J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*

H INFINITY CONTROL

This article describes an optimal multivariable control system design technique for achieving robustness under external disturbances and model uncertainties. We show that many robust control problems can be formulated as $\mathscr{H}_\infty$ norm optimization problems, and we describe analytically their solutions. We also give guidelines to the choice of design parameters and insights to this optimal and robust control theory.

One of the motivations for the original introduction of $\mathscr{H}_\infty$ methods by Zames (1) was to bring plant uncertainty, specified in the frequency domain, back to the center stage as it had been in classical control, in contrast to analytic methods such as linear quadratic Gaussian (*LQG*) control. The $\mathscr{H}_\infty$ norm was found to be appropriate for specifying both the level of plant uncertainty and the signal gain from disturbance inputs to error outputs in the controlled system.

The "standard" $\mathscr{H}_\infty$ optimal control problem is concerned with the feedback system shown in Fig. 1(a) where w represents an external disturbance, y is the measurement available to the controller, u is the output from the controller, and z is an error signal that should be kept small. The transfer function matrix G represents not only the conventional plant to be controlled but also any weighting functions included to specify the desired performance, which will be discussed in more detail later. Suppose that G is partitioned consistent with the inputs w and u and outputs z and y as

$$G(s) = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}$$

The closed loop transfer function from w to z, denoted by T_{zw}, is defined to be the linear fractional transformation (*LFT*) of G on K:

$$T_{zw} = \mathscr{F}_\ell(G, K) := G_{11} + G_{12}K(I - G_{22}K)^{-1}G_{21}$$

The $\mathscr{H}_\infty$ optimal control problem is then to design a stabilizing controller K, so as to minimize the $\mathscr{H}_\infty$ norm of T_{zw}, which is defined in the next section and is denoted by $\|T_{zw}\|_\infty$. The $\mathscr{H}_\infty$ norm gives the maximum energy gain, or sinusoidal gain of the system. This is in contrast to the $\mathscr{H}_2$ norm $\|T_{zw}\|_2$, which for example gives the variance of the output given white noise disturbances. The important property of the $\mathscr{H}_\infty$ norm comes from the application of the small gain theorem, which states that if $\|T_{zw}\|_\infty \le \gamma$, then the system in Fig. 1(b) will be stable for all stable Δ with $\|\Delta\|_\infty < 1/\gamma$. It is probably the case that this robust stability consequence was one of the main motivations for the development of $\mathscr{H}_\infty$ methods. The synthesis of controllers that achieve an $\mathscr{H}_\infty$ norm specification hence gives a well-defined practical and mathematical problem.

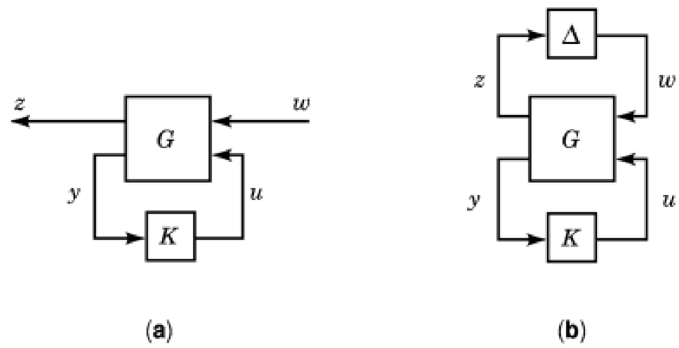

Fig. 1. Most control systems can be put in this unified framework where w represents an external disturbance, y is the measurement available to the controller, u is the output from the controller, and z is an error signal that it is desired to keep small.

$\mathcal{H}_2$ and $\mathcal{H}_\infty$ Norms

We consider a q-input and p-output dynamical system with the matrix transfer function $G(s)$. Let $G(s)$ have the following stabilizable and detectable state space realization:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned}$$

We shall denote this state-space realization by

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] := C(sI - A)^{-1}B + D$$

Many control design problems can be regarded as finding a suitable controller so that the undesirable responses of the system are made small in some sense. There are obviously many ways to define the smallness for a given control problem. Here we are mainly interested in one way of defining the smallness: the $\mathcal{H}_\infty$ norm. For comparison we shall also mention another more classical way of defining the smallness in terms of the $\mathcal{H}_2$ norm.

Let $\mathcal{RH}_2$ denote the set of strictly proper and real rational stable transfer matrices. In terms of state-space realizations, $\mathcal{RH}_2$ is simply the set of finite dimensional systems with $D = 0$ and stable A. The $\mathcal{H}_2$ norm of a $G(s) \in \mathcal{RH}_2$ is defined as

$$\begin{aligned} \|G\|_2 &:= \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}[G^*(j\omega)G(j\omega)]\,d\omega} \\ &= \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \sum_{i=1}^{\min\{p,q\}} \sigma_i^2(G(j\omega))\,d\omega} \end{aligned}$$

where trace(M), $M*$, and $\sigma_i(M)$ denote, respectively, the trace, the complex conjugate transpose, and the ith singular value of the matrix M. This norm can also be computed in the time domain as

$$\|G\|_2 = \sqrt{\int_0^\infty \text{trace}\{g^T(t)g(t)\}\,dt}$$

where $g(t)$ is the inverse Laplace transform of $G(s)$. Thus the $\mathscr{H}_2$ norm of a system is a measure of the total energy of the system impulse response. It can be computed using state-space realization as

$$\|G\|_2 = \sqrt{\text{trace}(B^T QB)} = \sqrt{\text{trace}(CPC^T)}$$

where Q and P are observability Gramian and controllability Gramian, which can be obtained from the following Lyapunov equations

$$AP + PA^T + BB^T = 0 \quad A^T Q + QA + C^T C = 0$$

Let $\mathscr{RH}_\infty$ denote the set of proper (but not necessarily strictly proper) and real rational stable transfer matrices. In terms of state-space realizations, $\mathscr{RH}_\infty$ includes all finite dimensional systems with stable A matrices. The $\mathscr{H}_\infty$ norm of a transfer matrix $G(s) \in \mathscr{RH}_\infty$ is defined as

$$\|G\|_\infty := \sup_{\omega\in\mathbb{R}} \bar{\sigma}\,[G(j\omega)]$$

where $\bar{\sigma}(M)$ denotes the largest singular value of a matrix M. When $G(s)$ is a single input and single output system, the $\mathscr{H}_\infty$ norm of the $G(s)$ is simply the peak value on the Bode magnitude plot of the frequency response $G(j\omega)$. It can also be regarded as the largest possible amplification factor of the system's steady state response to sinusoidal excitations. For example, the steady state response of the system with respect to a sinusoidal input $u(t) = U\sin(\omega_0 t + \phi)$ is

$$y(t) = U|G(j\omega_0)|\sin(\omega_0 t + \phi + \angle G(j\omega_0))$$

and thus the maximum possible amplification factor is

$$\sup_{\omega_0} |G(j\omega_0)|$$

which is precisely the $\mathscr{H}_\infty$ norm of the transfer function. As an example, consider a standard second-order system

$$G(s) = \frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2}, \quad 0 < \xi < 1/\sqrt{2}$$

Then $\omega_{\max} = \omega_n\sqrt{1-2\xi^2}$ $1 - 2\xi^2$ and $\|G\|_\infty = |G(j\omega_{\max})| = 1/(2\xi\sqrt{1-\xi^2}$ $1 - \xi^2)$. If $G(s)$ is the description of a structure vibration, then $\omega_{\max}$ would be the most dangerous exciting frequency.

In the multiple input and multiple output (*MIMO*) case, the $\mathscr{H}_\infty$ norm of a transfer matrix $G \in \mathscr{RH}_\infty$ is the peak value on the largest singular value Bode plot of the frequency response $G(j\omega)$. Analogous to the scalar case, the $\mathscr{H}_\infty$ norm of $G(s)$ can also be regarded as the largest possible amplification factor of the system's steady state response to sinusoidal excitations in the following sense: Let the sinusoidal inputs be

$$u(t) = \begin{bmatrix} u_1 \sin(\omega_0 t + \phi_1) \\ u_2 \sin(\omega_0 t + \phi_2) \\ \vdots \\ u_q \sin(\omega_0 t + \phi_q) \end{bmatrix}, \quad \hat{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_q \end{bmatrix}$$

Then the steady state response of the system can be written as

$$y(t) = \begin{bmatrix} y_1 \sin(\omega_0 t + \theta_1) \\ y_2 \sin(\omega_0 t + \theta_2) \\ \vdots \\ y_p \sin(\omega_0 t + \theta_p) \end{bmatrix}, \quad \hat{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix}$$

for some y_i, θ_i, $i = 1, 2, \ldots, p$. Furthermore,

$$\|G\|_\infty = \sup_{\phi_i, \omega_0, \hat{u}} \frac{\|\hat{y}\|}{\|\hat{u}\|}$$

where $\|\cdot\|$ is the Euclidean norm.

