\sigma_{\min}(L(j\omega)) \geq l(\omega) > 1,$$

where l is some frequency dependent bound. For cutoff frequencies, singular value loop shaping specifications have the form

$$\sigma_{\max}(L(j\omega)) \leq u(\omega) < 1,$$

where u is some frequency dependent bound.

These specifications are often depicted on a singular value Bode plot, as in figure 4.

(This discussion assumes that there are at least as many actuators as actuators, *i.e.*, $n_{\text{act}} \geq n_{\text{sens}}$. If not, the in-band specifications above are guaranteed to be infeasible since at all frequencies at least one singular value of L is zero. In this case, similar specifications can be imposed on $\tilde{L}$.)

It is difficult to formulate margin specifications that are directly analogous to the gain or phase margin constraints used in the SASS case. The general idea is to ensure that $L + I$ stays "sufficiently invertible" in the crossover band. One effective method simply limits the minimum singular value of this matrix:

$$\sigma_{\min}(L + I) \geq r > 0,$$

or equivalently,

$$\sigma_{\max}(S) \leq 1/r.$$

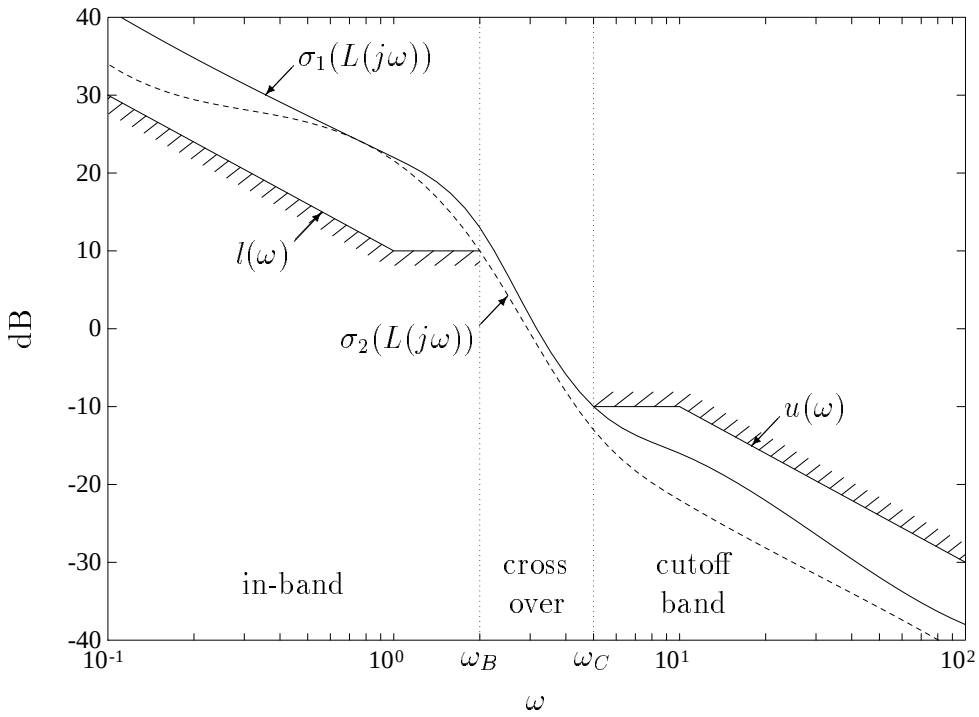


Figure 4 Examples of in-band and cutoff specifications on the loop gain L in a system with two sensors. In the in-band region, $\omega \leq \omega_B$, both singular values of L are required to exceed the lower bound. In the cutoff region, $\omega \geq \omega_C$, both singular values of L are required to be below the upper bound. In the crossover region each singular value of L crosses 0dB.

2 A Closed-Loop Convex Formulation

A design specification is *closed-loop convex* if it is equivalent to some closed-loop transfer function or matrix (*e.g.*, the sensitivity S) belonging to a convex set.

As a specific example, consider the specification

$$|L(j\omega)| \geq 3 \quad \text{for } 0 \leq \omega \leq 1. \tag{1}$$

We will see that this is equivalent to

$$|S(j\omega) - 1/8| \leq 3/8 \quad \text{for } 0 \leq \omega \leq 1. \tag{2}$$

Now, the set of transfer functions S that satisfy (2) is convex, since if $S^{(\mathrm{a})}$ and $S^{(\mathrm{b})}$ both satisfy (2), then so does $(S^{(\mathrm{a})} + S^{(\mathrm{b})})/2$. Therefore the specification (1) is closed-loop convex. See [4, 5] for extensive discussions.

The main result of this paper is that many classical and singular value loop-shaping specifications are closed-loop convex.

