1

Introduction

1.1 Goals and origins of $\mathcal{H}_\infty$ optimal control

Most engineering undergraduates are taught to design proportional-integral-derivative (PID) compensators using a variety of different frequency response techniques. With the help of a little laboratory experience, students soon realize that a typical design study involves juggling with conflicting design objectives such as the gain margin and the closed-loop bandwidth until an acceptable controller is found. In many cases these "classical" controller design techniques lead to a perfectly satisfactory solution and more powerful tools hardly seem necessary. Difficulties arise when the plant dynamics are complex and poorly modelled, or when the performance specifications are particularly stringent. Even if a solution is eventually found, the process is likely to be expensive in terms of design engineer's time.

When a design team is faced with one of these more difficult problems, and no solution seems forthcoming, there are two possible courses of action. These are either to compromise the specifications to make the design task easier, or to search for more powerful design tools. In the case of the first option, reduced performance is accepted without ever knowing if the original specifications could have been satisfied, as classical control design methods do not address existence questions. In the case of the second option, more powerful design tools can only help if a solution exists.

Any progress with questions concerning achievable performance limits and the existence of satisfactory controllers is bound to involve some kind of optimization theory. If, for example, it were possible to optimize the settings of a PID regulator, the design problem would either be solved or it would become apparent that the specifications are impossible to satisfy (with a PID regulator). We believe that answering existence questions is an important component of a good design methodology. One does not want to waste time trying to solve a problem that has no solution, nor does one want to accept specification compromises without knowing that these are necessary. A further benefit of optimization is that it provides an absolute scale of merit against which any design can be measured—if a design is already all but perfect, there is little point in trying to improve it further.

The aim of this book is to develop a theoretical framework within which one may address complex design problems with demanding specifications in a systematic way.

Wiener-Hopf-Kalman optimal control

The first successes with control system optimization came in the 1950s with the introduction of the Wiener-Hopf-Kalman (WHK) theory of optimal control.[1] At roughly the same time the United States and the Soviet Union were funding a massive research program into the guidance and maneuvering of space vehicles. As it turned out, the then new optimal control theory was well suited to many of the control problems that arose from the space program. There were two main reasons for this:

1. The underlying assumptions of the WHK theory are that the plant has a known linear (and possibly time-varying) description, and that the exogenous noises and disturbances impinging on the feedback system are stochastic in nature, but have known statistical properties. Since space vehicles have dynamics that are essentially ballistic in character, it is possible to develop accurate mathematical models of their behavior. In addition, descriptions for external disturbances based on white noise are often appropriate in aerospace applications. Therefore, at least from a modelling point of view, the WHK theory and these applications are well suited to each other.

2. Many of the control problems from the space program are concerned with resource management. In the 1960s, aerospace engineers were interested in minimum fuel consumption problems such as minimizing the use of retrorockets. One famous problem of this type was concerned with landing the lunar excursion module with a minimum expenditure of fuel. Performance criteria of this type are easily embedded in the WHK framework that was specially developed to minimize quadratic performance indices.

Another revolutionary feature of the WHK theory is that it offers a true synthesis procedure. Once the designer has settled on a quadratic performance index to be minimized, the WHK procedure supplies the (unique) optimal controller without any further intervention from the designer. In the euphoria that followed the introduction of optimal control theory, it was widely believed that the control system

[1] Linear Quadratic Gaussian (LQG) optimal control is the term now most widely used for this type of optimal control.

designer had finally been relieved of the burdensome task of designing by trial and error. As is well known, the reality turned out to be quite different.

The wide-spread success of the WHK theory in aerospace applications soon led to attempts to apply optimal control theory to more mundane industrial problems. In contrast to experience with aerospace applications, it soon became apparent that there was a serious mismatch between the underlying assumptions of the WHK theory and industrial control problems. Accurate models are not routinely available and most industrial plant engineers have no idea as to the statistical nature of the external disturbances impinging on their plant. After a ten year re-appraisal of the status of multivariable control theory, it became clear that an optimal control theory that deals with the question of plant modelling errors and external disturbance uncertainty was required.

Worst-case control and $\mathcal{H}_\infty$ optimization

$\mathcal{H}_\infty$ optimal control is a frequency-domain optimization and synthesis theory that was developed in response to the need for a synthesis procedure that *explicitly* addresses questions of modelling errors. The basic philosophy is to treat the worst case scenario: if you don't know what you are up against, plan for the worst and optimize. For such a framework to be useful, it must have the following properties:

1. It must be capable of dealing with plant modelling errors and unknown disturbances.
2. It should represent a natural extension to existing feedback theory, as this will facilitate an easy transfer of intuition from the classical setting.
3. It must be amenable to meaningful optimization.
4. It must be able to deal with multivariable problems.