The $\mathscr{H}_\infty$ norm of a stable transfer matrix can also be thought of as the maximum amplification factor of the input energy of the system at the output of the system. More precisely,

$$\|G\|_\infty = \sup_{u \in \mathscr{L}_2[0,\infty)} \frac{\|g * u\|_2}{\|u\|_2}$$

where $*$ denotes the time domain convolution and $\mathscr{L}_2[0, \infty)$ denotes the space of all square integrable functions with the norm defined as $\|f\|_2 := \sqrt{\int_0^\infty \|f(t)\|^2 dt}$. Thus it is important to make the $\mathscr{H}_\infty$ norm of all undesirable transfer functions small in a feedback control system. That is one of the motivations for the development of $\mathscr{H}_\infty$ control theory.

This discussion shows that the $\mathscr{H}_\infty$ norm of a transfer function can, in principle, be obtained either graphically or experimentally. To get an estimate, set up a fine grid of frequency points, $\{\omega_1, \ldots, \omega_N\}$. Then an estimate for $\|G\|_\infty$ is

$$\|G\|_\infty \approx \max_{1 \le k \le N} \overline{\sigma}\{G(j\omega_k)\}$$

This value is usually read directly from a singular value Bode plot. The $\mathscr{H}_\infty$ norm can also be computed directly using state-space representations. Let

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathscr{RH}_\infty$$

Then

$$\begin{aligned}\|G\|_\infty = \inf\{\gamma : \bar{\sigma}(D) < \gamma \\ \text{and } H \text{ has no eigenvalues on the imaginary axis}\end{aligned}$$

where

$$\begin{aligned} H &= \begin{bmatrix} A + BR^{-1}D^TC & BR^{-1}B^T \\ -C^T(I + DR^{-1}D^T)C & -(A + BR^{-1}D^TC)^T \end{bmatrix} \\ R &= \gamma^2 I - D^T D \end{aligned}$$

Hence, the $\mathscr{H}_\infty$ norm of a matrix transfer function can be calculated to the specified accuracy by a bisection search.

Example. Consider a mass/spring/damper system as shown in Fig. 2. The dynamical system can be described by the following differential equations:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\dfrac{k_1}{m_1} & \dfrac{k_1}{m_1} & -\dfrac{b_1}{m_1} & \dfrac{b_1}{m_1} \\ \dfrac{k_1}{m_2} & -\dfrac{k_1 + k_2}{m_2} & \dfrac{b_1}{m_2} & -\dfrac{b_1 + b_2}{m_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \dfrac{1}{m_1} & 0 \\ 0 & \dfrac{1}{m_2} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$$

Suppose that $G(s)$ is the transfer matrix from (F_1, F_2) to (x_1, x_2) and suppose $k_1 = 1$, $k_2 = 4$, $b_1 = 0.2$, $b_2 = 0.1$, $m_1 = 1$, and $m_2 = 2$ with appropriate units. Then the $\mathscr{H}_2$ norm of this transfer matrix is $\|G(s)\|_2 = 2.56$, whereas the $\mathscr{H}_\infty$ norm of this transfer matrix is $\|G(s)\|_\infty = 11.47$, which is shown as the peak of the largest singular value Bode plot in Fig. 3. Since the peak is achieved at $\omega_{\max} = 0.8483$, exciting the system using the

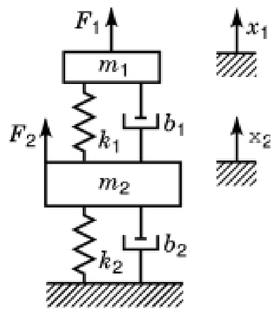

Fig. 2. A two mass/spring/damper system with the external forces F_1 and F_2 as inputs and the positions of the two masses as outputs.

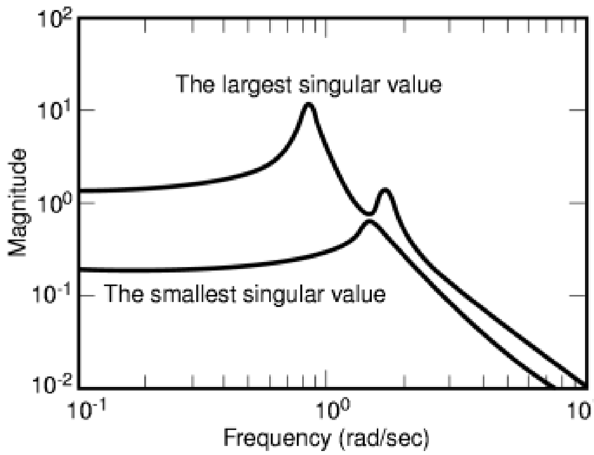

Fig. 3. The singular value Bode plot of the two mass/spring/damper system. The $\mathscr{H}_\infty$ norm $\|G\|_\infty$ is the peak of the largest singular value plot of $G(j\omega)$.

following sinusoidal input

$$\begin{bmatrix} F_1 \\ F_2 \end{bmatrix} = \begin{bmatrix} 0.9614\sin(0.8483t) \\ 0.2753\sin(0.8483t - 0.12) \end{bmatrix}$$

gives the steady state response of the system as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 11.47 \times 0.9614\sin(0.8483t - 1.5483) \\ 11.47 \times 0.2753\sin(0.8483t - 1.4283) \end{bmatrix}$$

This shows that the system response will be amplified 11.47 times for an input signal at the frequency $\omega_{\max}$, which could be undesirable if F_1 and F_2 are disturbance force and x_1 and x_2 are the positions to be kept steady. We will see later how to design an $\mathscr{H}_\infty$ feedback controller to suppress this kind of vibration.

We note that the state-space computational method is usually much more accurate than the graphical method. Consider, for example, the standard second-order system again with $\omega_n = 1$ and $\xi = 0.01$. By the analytic formula or the state-space computational method, we get that the $\mathscr{H}_\infty$ norm is 50.0025. To estimate the $\mathscr{H}_\infty$ norm graphically, we set up a frequency grid to compute the frequency response of G over a suitable range of frequency. Take, for example, 200 points in the frequency range of [0.1, 10] uniformly on the log scale, then we get an estimate of the norm $\approx$ 33.0743. This shows clearly that the graphical method may lead to a wrong answer for a lightly damped system if the frequency grid is not sufficiently dense. Indeed, we would get the $\mathscr{H}_\infty$ norm $\approx$ 43.5567, 48.1834, and 49.5608 from the graphical method if 400, 800, and 1600 frequency points are used, respectively.

Weighted $\mathscr{H}_\infty$ Performance

We now consider how to formulate some performance objectives into mathematically tractable problems. It is well known that the performance objectives of a feedback system can usually be specified in terms of requirements on the sensitivity functions and/or complementary sensitivity functions or in terms of some other closed-loop transfer functions. For instance, the performance criteria for a scalar system may be specified as requiring

$$\begin{cases} |S(j\omega)| \le \epsilon, & \forall \omega \le \omega_0, \\ |S(j\omega)| \le M, & \forall \omega > \omega_0 \end{cases}$$

with $S(j\omega) = 1/(1 + P(j\omega)K(j\omega))$ where P is the plant and K is the controller. However, it is much more convenient to reflect the system performance objectives by choosing appropriate weighting functions. For example, this performance objective can be written as

$$|W_e(j\omega)S(j\omega)| \le 1, \quad \forall \omega$$

with

$$|W_e(j\omega)| = \begin{cases} 1/\epsilon, & \forall \omega \le \omega_0 \\ 1/M, & \forall \omega > \omega_0 \end{cases}$$

In order to use W_e in control design, a rational transfer function $W_e(s)$ is usually used to approximate this frequency response.

The advantage of using weighted performance specifications is obvious in multivariable system design. First, some components of a vector signal are usually more important than others. Second, each component of the signal may not be measured in the same units; for example, some components of the output error signal may be measured in terms of length, and the others may be measured in terms of voltage. Weighting functions are essential to make these components comparable. Also, we might be primarily interested in rejecting errors in a certain frequency range (e.g., low frequencies); hence, some frequency-dependent weights must be chosen.