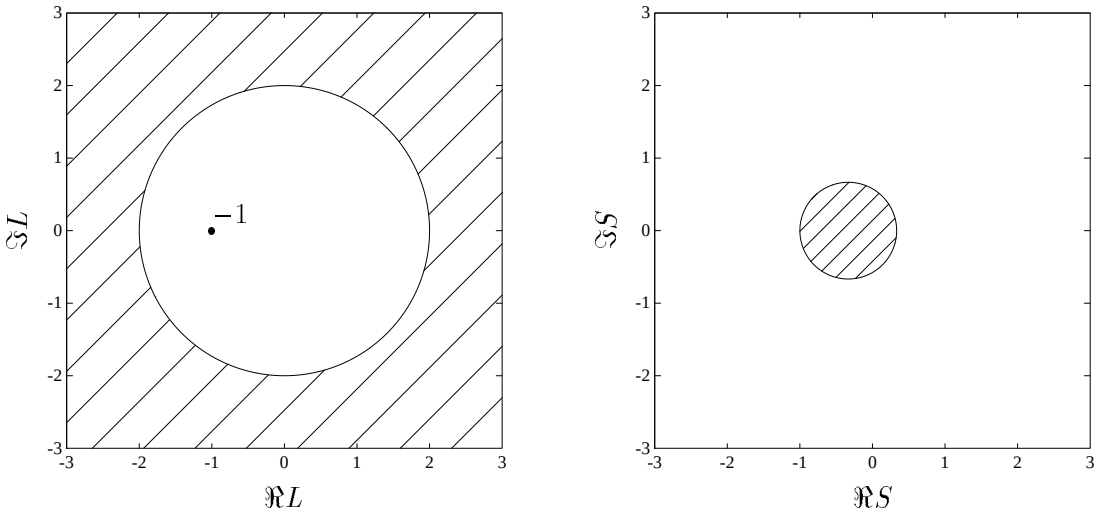
2.1 SASS case

2.1.1 In-band specifications

We first consider the in-band specification $|L(j\omega)| \geq \alpha$, where $\alpha > 1$. It is closed-loop convex since it is equivalent to the following convex specification on the sensitivity S, (a closed-loop transfer function):

$$|L(j\omega)| \geq \alpha > 1 \iff \left|S(j\omega) - \frac{1}{\alpha^2 - 1}\right| \leq \frac{\alpha}{\alpha^2 - 1}. \tag{3}$$

In other words, requiring $|L| \geq \alpha > 1$ is equivalent to requiring the sensitivity to lie inside a circle centered at $1/(\alpha^2 - 1)$ with radius $\alpha/(\alpha^2 - 1)$. Note that requiring L to be large corresponds to restricting the sensitivity S to lie in a disk that includes the point 0, but is not exactly centered at 0. Figure 5 illustrates this correspondence for $\alpha = 2$.



(a) (b)

Figure 5 The region $|L| \geq 2$ in the L-plane is shown in (a). The loop shaping specification $|L| \geq 2$ requires the Nyquist plot of L to lie in the shaded region in (a). The corresponding region in the S-plane is shown in (b), which is a disk that includes but is not centered at 0. This region is convex, and hence the loop gain specification $|L| \geq 2$ is closed-loop convex.

2.1.2 Cutoff specifications

We now consider the cutoff specification $|L(j\omega)| \leq \alpha$, where $\alpha < 1$. It is also closed-loop convex since it is equivalent to the following convex specification on the sensitivity:

$$|L(j\omega)| \leq \alpha < 1 \iff \left|S(j\omega) - \frac{1}{1 - \alpha^2}\right| \leq \frac{\alpha}{1 - \alpha^2}. \tag{4}$$

In other words, requiring $|L| \leq \alpha < 1$ is equivalent to requiring the sensitivity to lie in a disk centered at $1/(1-\alpha^2)$ with radius $\alpha/(1-\alpha^2)$. Note that requiring L to be small corresponds to restricting the sensitivity S to lie in a disk that includes the point 1, but is not exactly centered at 1. Figure 6 illustrates this correspondence for $\alpha = 0.5$.