Consider a standard feedback diagram in Fig. 4. The weighting functions in Fig. 4 are chosen to reflect the design objectives and knowledge on the disturbances and sensor noise. For example, W_d and W_i may be chosen to reflect the frequency contents of the disturbances d and d_i. The weighting matrix W_n is used to model the frequency contents of the sensor noise, whereas W_e may be used to reflect the requirements on the shape of certain closed-loop transfer functions (e.g., the shape of the output sensitivity function). Similarly, W_u may be used to reflect some restrictions on the control or actuator signals, and the dashed precompensator W_r is an

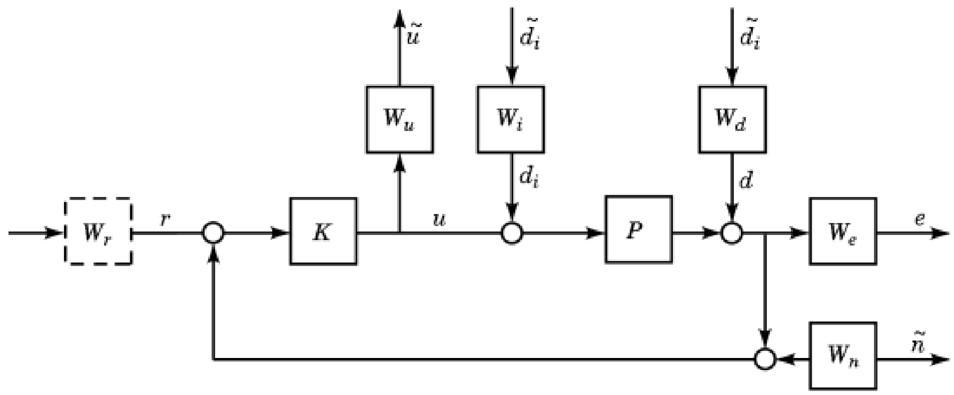

Fig. 4. Standard feedback configuration with disturbance weights and performance weights. W_i, W_d, and W_n represent the frequency contents of the input disturbance, output disturbance, and the sensor noise. W_e represents the disturbance rejection requirement, and W_u puts the limit on the control effort. W_r shapes the input signal.

optional element used to achieve deliberate command shaping or to represent a nonunity feedback system in equivalent unity feedback form.

A typical control design may involve making the sensitivity function small over a suitable frequency range while keeping the control effort within a reasonable limit. This may be mathematically formulated as minimizing

$$\sup_{\|\tilde{d}\|_2 \le 1} \|e\|_2 = \|W_e(I+PK)^{-1}W_d\|_\infty$$

subject to some restrictions on the control energy or control bandwidth:

$$\sup_{\|\tilde{d}\|_2 \le 1} \|\tilde{u}\|_2 = \|W_u K(I+PK)^{-1}W_d\|_\infty$$

Or more frequently, one may introduce a parameter ρ and a mixed criterion

$$\sup_{\|\tilde{d}\|_2 \le 1} \left\{\|e\|_2^2 + \rho^2\|\tilde{u}\|_2^2\right\} = \left\|\begin{bmatrix} W_e(I+PK)^{-1}W_d \\ \rho W_u K(I+PK)^{-1}W_d \end{bmatrix}\right\|_\infty^2$$

Note that ρ can be absorbed into W_u, so there is no loss of generality in assuming $\rho = 1$. Finding a controller so that the $\mathscr{H}_\infty$ norm of a certain closed-loop transfer function, such as the preceding one, is minimized is the $\mathscr{H}_\infty$ control problem.

Similar $\mathscr{H}_2$ norm minimization problems can be formulated if the disturbance is modeled as white noise and the performance is measured in terms of output power.

Robust Stabilization

Another way that a weighted $\mathscr{H}_\infty$ norm minimization problem can arise naturally is when we consider robust stability and robust performance of a closed-loop system with model uncertainties. For example, consider a unity feedback system with a family of additively perturbed uncertain dynamical systems:

$$\{P + W_1 \Delta W_2, W_1, W_2, \Delta \in \mathscr{H}_\infty\}$$

and assume that K stabilizes the nominal plant P. Then by the small gain theorem, the uncertain system is stable for all admissible Δ with $\|\Delta\|_\infty < 1/\gamma$ if and only if

$$\|W_2 K (I + PK)^{-1} W_1\|_\infty \leq \gamma$$

Therefore, a related synthesis problem is to find a controller K so that this inequality holds.

As another example, let $P = \tilde{M}^{-1}\tilde{N}$ be a normalized coprime factorization, that is,

$$\tilde{M}(s), \quad \tilde{N}(s) \in \mathscr{RH}_\infty, \quad \tilde{M}(s)\tilde{M}^T(-s) + \tilde{N}(s)\tilde{N}^T(-s) = I$$

Now consider a family of coprime factor perturbed uncertain systems:

$$\mathbf{F_p} := \{(\tilde{M} + \tilde{\Delta}_M)^{-1}(\tilde{N} + \tilde{\Delta}_N), \tilde{\Delta}_M, \tilde{\Delta}_N \in \mathscr{RH}_\infty, \|[\tilde{\Delta}_N \ \tilde{\Delta}_M]\|_\infty < \epsilon\}$$

A controller K stabilizing the nominal system P will robustly stabilize the family $\mathbf{F_p}$ if and only if

$$\left\| \begin{bmatrix} I \\ K \end{bmatrix} (I + PK)^{-1}\tilde{M}^{-1} \right\|_\infty \leq 1/\epsilon$$

Similarly, many other robust stability problems can be formulated. It should be noted that the $\mathscr{H}_2$ norm cannot be used in the robust stability test because it does not satisfy a key multiplicative property (i.e., $\|G_1 G_2\|_2 \nleq \|G_1\|_2 \|G_2\|_2$ in general).

Selection of Weighting Functions

As we mentioned previously, a very important step in the $\mathscr{H}_\infty$ control design process is to choose the appropriate weights. The appropriate choice of weights for a particular problem is not trivial. On many occasions, the weights are chosen purely as a design parameter without any physical bases, so these weights may be treated as tuning parameters that are chosen by the designer to achieve the best compromise between the conflicting objectives. Hence, the selection of weighting functions for a specific design problem often involves ad hoc fixing, iterations, and fine-tuning. It should be guided by the expected system inputs and the relative importance of the outputs. It is very hard to give a general formula for the weighting functions that will work in every case. Nevertheless, we shall try to give some guidelines in this section by looking at a typical single input single output (*SISO*) problem.

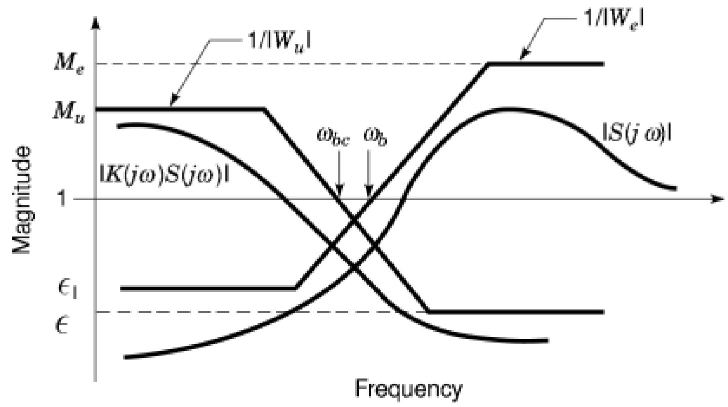

Fig. 5. The desired shapes of S and KS and their upper bounds.

Consider the SISO feedback system shown in Fig. 4. Then the tracking error is

$$e = r - y = S(r - d) + Tn - SPd_i$$

where $S = (I + PK)^{-1}$ is the output sensitivity function and $T = I - S = PK(I + PK)^{-1}$ is the output complementary sensitivity function. Note that tracking error is closely related to the low-frequency gain of S. In particular, we must keep $|S|$ small over the range of frequencies, typically low frequencies where r and d are significant. For example, if we need the steady state error with respect to a step input to be no greater than ϵ, then we need $|S(0)| \leq \epsilon$. Hence, the steady state tracking requirement can be fulfilled by constraining the low-frequency gain of S. From classical control theory, we know that the dynamical quality of the system time response can be quantified by rise time, settling time, and percent overshoot. Furthermore, the speed of the system response is inversely proportional to the closed-loop bandwidth and the overshoot of the system response increases with the resonant peak sensitivity defined as $M_s := \|S\|_\infty$. Let $\omega_b = \min\{\omega: |S(j\omega)| \geq 1\}$. Then we can regard ω_b as the closed-loop bandwidth because, beyond this frequency, the closed-loop system will not be able to track the reference and the disturbance will actually be amplified.

Now suppose that we are given the time domain performance specifications. Then we can determine the corresponding requirements in frequency domain in terms of the low-frequency gain, the bandwidth ω_b, and the peak sensitivity M_s. Hence, a good control design should result in a sensitivity function S satisfying all these requirements as shown in Fig. 5. These requirements can be approximately represented as

$$|S(s)| \leq \left|\frac{1}{W_e(s)}\right|, \quad s = j\omega, \quad \forall \omega, \quad W_e(s) = \frac{s/M_s + \omega_b}{s + \omega_b \epsilon}$$

If a steeper transition between low frequency and high frequency is desired, the weight W_e can be modified as follows:

$$W_e(s) = \left(\frac{s/\sqrt[k]{M_s} + \omega_b}{s + \omega_b \sqrt[k]{\epsilon}}\right)^k$$

for some integer $k \geq 1$.