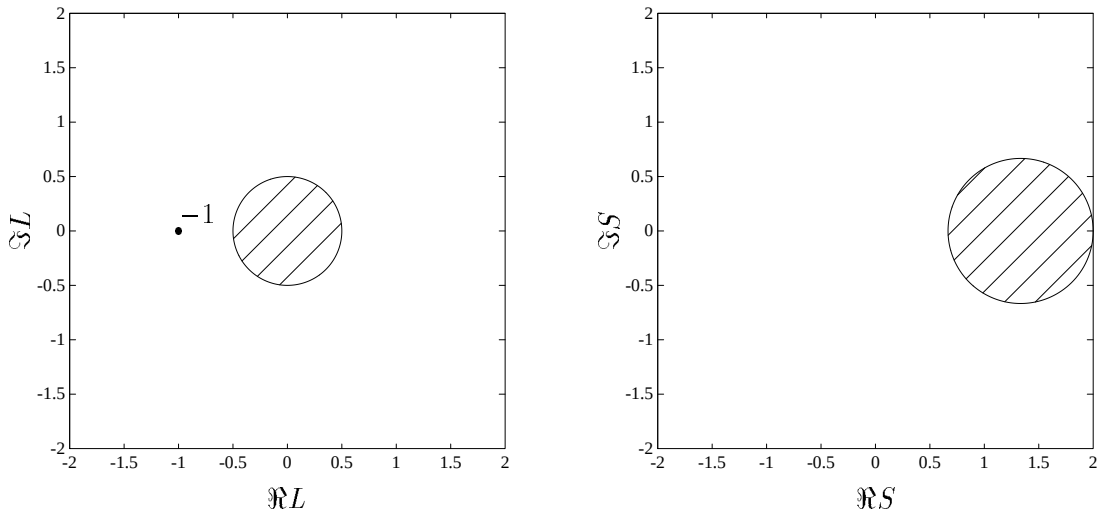


Figure 6 The region $|L| \leq 0.5$ in the L-plane is shown in (a). The loop shaping specification $|L| \leq 0.5$ requires the Nyquist plot of L to lie in the shaded region in (a). The corresponding region in the S-plane is shown in (b), which is a disk that includes but is not centered at 1. This region is convex, and hence the loop gain specification $|L| \leq 0.5$ is closed-loop convex.

The results (3) and (4) are easily established. Since we give a careful proof for the more general MAMS case, we give a simple discussion here. The in-band loop specification (3) requires L to lie outside a circle of radius α in the complex plane. Since $\alpha > 1$, the critical point -1 lies in the interior of this circle. Since S and L are related by the bilinear transformation $S = 1/(1+L)$, this circle maps to another circle in the S-plane. To find this circle, we note that the points $L = \pm\alpha$ map to $S = 1/(1 \pm \alpha)$, and the circle must be symmetric with respect to the real axis. Moreover since the critical point -1 is mapped to $S = \infty$, the exterior of the $|L| = \alpha$ circle maps to the interior of the circle in the S-plane.

The argument in the case of the cutoff specification (4) is similar, except that the critical point -1 is outside the $|L| = \alpha$ circle and so its interior maps to the interior of the corresponding circle in the S-plane.

We note that the specifications requiring L to be "not too big,"

$$|L(j\omega)| \leq \alpha \quad \text{where } \alpha > 1,$$

and requiring L to be "not too small,"

$$|L(j\omega)| \geq \alpha \quad \text{where } \alpha < 1,$$

are not closed-loop convex, since these specifications are equivalent to $S(j\omega)$ lying outside of the shaded disks in figures 5(b) and 6(b). These specifications, however, are not likely to be used in a practical design. It is interesting that the sensible specifications on $|L|$, given in (3) and (4), turn out to coincide exactly with the specifications on $|L|$ that are closed-loop convex.

2.1.3 Phase margin specifications

A common form for a margin specification limits the phase of the loop transfer function in the crossover band:

$$\theta_{\min} \leq \angle L(j\omega) \leq \theta_{\max}.$$

where $-180° < \theta_{\min} < 0°$ and $0° < \theta_{\max} < 180°$. It turns out that such a specification is closed-loop convex if and only if $\theta_{\max} - \theta_{\min} \leq 180°$, in which case S must lie in the intersection of two disks:

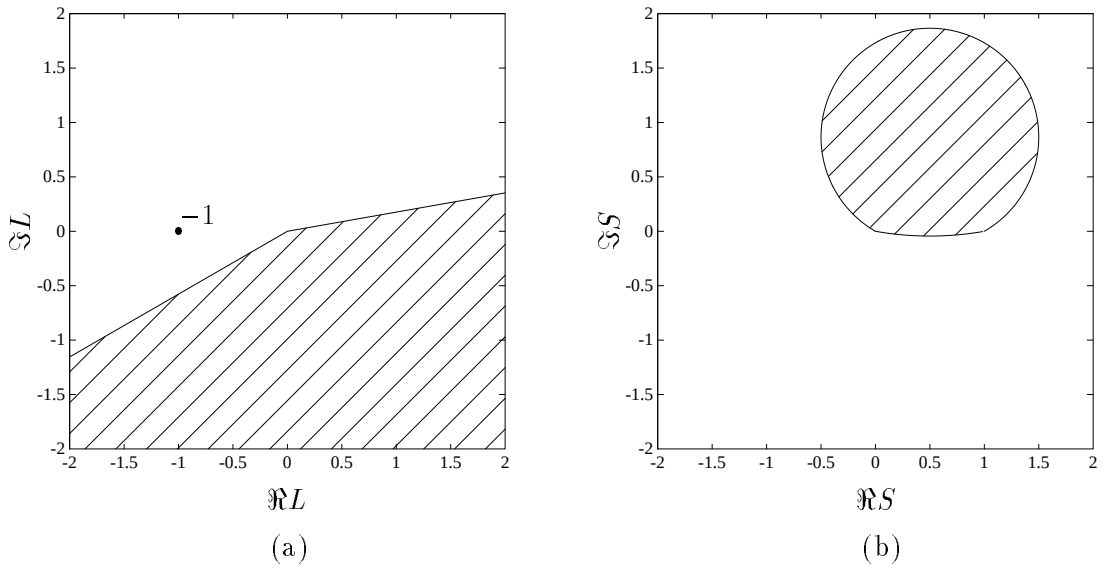
$$\theta_{\min} \leq \angle L(j\omega) \leq \theta_{\max} \iff |2S(j\omega) - (1 + j/\tan\theta_{\max})| \leq 1/\sin\theta_{\max} \text{ and} \quad (5)$$
$$|2S(j\omega) - (1 + j/\tan\theta_{\min})| \leq 1/\sin -\theta_{\min}. \quad (6)$$

This is shown in figure 7 for the case $\theta_{\min} = -150°$, $\theta_{\max} = 10°$. The phase margin specification $-150° \leq \angle L \leq 10°$ is equivalent to requiring the sensitivity S to lie in the convex set shown in figure 7(b).