The selection of control weighting function W_u follows similarly from the preceding discussion by considering the control signal equation

$$u = KS(r - n - d) - Td_i$$

The magnitude of $|KS|$ in the low-frequency range is essentially limited by the allowable cost of control effort and saturation limit of the actuators; hence, in general the maximum gain M_u of KS can be fairly large, whereas the high-frequency gain is essentially limited by the controller bandwidth ω_{bc} and the (sensor) noise frequencies. Ideally, we would like to roll off as fast as possible beyond the desired control bandwidth so that the high-frequency noises are attenuated as much as possible. Hence, a candidate weight W_u would be

$$W_u = \frac{s + \omega_{bc}/M_u}{\omega_{bc}}$$

However, the standard $\mathscr{H}_\infty$ control design techniques cannot be applied directly to a problem with an improper control weighting function. Hence, we shall introduce a faraway pole to make W_u proper:

$$W_u = \frac{s + \omega_{bc}/M_u}{\epsilon_1 s + \omega_{bc}}$$

for a small $\epsilon_1 > 0$ as shown in Fig. 5. Similarly, if a faster rolloff is desired, we may choose

$$W_u = \left(\frac{s + \omega_{bc}/\sqrt[k]{M_u}}{\sqrt[k]{\epsilon_1}s + \omega_{bc}}\right)^k$$

for some integer $k \geq 1$.

The weights for *MIMO* problems can be initially chosen as diagonal matrices with each diagonal term chosen in the preceding form.

General Problem Formulation and Solutions

All the disturbance rejection problems and robust stabilization problems discussed in the previous sections can be put in a unified framework of linear fractional transformation as shown in Fig. 1(a). For example,

$$\begin{bmatrix} W_e(I+PK)^{-1}W_d \\ W_uK(I+PK)^{-1}W_d \end{bmatrix} = \mathscr{F}_\ell(G_1, K)$$

$$\text{with } G_1 = \begin{bmatrix} \begin{bmatrix} W_eW_d \\ 0 \end{bmatrix} & \begin{bmatrix} -W_eP \\ W_u \end{bmatrix} \\ W_d & -P \end{bmatrix}$$

$$W_2K(I+PK)^{-1}W_1 = \mathscr{F}_\ell(G_2, K) \quad \text{with } G_2 = \begin{bmatrix} 0 & W_2 \\ W_1 & -P \end{bmatrix}$$

and

$$\begin{bmatrix} I \\ K \end{bmatrix} (I+PK)^{-1}\tilde{M}^{-1} = \mathscr{F}_\ell(G_3, K) \quad \text{with } G_3 = \begin{bmatrix} \begin{bmatrix} \tilde{M}^{-1} \\ 0 \end{bmatrix} & \begin{bmatrix} -P \\ I \end{bmatrix} \\ \tilde{M}^{-1} & -P \end{bmatrix}$$

For the following discussion, let us assume that state-space models of G and K are available and that their realizations are assumed to be stabilizable and detectable. We say that a controller is *admissible* if it internally stabilizes the system.

Optimal $\mathscr{H}_2$ Control:
Find an admissible controller $K(s)$ such that $\|T_{zw}\|_2$ is minimized.

Optimal $\mathscr{H}_\infty$ Control:
Find an admissible controller $K(s)$ such that $\|T_{zw}\|_\infty$ is minimized.

It should be noted that the optimal $\mathscr{H}_\infty$ controllers as defined here are generally not unique for MIMO systems. Furthermore, finding an optimal $\mathscr{H}_\infty$ controller is often both numerically and theoretically complicated. This is certainly in contrast to the standard LQG or $\mathscr{H}_2$ theory, in which the optimal controller is unique and can be obtained by solving two Riccati equations without iterations. Knowing the achievable optimal (minimum) $\mathscr{H}_\infty$ norm may be useful theoretically because it sets a limit on what we can achieve. In practice, however, it is often not necessary and sometimes even undesirable to design an optimal controller, and it is usually much cheaper to obtain controllers that are very close in the norm sense to the optimal ones, which will be called *suboptimal controllers.* A suboptimal controller may also have other nice properties (e.g., lower bandwidth) over the optimal ones.

Suboptimal $\mathscr{H}_\infty$ Control:
Given $\gamma > 0$, find an admissible controller $K(s)$, if there are any, such that $\|T_{zw}\|_\infty < \gamma$.

For these reasons mentioned, we shall focus our attention on the suboptimal $\mathscr{H}_\infty$ control. We shall assume that the realization of the transfer matrix G takes the following form:

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right]$$

which is compatible with the dimensions of $z(t) \in \mathbb{R}^{p_1}$, $y(t) \in \mathbb{R}^{p_2}$, $w(t) \in \mathbb{R}^{m1}$, and $u(t) \in \mathbb{R}^{m2}$ and the state $x(t) \in \mathbb{R}^n$. We make the following assumptions:

1. (A, B_2) is stabilizable, and (C_2, A) is detectable.
2. $R_1 = D_{12}^T D_{12} > 0$ and $R_2 = D_{21} D_{21}^T > 0$.
3. $\begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all ω.
4. $\begin{bmatrix} A - j\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ has full row rank for all ω.

Assumption 1 is necessary for the existence of stabilizing controllers. Assumption 2 means that the penalty on $z = C_1x + D_{12}u$ includes a nonsingular penalty on the control u, that the exogenous signal w includes both plant disturbance and sensor noise, and that the sensor noise weighting is nonsingular. Relaxation of assumption 2 leads to singular control problems.

Assumptions 3 and 4 are made for a technical reason: together with assumption 1 they guarantee that the two algebraic Riccati equations in the corresponding LQG or $\mathscr{H}_2$ problem have the desired solutions. Dropping assumptions 3 and 4 would make the solution very complicated.

Define

$$A_x = A - B_2R_1^{-1}D_{12}^TC_1, \quad A_y = A - B_1D_{21}^TR_2^{-1}C_2,$$
$$P = B_1(I - D_{21}^TR_2^{-1}D_{21})B_1^T, \quad Q = C_1^T(I - D_{12}R_1^{-1}D_{12}^T)C_1$$

The following $\mathscr{H}_2$ and $\mathscr{H}_\infty$ control results can be found in Refs. 2 and 3.

Theorem 1. There exists a unique optimal controller

$$K_2(s) := \left[\begin{array}{c|c} A + B_2F_2 + L_2C_2 & -L_2 \\ \hline F_2 & 0 \end{array}\right]$$

that minimizes $\|T_{zw}\|_2$ where

$$F_2 := -R_1^{-1}(D_{12}^TC_1 + B_2^TX_2), \quad L_2 := -(B_1D_{21}^T + Y_2C_2^T)R_2^{-1}$$

$X_2 \geq 0$ and $Y_2 \geq 0$ are the solutions to

$$X_2A_x + A_x^TX_2 - X_2B_2R_1^{-1}B_2^TX_2 + Q = 0$$
$$Y_2A_y^T + A_yT_2 - Y_2C_2^TR_2^{-1}C_2Y_2 + P = 0$$

such that both $A_x - B_2R^{-1}{}_1B^T{}_2X_2$ and $A_y - Y_2C^T{}_2R^{-1}{}_2C_2$ are stable. Moreover,

$$\min\|T_{zw}\|_2 = \sqrt{\operatorname{trace}(B_1^TX_2B_1) + \operatorname{trace}(F_2Y_2F_2^TR_1)}$$

Theorem 2. Suppose G satisfies assumptions 1–4. Then there exists an admissible controller K_∞ such that $\|\mathscr{F}_\ell(G, K_\infty)\|_\infty < \gamma$ (i.e., $\|T_{zw}\|_\infty < \gamma$) if and only if

(1) there exists an $X_\infty \geq 0$ such that

$$X_\infty A_x + A_x^TX_\infty + X_\infty(B_1B_1^T/\gamma^2 - B_2R_1^{-1}B_2^T)X_\infty + Q = 0$$

and $A_x + (B_1B^T{}_1/\gamma^2 - B_2R^{-1}{}_1B^T{}_2)X_\infty$ is stable;

(2) there exists a $Y_\infty \geq 0$ such that

$$Y_\infty A_y^T + A_yY_\infty + Y_\infty(C_1^TC_1/\gamma^2 - C_2^TR_2^{-1}C_2)Y_\infty + P = 0$$

and $A_y + Y_\infty(C^T{}_1C_1/\gamma^2 - C^T{}_2R^{-1}{}_2C_2)$ is stable;

(3) $\rho(X_\infty Y_\infty) < \gamma^2$ where $\rho(\cdot)$ denotes the spectral radius.

Furthermore, if these conditions are satisfied, all internally stabilizing controllers $K_\infty(s)$ satisfying $\|\mathscr{F}_\ell(G, K_\infty)\|_\infty < \gamma$ can be parameterized as

$$K_\infty = \mathscr{F}_\ell(M_\infty, Q)$$

for any $Q \in \mathscr{R}\mathscr{H}_\infty$ such that $\|Q\|_\infty < \gamma$ where

$$M_\infty = \left[\begin{array}{c|cc} \hat{A}_\infty & -Z_\infty L_\infty & Z_\infty(B_2 + Y_\infty C_1^T D_{12})R_1^{-1/2} \\ \hline F_\infty & 0 & R_1^{-1/2} \\ -R_2^{-1/2}(C_2 + D_{21}B_1^T X_\infty/\gamma^2) & R_2^{-1/2} & 0 \end{array}\right]$$

$$\begin{aligned} F_\infty &= -R_1^{-1}(D_{12}^T C_1 + B_2^T X_\infty) \\ L_\infty &= -(B_1 D_{21}^T + Y_\infty C_2^T)R_2^{-1} \\ Z_\infty &= (I - \gamma^{-2} Y_\infty X_\infty)^{-1} \\ \hat{A}_\infty &= A + B_1 B_1^T X_\infty/\gamma^2 + B_2 F_\infty + Z_\infty L_\infty(C_2 + D_{21}B_1^T X_\infty/\gamma^2) \end{aligned}$$

The controller obtained by setting $Q = 0$

$$K_\infty = \left[\begin{array}{c|c} \hat{A}_\infty & -Z_\infty L_\infty \\ \hline F_\infty & 0 \end{array}\right]$$

is called the *central controller*. Comparing the $\mathscr{H}_\infty$ central controller to the $\mathscr{H}_2$ optimal controller, we can see that the $\mathscr{H}_\infty$ central controller will approach the optimal $\mathscr{H}_2$ controller as $\gamma \to \infty$.

Example. Consider again the two mass/spring/damper system shown in Fig. 2. Assume that F_1 is the control force, F_2 is the disturbance force, and the measurements of x_1 and x_2 are corrupted by measurement noise:

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + W_n \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}$$

$$W_n = \begin{bmatrix} \dfrac{0.01(s+10)}{s+100} & 0 \\ 0 & \dfrac{0.01(s+10)}{s+100} \end{bmatrix}$$

Our objective is to design a control law so that the effect of the disturbance force F_2 on the positions of the two masses x_1 and x_2 are reduced in a frequency range $0 \le \omega \le 2$. The problem can be set up as shown in Fig. 6, where

$$W_e = \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix}$$

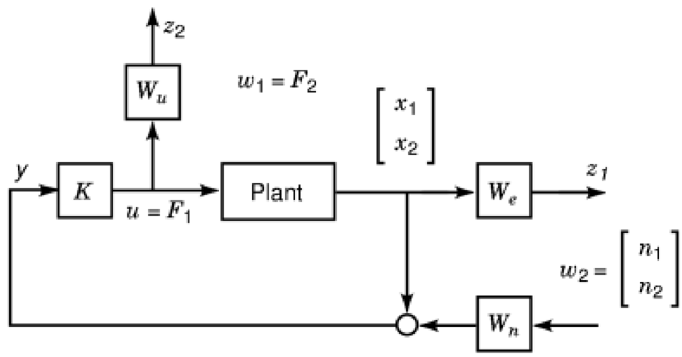

Fig. 6. Rejecting the disturbance force F_2 of the two mass/spring/damper system by a feedback control of F_1.

is the performance weight and W_u is the control weight. In order to limit the control force, we shall choose

$$W_u = \frac{s+5}{s+50}$$

Now let $u = F_1$,

$$w = \begin{bmatrix} F_2 \\ n_1 \\ n_2 \end{bmatrix}$$

then the problem can be formulated in a LFT form with

$$G(s) = \begin{bmatrix} \begin{bmatrix} W_e P_1 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} W_e P_2 \\ W_u \end{bmatrix} \\ [P_1 \quad W_n] & P_2 \end{bmatrix}$$

where P_1 and P_2 denote the transfer matrices from F_1 and F_2 to

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

respectively. Let

$$W_1 = \frac{5}{s/2+1}, \quad W_2 = 0$$

that is, we want to reject only the effect of the disturbance force F_2 on the position x_1. Then the optimal $\mathscr{H}_2$ performance is $\|\mathscr{F}_\ell(G, K_2)\|_2 = 2.6584$, and the $\mathscr{H}_\infty$ performance with the optimal $\mathscr{H}_2$ controller is $\|\mathscr{F}_\ell(G, K_2)\|_\infty = 2.6079$, whereas the optimal $\mathscr{H}_\infty$ performance with an $\mathscr{H}_\infty$ controller is $\|\mathscr{F}_\ell(G, K_\infty)\|_\infty = 1.6101$. This means

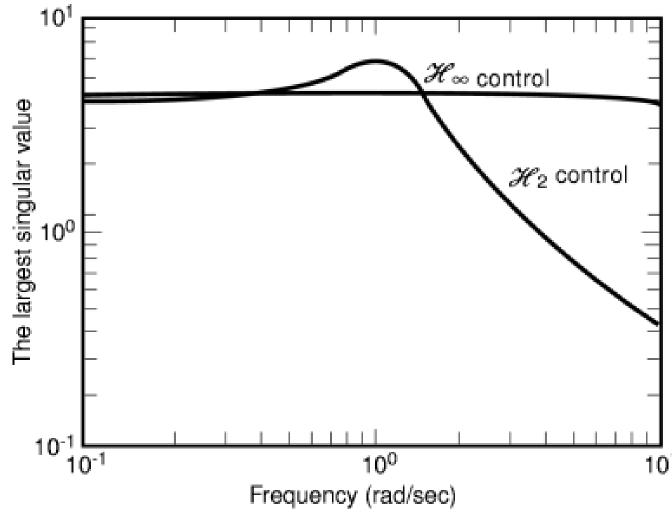

Fig. 7. The largest singular value Bode plot of the closed-loop two mass/spring/damper system with an $\mathscr{H}_2$ controller and an $\mathscr{H}_\infty$ controller.

that the effect of the disturbance force F_2 in the desired frequency range $0 \le \omega \le 2$ will be effectively reduced with the $\mathscr{H}_\infty$ controller K_∞ by 5/1.6101 = 3.1054 times at x_1. On the other hand, let

$$W_1 = 0, \quad W_2 = \frac{5}{s/2+1}$$

that is, we want to reject only the effect of the disturbance force F_2 on the position x_2. Then the optimal $\mathscr{H}_2$ performance is $\|\mathscr{F}_\ell(G, K_2)\|_2 = 0.1659$, and the $\mathscr{H}_\infty$ performance with the optimal $\mathscr{H}_2$ controller is $\|\mathscr{F}_\ell(G, K_2)\|_\infty = 0.5202$, whereas the optimal $\mathscr{H}_\infty$ performance with an $\mathscr{H}_\infty$ controller is $\|\mathscr{F}_\ell(G, K_\infty)\|_\infty = 0.5189$. This means that the effect of the disturbance force F_2 in the desired frequency range $0 \le \omega \le 2$ will be effectively reduced with the $\mathscr{H}_\infty$ controller K_∞ by 5/0.5189 = 9.6358 times at x_2. Finally, set

$$W_1 = W_2 = \frac{5}{s/2+1}$$

that is, we want to reject the effect of the disturbance force F_2 on both x_1 and x_2. Then the optimal $\mathscr{H}_2$ performance is $\|\mathscr{F}_\ell(G, K_2)\|_2 = 4.087$, and the $\mathscr{H}_\infty$ performance with the optimal $\mathscr{H}_2$ controller is $\|\mathscr{F}_\ell(G, K_2)\|_\infty = 6.0921$, whereas the optimal $\mathscr{H}_\infty$ performance with an $\mathscr{H}_\infty$ controller is $\|\mathscr{F}_\ell(G, K_\infty)\|_\infty = 4.3611$. This means that the effect of the disturbance force F_2 in the desired frequency range $0 \le \omega \le 2$ will be effectively reduced only with the $\mathscr{H}_\infty$ controller K_∞ by 5/4.3611 = 1.1465 times at both x_1 and x_2. This result shows clearly that it is very hard to reject the disturbance effect on both positions at the same time. The largest singular value Bode plots of the closed-loop system are shown in Fig. 7. We note that the $\mathscr{H}_\infty$ controller typically gives a relatively flat frequency response because it tries to minimize the peak of the frequency response. On the other hand, the $\mathscr{H}_2$ controller typically produces a frequency response that rolls off fast in the high-frequency range but with a large peak in the low-frequency range.

$\mathscr{H}_\infty$ Filtering

In this section we shall illustrate how an $\mathscr{H}_\infty$ filtering problem can be converted to a special $\mathscr{H}_\infty$ control problem. Suppose that a dynamic system is described by the following equations:

$$\begin{aligned}\dot{x} &= Ax + B_1 w, \quad x(0) = 0\\ z &= C_1 x\\ y &= C_2 x + D_{21} w\end{aligned}$$

The filtering problem is to find an estimate $\hat{z}$ of z in some sense using the measurement of y. The restriction on the filtering problem is that the filter has to be causal so that it can be realized (i.e., $\hat{z}$ must be generated by a causal system acting on the measurements). We will further restrict our filter to be *unbiased,* that is, given $T > 0$, the estimate $\hat{z}(t) = 0$ $\forall t \in [0, T]$ if $y(t) = 0$, $\forall t \in [0, T]$. Now we can state our $\mathscr{H}_\infty$ filtering problem.