2.1.4 General circle specifications

All of the specifications above—in-band, cutoff, and phase margin, are special cases of *general circle specifications.* Consider any generalized circle in the complex plane (*i.e.*, a circle or a line, which we consider a "circle" centered at ∞) that does not pass through the critical point -1. Such a circle divides the complex plane into two regions, one of which includes the critical point -1. The specification that the loop transfer function must lie in the region that does not contain -1 is what we call a *generalized circle constraint*, and is readily shown (by a mapping argument) to be closed-loop convex since it is equivalent to S lying inside a circle or half-plane.

The in-band and cutoff specifications are of this form with the circle given by $|L| = \alpha$; in each case the specification requires that the loop transfer function avoid the region that includes the critical point -1. The phase margin constraint can be expressed as the simultaneous satisfaction of the two generalized circle constraints corresponding to the lines that pass through the origin at the angles $\theta_{\min}$ and $\theta_{\max}$,



(a) (b)

Figure 7 The region $-150^\circ \leq \angle L \leq 10^\circ$ in the L-plane is shown in (a). The phase margin specification $-150^\circ \leq \angle L \leq 10^\circ$ requires the Nyquist plot of L to lie in the shaded region in (a). The corresponding region in the S-plane is the intersection of two disks, shown in (b), and hence the phase margin specification $-150^\circ \leq \angle L \leq 10^\circ$ is closed-loop convex.

respectively. This explains why the phase margin constraint is equivalent to the sensitivity lying inside the intersection of two disks (see (6)).

We note that generalized circle constraints have appeared in many contexts. Examples include the circle criterion, used in stability and robustness analysis of nonlinear systems (see [30, 20, 27, 24]), the Popov criterion (with a fixed Popov parameter) [23], and many of the specifications in [7, 29, 1]. In the remainder of this section we discuss two particular generalized circle constraints.

One useful generalized circle constraint excludes L from a disk about the critical point -1:

$$|L(j\omega) + 1| \geq \alpha, \tag{7}$$

where $\alpha > 0$. This specification is equivalent to

$$|S(j\omega)| \leq 1/\alpha,$$

which is just a limit on the magnitude of the sensitivity. The case $\alpha = 1$ is shown in figure 8.

Specifications of the form (7) can be used to guarantee a classical phase margin. Since the bounds $l(\omega)$ and $u(\omega)$ are not equal to one in the in-band and cutoff regions, $|L(j\omega)|$ can equal one only in the transition regions. If the specification (7) is imposed at all frequencies in the transition regions, then whenever $|L(j\omega)| = 1$, we have

$|L(j\omega)+1| \geq \alpha$, which implies a phase margin of at least $2\arcsin(\alpha/2)$ (and 180° for $\alpha > 2$).

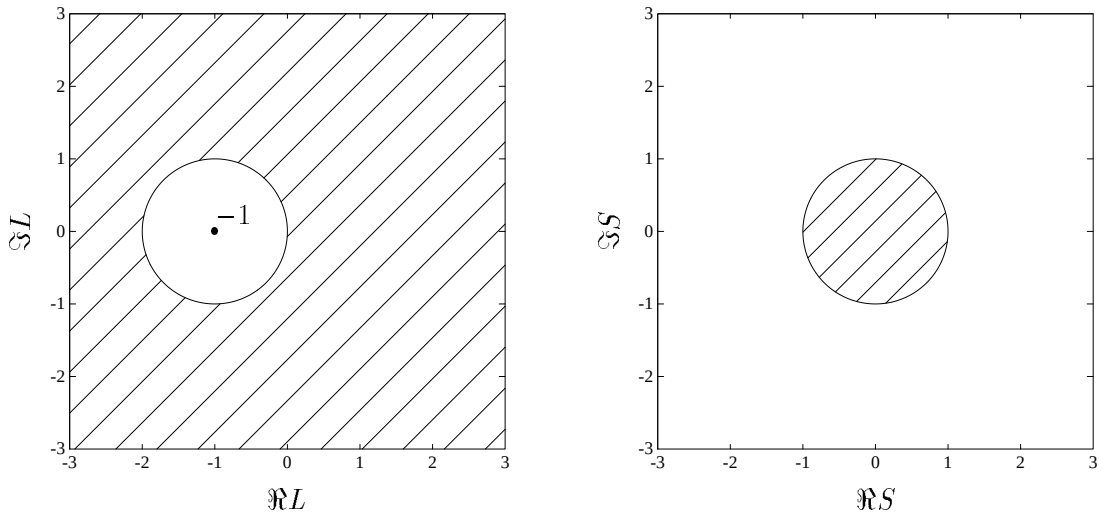


Figure 8 The region $|L+1| \geq 1$ in the L-plane is shown in (a). The specification $|L+1| \geq 1$ requires the Nyquist plot of L to lie in the shaded region in (a), *i.e.*, to maintain a distance of at least 1 from the critical point -1. The corresponding region in the S-plane is shown in (b), which is a disk around 0. This region is convex, and hence the loop gain specification $|L+1| \geq 1$ is closed-loop convex.