$\mathscr{H}_\infty$ Filtering:

Given a $\gamma > 0$, find a causal filter $F(s) \in \mathscr{R}\mathscr{H}_\infty$ if it exists such that

$$J := \sup_{w \in \mathscr{L}_2[0,\infty)} \frac{\|z - \hat{z}\|_2}{\|w\|_2} < \gamma$$

with $\hat{z} = F(s)y$.

This $\mathscr{H}_\infty$ filtering problem can also be formulated in an LFT framework because

$$z - \hat{z} = \mathscr{F}_\ell(G, F)w, \quad G(s) = \left[\begin{array}{c|cc} A & B_1 & 0 \\ \hline C_1 & 0 & -I \\ C_2 & D_{21} & 0 \end{array}\right]$$

Hence, the filtering problem can be regarded as a special $\mathscr{H}_\infty$ problem. However, comparing this filtering problem to the control problems, we can see that there is no internal stability requirement in the filtering problem. Hence, the solution to this filtering problem can be obtained from the $\mathscr{H}_\infty$ solution in the last section by setting $B_2 = 0$ and dropping the internal stability requirement. Thus, a rational causal filter $F(s)$ is given by

$$\hat{z} = F(s)y = \left[\begin{array}{c|c} A + L_\infty C_2 & -L_\infty \\ \hline C_1 & 0 \end{array}\right] y$$

Understanding $\mathscr{H}_\infty$ Control

Most existing derivations and proofs of the $\mathscr{H}_\infty$ control results given in Theorem 2 are mathematically quite complex. Some algebraic derivations are simple but they provide no insight to the theory for control engineers. In this section, we shall present an intuitive but nonrigorous derivation of the $\mathscr{H}_\infty$ results by using only some basic system theoretic concept such as state feedback and state estimation. In fact, we shall construct intuitively the output feedback $\mathscr{H}_\infty$ central controller by combining an $\mathscr{H}_\infty$ state feedback and an observer.

A key fact we shall use is the so-called bounded real lemma, which states that for a system $z = G(s)w$ with state space realization $G(s) = C(sI - A)^{-1} B \in \mathscr{H}_\infty$, $\|G\|_\infty < \gamma$, which is essentially equivalent to

$$\int_0^\infty \left(\|z\|^2 - \gamma^2\|w\|^2\right) dt < 0, \quad \forall w \neq 0$$

if and only if there is an $X = X' \geq 0$ such that

$$XA + A'X + XBB'X/\gamma^2 + C'C = 0$$

and $A + BB'X/\gamma^2$ is stable. Dually, there is a $Y = Y' \geq 0$ such that

$$YA' + AY + YC'CY/\gamma^2 + BB' = 0$$

and $A + YC'C/\gamma^2$ is stable.

Note that the system has the following state space realization:

$$\dot{x} = Ax + B_1 w + B_2 u, \quad z = C_1 x + D_{12} u, \quad y = C_2 x + D_{21} w$$

To keep the presentation simple, we shall make some additional assumptions: $D'_{12}C_1 = 0$, $B_1D'_{21} = 0$, $D'_{12}D_{12} = I$, and $D_{21}D'_{21} = I$.

We shall first consider state feedback $u = Fx$. Then the closed-loop system becomes

$$\dot{x} = (A + B_2F)x + B_1 w, \quad z = (C_1 + D_{12}F)x$$

By the bounded real lemma, $\|T_{zw}\|_\infty < \gamma$ implies that there exists an $X = X' \geq 0$ such that

$$X(A + B_2F) + (A + B_2F)'X + XB_1B_1'X/\gamma^2 + (C_1 + D_{12}F)'(C_1 + D_{12}F) = 0$$

which is equivalent, by completing the square with respect to F, to

$$XA + A'X + XB_1B_1'X/\gamma^2 - XB_2B_2'X + C_1'C_1 + (F + B_2'X)'(F + B_2'X) = 0$$

Intuition suggests that we can take

$$F = -B_2'X$$

which gives

$$XA + A'X + XB_1B_1'X/\gamma^2 - XB_2B_2'X + C_1'C_1 = 0$$

This is exactly the X_∞ Riccati equation under the preceding simplified conditions. Hence, we can take $F = F_\infty$ and $X = X_\infty$.

Next, suppose that there is an output feedback stabilizing controller such that $\|T_{zw}\|_\infty < \gamma$. Then $x(\infty) = 0$ because the closed-loop system is stable. Consequently, we have

$$\begin{aligned}\int_0^\infty \left(\|z\|^2 - \gamma^2\|w\|^2\right) dt &= \int_0^\infty \left(\|z\|^2 - \gamma^2\|w\|^2 + \frac{d}{dt}(x'X_\infty x)\right) dt \\ &= \int_0^\infty \left(\|z\|^2 - \gamma^2\|w\|^2 + \dot{x}'X_\infty x + x'X_\infty \dot{x}\right) dt\end{aligned}$$

Substituting $\dot{x} = Ax + B_1w + B_2u$ and $z = C_1x + D_{12}u$ into the above integral and using the X_∞ equation, and finally completing the squares with respect to u and w, we get

$$\int_0^\infty \left(\|z\|^2 - \gamma^2\|w\|^2\right) dt = \int_0^\infty \left(\|v\|^2 - \gamma^2\|r\|^2\right) dt$$

where $v = u + B'_2X_\infty x = u - F_\infty x$ and $r = w - \gamma^{-2} B'_1X_\infty x$. Substituting w into the system equations, we have the new system equations

$$\begin{aligned}\dot{x} &= (A + B_1B_1'X_\infty/\gamma^2)x + B_1r + B_2u \\ v &= -F_\infty x + u \\ y &= C_2x + D_{21}r\end{aligned}$$

Hence the original $\mathscr{H}_\infty$ control problem is equivalent to finding a controller so that $\|T_{vr}\|_\infty < \gamma$ or

$$\int_0^\infty \left(\|u - F_\infty x\|^2 - \gamma^2\|r\|^2\right) dt < 0$$

Obviously, this also suggests intuitively that the state feedback control can be $u = F_\infty x$ and a worst state feedback disturbance would be $w_* = \gamma^{-2} B'_1X_\infty x$. Since full state is not available for feedback, we have to implement the control law using estimated state:

$$u = F_\infty \hat{x}$$

where $\hat{x}$ is the estimate of x. A standard observer can be constructed from the new system equations as

$$\dot{\hat{x}} = (A + B_1B_1'X_\infty/\gamma^2)\hat{x} + B_2u + L(C_2\hat{x} - y)$$

where L is the observer gain to be determined. Let $e := x - \hat{x}$ Then

$$\begin{aligned}\dot{e} &= (A + B_1B_1'X_\infty/\gamma^2 + LC_2)e + (B_1 + LD_{21})r \\ v &= -F_\infty e\end{aligned}$$

Since it is assumed that $\|T_{vr}\|_\infty < \gamma$, it follows from the dual version of the bounded real lemma that there exists a $Y \geq 0$ such that

$$\begin{aligned} &Y(A + B_1B_1'X_\infty/\gamma^2 + LC_2)' + (A + B_1B_1'X_\infty/\gamma^2 + LC_2)Y \\ &\qquad + YF_\infty'F_\infty Y/\gamma^2 + (B_1 + LD_{21})(B_1 + LD_{21})' = 0 \end{aligned}$$

The above equation can be written as

$$\begin{aligned} &(A + B_1B_1'X_\infty/\gamma^2)' + (A + B_1B_1'X_\infty/\gamma^2)Y + YF_\infty'F_\infty Y/\gamma^2 \\ &\qquad + B_1B_1' - YC_2'C_2Y + (L + YC_2')(L + YC_2')' = 0 \end{aligned}$$

Again, intuition suggests that we can take

$$L = -YC_2'$$

which gives

$$\begin{aligned} &Y(A + B_1B_1'X_\infty/\gamma^2)' + (A + B_1B_1'X_\infty/\gamma^2)Y \\ &\qquad + YF_\infty'F_\infty Y/\gamma^2 - YC_2'C_2Y + B_1B_1' = 0 \end{aligned}$$

It is easy to verify that

$$Y = Y_\infty(I - \gamma^{-2}X_\infty Y_\infty)^{-1}$$

where Y_∞ is as given in Theorem 2. Since $Y \geq 0$, we must have

$$\rho(X_\infty Y_\infty) < \gamma^2$$

Hence the controller is given by

$$\begin{aligned} \dot{\hat{x}} &= (A + B_1B_1'X_\infty/\gamma^2)\hat{x} + B_2u + L(C_2\hat{x} - y) \\ u &= F_\infty\hat{x} \end{aligned}$$

which is exactly the $\mathscr{H}_\infty$ central controller given in Theorem 2 under the simplified conditions.

We can see that the $\mathscr{H}_\infty$ central controller can be obtained by connecting a state feedback with a state estimate under the worst state feedback disturbance.

$\mathscr{H}_\infty$ Loop Shaping

Consider the family of uncertain systems $\mathbf{F_p}$ again. It is now clear that finding a controller K such that it robustly stabilizes the family $\mathbf{F_p}$ is a standard $\mathscr{H}_\infty$ norm minimization problem. Now suppose P has a stabilizable and detectable state-space realization given by

$$P = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$$

and let $Y \geq 0$ be the solution to

$$AY + YA^T - YC^TCY + BB^T = 0$$

Then the left coprime factorization $(\tilde{M}, \tilde{N})$ given by

$$[\tilde{N} \quad \tilde{M}] = \left[\begin{array}{c|cc} A - YC^TC & B & -YC^T \\ \hline C & 0 & I \end{array}\right]$$

is a normalized left coprime factorization. Furthermore,

$$\left\| \begin{bmatrix} I \\ K \end{bmatrix} (I + PK)^{-1}\tilde{M}^{-1} \right\|_\infty = \left\| \begin{bmatrix} I \\ K \end{bmatrix} (I + PK)^{-1}[I \quad P] \right\|_\infty$$

Define

$$b_{P,K} := \begin{cases} \left(\left\| \begin{bmatrix} I \\ K \end{bmatrix} (I + PK)^{-1}[I \quad P] \right\|_\infty \right)^{-1}, & \text{if } K \text{ stabilizes } P \\ 0, & \text{otherwise} \end{cases}$$

Then $b_{P,K} > 0$ implies that K also stabilizes robustly the following family of uncertain systems:

(1) $\tilde{P} = P + \Delta_a$ such that P and $\tilde{P}$ have the same number of unstable poles and $\|\Delta_a\|_\infty < b_{P,K}$.
(2) $\tilde{P} = (I + \Delta_m)P$ such that P and $\tilde{P}$ have the same number of unstable poles and $\|\Delta_m\|_\infty < b_{P,K}$.
(3) $\tilde{P} = (I + \Delta_f)^{-1}P$ such that P and $\tilde{P}$ have the same number of unstable poles and $\|\Delta_f\|_\infty < b_{P,K}$.

These conclusions also hold when the roles of plant and controller are interchanged. The number $b_{P,K}$ can also be related to the classical gain and phase margins of a SISO system:

$$\text{Gain Margin} \geq \frac{1 + b_{P,K}}{1 - b_{P,K}}, \quad \text{Phase Margin} \geq 2\arcsin(b_{P,K})$$

Hence, $b_{P,K}$ is a good measure of a system's robustness. Define

$$b_{\text{opt}} := \sup_K b_{P,K}$$

Obviously, b_{opt} is the largest admissible size of perturbation ϵ so that the system is stable. It follows from the standard $\mathscr{H}_\infty$ solution that

$$b_{\text{opt}}(P) = \frac{1}{\sqrt{1 - \lambda_{\max}(YQ)}}$$

where Q is the solution to the following Lyapunov equation

$$Q(A - YC^TC) + (A - YC^TC)^TQ + C^TC = 0$$

Moreover, for any $\gamma > 1/b_{\text{opt}}(P)$, a controller achieving $b_{P,K} > 1/\gamma$ is given by

$$K(s) = \left[\begin{array}{c|c} A - BB^TX_\infty - YC^TC & -YC^T \\ \hline -B^TX_\infty & 0 \end{array}\right]$$

where

$$X_\infty = \frac{\gamma^2}{\gamma^2 - 1} Q \left(I - \frac{\gamma^2}{\gamma^2 - 1} YQ \right)^{-1}$$

This stabilization solution can be used to devise an $\mathscr{H}_\infty$ loop-sharing design method. The objective of this approach is to incorporate the simple performance/robustness trade-off obtained in the loop shaping with the guaranteed stability properties of $\mathscr{H}_\infty$ design methods. Recall that good performance controller design requires that

$$\overline{\sigma}((I + PK)^{-1}) \ll 1 \Leftrightarrow \underline{\sigma}(PK) \gg 1$$

particularly in some low-frequency range where $\sigma\_(PK)$ denotes the smallest singular value. And good robustness requires that

$$\overline{\sigma}(PK(I + PK)^{-1}) \gg 1 \Leftrightarrow \overline{\sigma}(PK) \ll 1$$

particularly in some high-frequency range.

The $\mathscr{H}_\infty$ loop-shaping design procedure is developed by McFarlane and Glover (4) and is stated in the next section.

Loop-Shaping Design Procedure.

(1) Loop Shaping: The singular values of the nominal plant are shaped, using a precompensator W_1 and/or a postcompensator W_2, to give a desired open-loop shape. The nominal plant P and the shaping functions W_1, W_2 are combined to form the shaped plant P_s, where $P_s = W_2PW_1$. We assume that W_1 and W_2 are such that P_s contains no hidden modes.

(2) Robust Stabilization: (a) If $b_{\text{opt}}(P) \ll 1$ return to (1) and adjust W_1 and W_2. (b) Select $\epsilon \leq b_{\text{opt}}(P_s)$; then synthesize a stabilizing controller K_∞, which satisfies $b_{Ps,K\infty} \geq \epsilon$.

(3) The final feedback controller K is then constructed by combining the $\mathscr{H}_\infty$ controller K_∞ with the shaping functions W_1 and W_2 such that $K = W_1K_\infty W_2$.

A typical design works as follows: the designer inspects the open-loop singular values of the nominal plant, and shapes these by pre- and/or postcompensation until nominal performance (and possibly robust stability) specifications are met. (Recall that the open-loop shape is related to closed-loop objectives.) A feedback controller K_∞ with associated stability margin (for the shaped plant) $\epsilon \le b_{\text{opt}}(P_s)$ is then synthesized. If $b_{\text{opt}}(P_s)$ is small, then the specified loop shape is incompatible with robust stability requirements, and should be adjusted accordingly; then K_∞ is reevaluated.

Note that in contrast to the classical loop-sharing approach, the loop shaping here is done without explicit regard for the nominal plant phase information. That is, closed-loop stability requirements are disregarded at this stage. Also in contrast with conventional $\mathscr{H}_\infty$ design, the robust stabilization is done without frequency weighting.

In fact, the preceding robust stabilization objective can also be interpreted as the more standard $\mathscr{H}_\infty$ problem formulation of minimizing the $\mathscr{H}_\infty$ norm of the frequency weighted gain from disturbances on the plant input and output to the controller input and output as follows:

$$\begin{aligned} b^{-1}_{P_s,K_\infty} &= \left\| \begin{bmatrix} I \\ K_\infty \end{bmatrix} (I + P_s K_\infty)^{-1} \tilde{M}_s^{-1} \right\|_\infty \\ &= \left\| \begin{bmatrix} I \\ K_\infty \end{bmatrix} (I + P_s K_\infty)^{-1} [I \quad P_s] \right\|_\infty \\ &= \left\| \begin{bmatrix} W_2 \\ W_1^{-1} K \end{bmatrix} (I + PK)^{-1} [W_2^{-1} \quad PW_1] \right\|_\infty \end{aligned}$$

This shows that the $\mathscr{H}_\infty$ loop-shaping design is equivalent to a standard $\mathscr{H}_\infty$ design with the shaping functions as weighting functions.

μ Synthesis

As we discussed at the beginning of this article, $\|\mathscr{F}_\ell(G, K)\|_\infty \le \gamma$ guarantees the robust stability of the uncertain system shown in Fig. 1(b) for any $\Delta(s) \in \mathscr{RH}_\infty$ with $\|\Delta\|_\infty < 1/\gamma$. However, if a system is built from components, which are themselves uncertain, then, in general, the uncertainty in the system level is structured, and this robust stability guarantee may be overly conservative. Because the interconnection model G can always be chosen so that $\Delta(s)$ is block diagonal, and, by absorbing any weights, $\|\Delta\|_\infty < 1$. Thus we can assume that $\Delta(s)$ takes the form of

$$\Delta(s) = \{\text{diag}[\delta_1 I_{r_1}, \ldots, \delta_s I_{r_S}, \Delta_1, \ldots, \Delta_F]: \quad \delta_i(s) \in \mathscr{RH}_\infty, \Delta_j \in \mathscr{RH}_\infty\}$$

with $\|\delta_i\|_\infty < 1$ and $\|\Delta_j\|_\infty < 1$. The robust stability analysis for systems with such *structured* uncertainty is not as simple but can be formally characterized by using the structured singular value, see Ref. 5.