As another example, we consider the specification

$$\Re L(j\omega) \geq -\alpha, \tag{8}$$

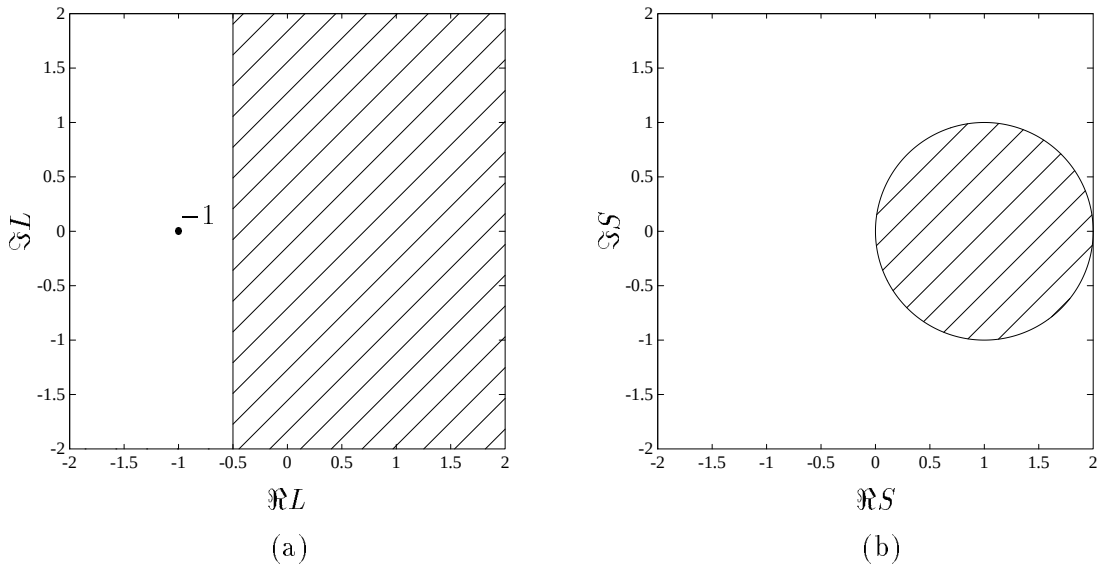
where $0 < \alpha < 1$, which can be expressed in terms of the sensitivity as

$$\Re L(j\omega) \geq -\alpha \iff |2S(j\omega) - 1/(1-\alpha)| \leq 1/(1-\alpha).$$

The case $\alpha = 0.5$ is shown in figure 9.

If the specification (8) is imposed at all frequencies, then the closed-loop system will remain stable even if the loop gain is increased by up to $1/\alpha$. Thus it can be interpreted as enforcing a (positive) gain margin of $20\log_{10}(1/\alpha)$dB.

The specification (8), imposed at all frequencies, can also be interpreted as a circle criterion condition that guarantees the system will remain stable if any memoryless nonlinearity in sector $[0,\ 1/\alpha]$ is introduced into the loop. For example, if the specification (8) is imposed at all frequencies (for any $\alpha < 1$) then actuator or sensor saturation cannot destabilize the control system.



(a) (b)

Figure 9 The region $\Re L \geq -0.5$ in the L-plane is shown in (a). The gain margin specification $\Re L \geq -0.5$ requires the Nyquist plot of L to lie in the shaded region in (a). The corresponding region in the S-plane is the disk shown in (b), and hence the gain margin specification $\Re L \geq -0.5$ is closed-loop convex.

2.2 MAMS Case

2.2.1 In-band and Cutoff Specifications

The analogous results for the MAMS case are:

$$\text{for } \alpha < 1: \qquad \sigma_{\max}(L) \leq \alpha \iff \sigma_{\max}((1-\alpha^2)S - I) \leq \alpha, \tag{9}$$

$$\text{for } \alpha > 1: \qquad \sigma_{\min}(L) \geq \alpha \iff \sigma_{\max}((1-\alpha^2)S - I) \leq \alpha \tag{10}$$

(we have suppressed the frequency arguments for simplicity). Note that the right-hand sides of (9) and (10) are the same. Thus, the inequality on the right-hand side expresses in one formula all reasonable in-band and cutoff loop shaping specifications:

$$\sigma_{\max}((1-\alpha^2)S - I) \leq \alpha \quad \text{both in-band } (\alpha > 1) \text{ and cutoff } (\alpha < 1). \tag{11}$$

(The same correspondences hold with $\tilde{L}$ and $\tilde{S}$.)