Define $\Delta \subset \mathbb{C}^{n\times n}$ as

$$\Delta = \{\text{diag}[\delta_1 I_{r_1}, \ldots, \delta_s I_{r_s}, \Delta_1, \ldots, \Delta_F]: \delta_i \in \mathbb{C}, \Delta_j \in \mathbb{C}^{m_j \times m_j}\}$$

Then for $M \in \mathbb{C}^{n\times n}$, the structured singular value of M, $\mu_\Delta(M)$, is defined as

$$\mu_{\boldsymbol{\Delta}}(M) := \{\min\{\overline{\sigma}(\Delta) \colon \Delta \in \boldsymbol{\Delta},\ \det(I - M\Delta) = 0\}\}^{-1}$$

The μ itself is not easy to compute. But good bounds can be obtained efficiently. Let

$$\mathscr{D} = \left\{ \begin{array}{l} \operatorname{diag}[D_1, \ldots, D_S, d_1 I_{m_1}, \ldots, d_{F-1} I_{m_{F-1}}, I_{m_F}]\colon \\ \quad D_i \in \mathbb{C}^{r_i \times r_i}, D_i = D_i^T > 0,\ d_j \in \mathbb{R},\ d_j > 0 \end{array} \right\}$$

Then for any $\Delta \in \Delta$ and $D \in D$, $D\Delta = \Delta D$ and

$$\mu_{\boldsymbol{\Delta}}(M) \leq \inf_{D\in\mathscr{D}} \overline{\sigma}(DMD^{-1})$$

and the equality holds if $2S + F \leq 3$. This bound can be used frequency by frequency to determine the system robust stability and performance with structured uncertainties. For example, the system in Fig. 1(b) is well posed, internally stable for all stable $\Delta(s)$ with $\Delta(s_0) \in \Delta$, $\forall$ Re$(s_0) \geq 0$, and $\|\Delta\|_\infty < 1/\gamma$ if and only if

$$\sup_{\omega\in\mathbb{R}} \mu_{\boldsymbol{\Delta}}(\mathscr{F}_\ell(G(j\omega), K(j\omega))) \leq \gamma$$

This result leads us to the following synthesis problem:

$$\min_K \sup_{\omega\in\mathbb{R}} \mu_{\boldsymbol{\Delta}}(\mathscr{F}_\ell(G, K))$$

This synthesis problem is not yet fully solved in the general case. A reasonable approach is to obtain a solution to an upper bound:

$$\min_K \inf_{D(s), D^{-1}(s)\in\mathscr{H}_\infty} \|D(s)\mathscr{F}_\ell(G(s), K(s))D^{-1}(s)\|_\infty$$

by iteratively solving for K and D. This is the so-called D-K iteration. The stable and minimum phase scaling matrix $D(s)$ is chosen such that $D(s)\Delta(s) = \Delta(s)D(s)$. For a fixed scaling transfer matrix D, $\min_K \|D\mathscr{F}_\ell(G, K)D^{-1}\|_\infty$ is a standard $\mathscr{H}_\infty$ optimization problem because

$$D(s)\mathscr{F}_\ell(G(s), K(s))D^{-1}(s) = \mathscr{F}_\ell\left(\begin{bmatrix} D & 0 \\ 0 & I \end{bmatrix} G \begin{bmatrix} D^{-1} & 0 \\ 0 & I \end{bmatrix}, K\right)$$

For a given stabilizing controller K, $\inf_{D,D^{-1}\in\mathscr{H}_\infty} \|D\mathscr{F}_\ell(G, K)D^{-1}\|_\infty$ is a standard convex optimization problem, and it can be solved pointwise in the frequency domain:

$$\sup_\omega \inf_{D_\omega\in\mathscr{D}} \overline{\sigma}[D_\omega \mathscr{F}_\ell(G, K)(j\omega)D_\omega^{-1}]$$

Then a $D(s)$ is found to approximate the magnitude frequency response D_ω uniformly (usually by curve fitting). D-K iterations proceed by performing this two-parameter minimization in sequential fashion: minimizing over

K with $D(s)$ fixed, minimizing pointwise over D with K fixed, minimizing again over K, and then again over D, and so on. With either K or D fixed, the global optimum in the other variable may be found using the μ and $\mathscr{H}_\infty$ solutions. Although the joint optimization of D and K is generally not convex and the global convergence is not guaranteed, many designs have shown that this approach works very well. In fact, this is probably the most effective design methodology available today for dealing with such complicated problems.

Additional Applications

There are many additional extensions and development in the $\mathscr{H}_\infty$ control theory. Here are some of them:

$\mathscr{H}_\infty$ loop-shaping techniques using ν-gap metric, see Ref. 6.
Robust control design in the gap metric, see Robust control and Ref. 6.
Linear matrix inequality (*LMI*) approach to $\mathscr{H}_\infty$ control, see Convex optimization.
Time-varying and finite horizon $\mathscr{H}_\infty$ control and game theoretical approach to $\mathscr{H}_\infty$ control, see Refs. 7 and 8.
Operator theoretic approach to $\mathscr{H}_\infty$ control, see Ref. 9.
Chain-scattering approach to $\mathscr{H}_\infty$ control, see Ref. 10.
$\mathscr{H}_\infty$ control with pole placement, see Ref. 11.
$\mathscr{H}_\infty$ controller reduction, see Refs. 12 and 13.
Linear parameter varying $\mathscr{H}_\infty$ control, see Ref. 14.
Sampled-Data $\mathscr{H}_\infty$ control, see Ref. 15.
$\mathscr{H}_\infty$ control for infinite dimensional systems, see Refs. 16 and 17.
$\mathscr{H}_\infty$ control for nonlinear systems, see Ref. 18.
Software and applications, see Ref. 19.

A comprehensive treatment of $\mathscr{H}_\infty$ control theory can be found in Refs. 13 and 20.

BIBLIOGRAPHY

1. G. Zames, Feedback and optimal sensitivity: model reference transformations, multiplicative semi-norms, and approximate inverses, *IEEE Trans. Autom. Control*, **26**: 301–320, 1981.
2. K. Glover and J. Doyle, State-space formulae for all stabilizing controllers that satisfy an $\mathscr{H}_\infty$ norm bound and relations to risk sensitivity, *Systems and Control Letters*, **11**: 167–172, 1988.
3. J. C. Doyle *et al.*, State-space solutions to standard $\mathscr{H}_2$ and $\mathscr{H}_\infty$ control problems, *IEEE Trans. Autom. Control*, **34**(8): 831–847, 1989.
4. D.C. McFarlane and K. Glover, A loop shaping design procedure using $\mathscr{H}_\infty$ synthesis, *IEEE Trans. Autom. Control*, **37**(6): 759–769, 1992.
5. A. Packard and J. C. Doyle, The complex structured singular value, *Automatica*, **29**: 71–109, 1993.
6. G. Vinnicombe, Frequency domain uncertainty and the graph topology, *IEEE Trans. Autom. Control*, **38**(9): 1371–1383, 1993.
7. M. Green and D. J. N. Limebeer, *Linear Robust Control*, Englewood Cliffs, NJ: Prentice-Hall, 1995.
8. T. Başaer and P. Bernhard, *$\mathscr{H}_\infty$-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach,* Systems and Control: Foundations and Applications, Boston: Birkhäuser, 1991.
9. B. A. Francis, A Course in $\mathscr{H}_\infty$ Control Theory, *Lecture Notes in Control and Information Sciences*, Vol. **88**, London: Springer-Verlag, 1987.
10. H. Kimura, *Chain-Scattering Approach to $\mathscr{H}_\infty$-Control*, Boston: Birkhäuser, 1997.

11. M. Chilali and P. Gahinet, $\mathscr{H}_\infty$ design with pole placement constraints: An LMI approach, *IEEE Trans. Autom. Control*, **41**(3): 358–367, 1996.
12. P. J. Goddard and K. Glover, Performance preserving frequency weighted controller approximation: A coprime factorization approach, *Proceedings of the 33nd Conference on Decision and Control*, Orlando, FL, pp. 2720–2725, December 1994.
13. K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Upper Saddle River, NJ: Prentice-Hall, 1996.
14. A. Packard, Gain scheduling via linear fractional transformations, *Systems and Control Lett.*, **22**: 79–92, 1994.
15. T. Chen and B. A. Francis, *Optimal Sampled-Data Control Systems*, London: Springer-Verlag, 1995.
16. R. F. Curtain and H. J. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*, New York: Springer-Verlag, 1995.
17. C. Foias, H. Özbay, and A. Tannenbaum, Robust Control of Infinite Dimensional Systems, *Lecture Notes in Control and Information Sciences*, Vol. **209**, London: Springer-Verlag, 1996.
18. A. van der Schaft, L_2-Gain and Passivity Techniques in Nonlinear Control, *Lecture Notes in Control and Information Sciences*, Vol. **218**, London: Springer-Verlag, 1996.
19. G. Balas *et al.*, *μ-Analysis and Synthesis Toolbox*, Natick, MA: The MathWorks, Inc., 1995.
20. K. Zhou and J. C. Doyle, *Essentials of Robust Control*, Upper Saddle River, NJ: Prentice-Hall, 1998.

KEMIN ZHOU
Louisiana State University