We now establish (9). Since $\alpha < 1$, S is nonsingular, and we have

$$\begin{aligned} \sigma_{\max}(L) \leq \alpha &\iff \sigma_{\max}(S^{-1} - I) \leq \alpha \\ &\iff (S^{-1} - I)^*(S^{-1} - I) \leq \alpha^2 I. \end{aligned}$$

Multiplying the last inequality by S^* on the left and S on the right, and multiplying by $1 - \alpha^2 > 0$ (since $\alpha < 1$), gives

$$\begin{aligned} \sigma_{\max}(L) \leq \alpha &\iff (1-\alpha^2)^2 S^*S - (1-\alpha^2)S^* - (1-\alpha^2)S + I \leq \alpha^2 I \\ &\iff \sigma_{\max}((1-\alpha^2)S - I) \leq \alpha, \end{aligned}$$

which is (9).

The in-band result (10) is established in a similar manner. Since $\alpha > 1$, S is nonsingular, and so

$$\begin{aligned} \sigma_{\min}(L) \geq \alpha &\iff \sigma_{\min}(S^{-1} - I) \geq \alpha \\ &\iff (S^{-1} - I)^*(S^{-1} - I) \geq \alpha^2 I. \end{aligned}$$

We proceed as before, except that $1 - \alpha^2 < 0$ (since $\alpha > 1$), so the inequality is reversed:

$$\begin{aligned} \sigma_{\min}(L) \geq \alpha &\iff (1-\alpha^2)^2 S^*S - (1-\alpha^2)S^* - (1-\alpha^2)S + I \leq \alpha^2 I \\ &\iff \sigma_{\max}((1-\alpha^2)S - I) \leq \alpha, \end{aligned}$$

which is (10).

2.2.2 General sector specifications

The in-band and cutoff specifications (9) and (10) are special forms of *general sector specifications*, which we now describe. Given complex matrices C and R such that $(I + C)^*(I + C) > R^*R$, the specification

$$(L - C)^*(L - C) \leq R^*R \tag{12}$$

is closed-loop convex. This specification can be interpreted as requiring L to be in a neighborhood of "radius" R about the "center" C that excludes $-I$. The specification (12) reduces to (9) when $C = 0$ and $R = \alpha I$.

Similarly, given complex matrices C and R such that $(I + C)^*(I + C) < R^*R$, then

$$(L - C)^*(L - C) \geq R^*R \tag{13}$$

is closed-loop convex. This specification can be interpreted as requiring L to be outside a neighborhood of "radius" R about C that includes $-I$. The specification (13) reduces to (10) when $C = 0$ and $R = \alpha I$.

The specifications (12) and (13) are closely connected to the conic sector conditions developed by Zames [31] and Safonov [24]. For example, if C, R, R^{-1}, and L are stable transfer matrices and (12) is imposed at all frequencies, then, using the terminology of Safonov, (12) implies that Graph(L) is inside Cone(C, R). These sector conditions form the basis of various MAMS generalizations of classical frequency domain stability and robustness criteria.

With $C = -I$ and $R = \alpha I$, (13) excludes L from a neighborhood about the critical point $-I$:

$$\sigma_{\min}(L + I) \geq \alpha.$$

This is equivalent to

$$\sigma_{\max}(S) \leq 1/\alpha, \tag{14}$$

which limits the size of the closed-loop sensitivity transfer matrix. Specifications such as (14) that limit the size of a closed-loop transfer matrix, when imposed at all frequencies, can be interpreted as circle criterion constraints that guarantee robustness in the face of various types and locations of nonlinearities.

2.3 Some Convex and Quasiconvex Functionals

We have so far considered loop-shaping *specifications*, which are constraints that a given loop transfer function or matrix either satisfies or not. When these specifications are closed-loop convex, then we can use (numerical) nondifferentiable convex optimization methods to determine whether or not the specifications can be achieved [4]. Practical design problems, however, are more often expressed using a combination of hard constraints (specifications) and soft objectives (performance indices), for example: "maximize the bandwidth subject to a fixed set of crossover and cutoff specifications." Of course such a problem can be solved by repeatedly determining whether fixed sets of specifications are feasible, for example, using a bisection on the objective. Many of the performance indices associated with loop-shaping design are *closed-loop quasiconvex* (when they are to be minimized) or *closed-loop quasiconcave* (when they are to be maximized), which means that these problems can be directly solved. We refer the reader to [4] for a precise definition of these terms and descriptions of numerical methods (*e.g.*, the ellipsoid method) that directly solve compensator design problems that are expressed in terms of closed-loop quasiconvex and quasiconcave performance indices.

Many of these performance indices are constructed in the following general way. We have a family of loop-shaping specifications that is indexed by some number u, in such a way that the specifications always become tighter as the parameter u is decreased. Our performance index is then given by the smallest value of u that the current design satisfies the corresponding specification. Similarly, when smaller values of u correspond to looser specifications, we take the performance index to be the largest value of u such that the corresponding specification is satisfied.

For example, consider the family

$$|L(j\omega)| \geq 10 \quad \text{for} \quad \omega \leq u \tag{15}$$

which is indexed by the number u. As u is decreased, the specification (15) becomes looser, that is, if a given compensator satisfies (15) for a given value of u then it satisfies it for any $\tilde{u} \leq u$. In (15), u can be thought of as the 20dB control bandwidth. The performance index, or 20dB control bandwidth, is the largest value of u such that (15) is satisfied.

Similarly, consider the family

$$|L(j\omega)| \geq u \quad \text{for} \quad \omega \leq 1. \tag{16}$$

The parameter u can be thought of as the minimum loop gain over the fixed bandwidth $0 \leq \omega \leq 1$. The performance index, or in-band disturbance rejection, is the largest value of u such that (16) is satisfied.

If the loop shaping specifications in the family indexed by u are all closed-loop convex, then the performance indices described above are *closed-loop quasiconcave* (when the largest u gives the performance index) and *closed-loop quasiconvex* (when the smallest u gives the performance index). For example, the 20dB control bandwidth, given by the largest u such that (15) holds, is closed-loop quasiconcave; whereas the in-band disturbance rejection, given by the minimum loop gain over the bandwidth $0 \leq \omega \leq 1$, is closed-loop quasiconvex.

Common performance indices associated with loop shaping design that are closed-loop quasiconvex or quasiconcave, as appropriate, include:

- *Bandwidth*, *i.e.*, the smallest frequency for which $|L(j\omega)|$ is less than 3dB, is quasiconcave.

- *System type*, *i.e.*, the multiplicity (possibly zero) of the pole in L at $s = 0$, is an integer valued quasiconcave performance index. In terms of loop shaping specifications, system type is constrained by forcing $l(\omega)$ to grow as $20t$dB/decade for small ω, where t is the type. (System type determines the multiplicity of the zero in S at $s = 0$.)

- *Classical error constant*, given by the absolute value of the first nonvanishing derivative of S at $s = 0$, is convex (and hence quasiconvex).

- *Cutoff frequency*, defined as the smallest frequency such that $|L(j\omega)|$ is less than some number $\alpha < 1$ (-3dB is typical), is quasiconvex.

- *Cutoff rolloff rate*, *i.e.*, the asymptotic rate at which $|L(j\omega)|$ decreases, (which for rational plants is a multiple of -20dB/decade), is quasiconcave.

3 Conclusions

We have shown that many classical and singular value loop shaping problems are closed-loop convex. Consequently, loop shaping problems can be solved by efficient numerical methods. In particular, it can be determined whether or not a compensator exists that satisfies a given set of loop shaping specifications. Loop shaping design problems that are formulated as classical optimization problems, *e.g.*, maximizing

bandwidth subject to given margin and cutoff specifications, can be solved by direct numerical methods for quasiconvex optimization.

A consequence of these observations is that closed-loop convex design methods can be used to do compensator design in a classical loop shaping framework which is familiar to many control engineers. In contrast with classical compensator design methods, in which the designer must decide how to vary *parameters* (such as poles, zeros, and gain) in such a way that the loop transfer function meets the specifications, the designer can directly manipulate the loop shaping *specifications*, since the step of finding a suitable compensator, or determining that none exists, can be automated.

We comment, however, that classical loop shaping is an indirect design technique originally developed before the advent of computers, and is not a particularly good method for compensator design, and especially, computer-aided compensator design. In our opinion, every specification that can be expressed in terms of the loop transfer function or matrix can be more directly expressed in terms of some closed-loop transfer function or transfer matrix, so that compensator design directly from closed-loop (convex) specifications is more direct and natural (see [3, 4]).

A closed-loop convex formulation of a classical design technique such as loop-shaping, even if it is indirect, does have an important practical use: it makes use of the large investment we currently have in teaching, understanding, experience with, and engineering intuition about, classical loop shaping.

We close by noting two important *limitations* of loop shaping via closed-loop convex methods. The first is that closed-loop convex methods generate high order compensators (see, for example, [4, 3]), whereas one of the main advantages of classical loop shaping design is that the designer uses "only as much compensator complexity as is needed to meet the specifications." However, the advent of cheap, high performance, digital signal processors has substantially reduced the relevance of compensator order. Also, even if a compensator designed by closed-loop convex methods is not implemented, knowledge that particular loop shaping specifications can or cannot be achieved is very valuable information to the designer. The designer then *knows* exactly how much performance is given up by using a low order compensator, or some other design method.

The second, and in our opinion more important, drawback of closed-loop convex loop shaping is that margin specifications for MAMS systems that are expressed in terms of singular values or more general sector conditions are often overly conservative when one considers more detail about the types of plant variations that can occur (see [9, 25, 26, 28, 12]). Scaling can greatly reduce the conservatism of these specifications. While a fixed scaling preserves the closed-loop convexity of loop shaping specifications, specifications involving optimal (variable) scaling are not closed-loop convex.

This drawback is not present for SASS control systems. In this case, very general robust performance problems turn out to be closed-loop convex (see [4, p246], [10, §4.3]). In chapters 7 and 8 of [10], these closed-loop convex problems are approxi-

mately transformed into classical loop shaping specifications.

References

[1] F. N. Bailey and C. H. Hui. Loop gain-phase shaping for single-input-single-output robust controllers. *IEEE Control Syst. Mag.*, pages 93–101, January 1991.

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

[3] S. Boyd, V. Balakrishnan, C. Barratt, N. Khraishi, X. Li, D. Meyer, and S. Norman. A new CAD method and associated architectures for linear controllers. *IEEE Trans. Aut. Control*, AC-33(3):268–283, March 1988.

[4] S. Boyd and C. Barratt. *Linear Controller Design: Limits of Performance.* Prentice-Hall, 1991.

[5] S. Boyd, C. Barratt, and S. Norman. Linear controller design: Limits of performance via convex optimization. *Proc. IEEE*, 78(3):529–574, March 1990.

[6] F. M. Callier and C. A. Desoer. *Multivariable Feedback Systems.* Springer-Verlag, 1982.

[7] Y-S. Cho and K. S. Narendra. An off-axis circle criterion for the stability of feedback systems with a monotonic nonlinearity. *IEEE Trans. Aut. Control*, AC-13:413–416, 1968.

[8] R. C. Dorf. *Modern Control Systems.* Addison-Wesley, 5th edition, 1988.

[9] J. Doyle. Analysis of feedback systems with structured uncertainties. *IEE Proc.*, 129-D(6):242–250, November 1982.

[10] J. Doyle, B. Francis, and A. Tannenbaum. *Feedback Control Theory.* Macmillan, 1992.

[11] J. Doyle and G. Stein. Multivariable feedback design: Concepts for a classical/modern synthesis. *IEEE Trans. Aut. Control*, AC-26(1):4–16, February 1981.

[12] M.K.H. Fan, A.L. Tits, and J.C. Doyle. Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics. *IEEE Trans. Aut. Control*, 36(1):25–38, January 1991.

[13] G. F. Franklin, J. D. Powell, and A. Emami-Naeni. *Feedback Control of Dynamic Systems*. Addison-Wesley, 1986.

[14] J. S. Freudenberg and D. P. Looze. *Frequency Domain Properties of Scalar and Multivariable Feedback Systems*. Lecture Notes in Control and Information Sciences. Springer-Verlag, 1988.

[15] J. W. Helton and A. Sideris. Frequency response algorithms for $\mathbf{h}_\infty$ optimization with time domain constraints. *IEEE Trans. Aut. Control*, 34(4):427–434, April 1989.

[16] I. M. Horowitz. *Synthesis of Feedback Systems*. Academic Press, New York, 1963.

[17] B. C. Kuo. *Automatic Control Systems*. Prentice-Hall, 4th edition, 1982.

[18] J. Lunze. *Robust Multivariable Feedback Control*. Prentice-Hall, 1989.

[19] J. M. Maciejowski. *Multivariable Feedback Design*. Addison-Wesley, 1989.

[20] K. S. Narendra and R. M. Goldwyn. A geometrical criterion for the stability of certain nonlinear nonautonomous systems. *IEEE Trans. Circuit Theory*, CT-11:406–408, September 1964.

[21] K. Ogata. *Modern Control Engineering*. Prentice-Hall, Englewood Cliffs, New Jersey, second edition, 1990.

[22] E. Polak and S. Salcudean. On the design of linear multivariable feedback systems via constrained nondifferentiable optimization in $\mathbf{H}_\infty$ spaces. *IEEE Trans. Aut. Control*, AC-34(3):268–276, 1989.

[23] V. M. Popov. Absolute stability of nonlinear systems of automatic control. *Automation and Remote Control*, 22:857–875, 1962.

[24] M. G. Safonov. *Stability and Robustness of Multivariable Feedback Systems*. MIT Press, Cambridge, 1980.

[25] M. G. Safonov. Stability margins of diagonally perturbed multivariable feedback systems. *IEE Proc.*, 129-D:251–256, 1982.

[26] M. G. Safonov and J. Doyle. Minimizing conservativeness of robust singular values. In S. G. Tzafestas, editor, *Multivariable Control*, pages 197–207. D. Reidel, 1984.

[27] I. W. Sandberg. A frequency-domain condition for the stability of feedback systems containing a single time-varying nonlinear element. *Bell Syst. Tech. J.*, 43(3):1601–1608, July 1964.

[28] G. Stein and J. C. Doyle. Beyond singular values and loop shapes. *J. Guidance and Control*, 14:5–16, 1991.

[29] J. C. Willems and G. L. Blankenship. Frequency domain stability criteria for stochastic systems. *IEEE Trans. Aut. Control*, AC-16:292–299, 1971.

[30] G. Zames. On the input-output stability of nonlinear time-varying feedback systems—Part II: Conditions involving circles in the frequency plane and sector nonlinearities. *IEEE Trans. Aut. Control*, AC-11:465–476, July 1966.

[31] G. Zames. On the input-output stability of nonlinear time-varying feedback systems—Part I: Conditions derived using concepts of loop gain, conicity, and positivity. *IEEE Trans. Aut. Control*, AC-11:228–238, April 1966.