In this chapter, we will introduce the infinity norm and $\mathcal{H}_\infty$ optimal control with the aid of a sequence of simple single-loop examples. We have carefully selected these in order to minimize the amount of background mathematics required of the reader in these early stages of study; all that is required is a familiarity with the *maximum modulus principle*. Roughly speaking, this principle says that if a function f (of a complex variable) is analytic inside and on the boundary of some domain $\mathcal{D}$, then the maximum modulus (magnitude) of the function f occurs on the boundary of the domain $\mathcal{D}$. For example, if a feedback system is closed-loop stable, the maximum of the modulus of the closed-loop transfer function over the closed right-half of the complex plane will always occur on the imaginary axis.

To motivate the introduction of the infinity norm, we consider the question of robust stability optimization for the feedback system shown in Figure 1.1. The transfer function g represents a nominal linear, time-invariant model of an open-loop system and the transfer function k represents a linear, time-invariant controller to be designed. If the "true" system is represented by $(1+\delta)g$, we say that the modelling

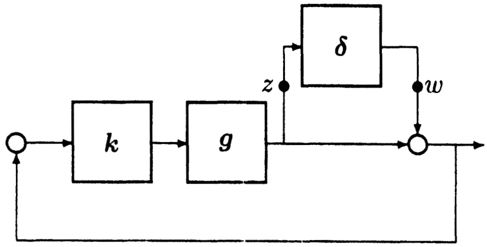

Figure 1.1: The problem of robust stability optimization.

error is represented by a multiplicative perturbation $\boldsymbol{\delta}$ at the plant output. For this introductory analysis, we assume that $\boldsymbol{\delta}$ is an unknown linear, time-invariant system.

Since

$$z = (1 - \boldsymbol{gk})^{-1}\boldsymbol{gk}w,$$

the stability properties of the system given in Figure 1.1 are the same as those given in Figure 1.2, in which

$$\boldsymbol{h} = (1 - \boldsymbol{gk})^{-1}\boldsymbol{gk}.$$

If the perturbation $\boldsymbol{\delta}$ and the nominal closed-loop system given by $\boldsymbol{h}$ are both

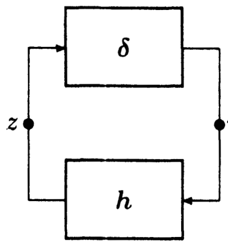

Figure 1.2: The small gain problem.

stable, the Nyquist criterion says that the closed-loop system is stable if and only if the Nyquist diagram of $\boldsymbol{h\delta}$ does not encircle the +1 point. We use the +1 point rather than the −1 point because of our positive feedback sign convention. Since the condition

$$\sup_{\omega} |\boldsymbol{h}(j\omega)\boldsymbol{\delta}(j\omega)| < 1. \tag{1.1.1}$$

ensures that the Nyquist diagram of $\boldsymbol{h\delta}$ does not encircle the +1 point, we conclude that the closed-loop system is stable provided (1.1.1) holds.

Since δ is unknown, it makes sense to replace (1.1.1) with an alternative sufficient condition for stability in which h and δ are separated. We could for example test the condition

$$\sup_\omega |h(j\omega)| \sup_\omega |\delta(j\omega)| < 1.$$

If δ is stable and bounded in magnitude, so that

$$\sup_\omega |\delta(j\omega)| = M,$$

the feedback loop given in Figure 1.1 will be stable provided a stabilizing controller can be found such that

$$\sup_\omega |h(j\omega)| < \frac{1}{M}.$$

The quantity $\sup_\omega |h(j\omega)|$ satisfies the axioms of a norm, and is known as the *infinity norm.* Specifically,

$$\|h\|_\infty = \sup_\omega |h(j\omega)|.$$

Electrical engineers will immediately recognize $\|h\|_\infty$ as the highest gain value on a Bode magnitude plot. The quantity $\|\cdot\|_\infty$ is a norm, since it satisfies the following axioms:

1. $\|h\|_\infty \geq 0$ with $\|h\|_\infty = 0$ if and only if $h = 0$.
2. $\|\alpha h\|_\infty = |\alpha| \|h\|_\infty$ for all scalars α.
3. $\|h + g\|_\infty \leq \|h\|_\infty + \|g\|_\infty$.

In addition, $\|\cdot\|_\infty$ satisfies

4. $\|hg\|_\infty \leq \|h\|_\infty \|g\|_\infty$.

The fourth property is the crucial submultiplicative property which is central to all the robust stability and robust performance work to be encountered in this book. Note that not all norms have this fourth property.

With this background, the optimal robust stability problem is posed as one of finding a stabilizing controller k that minimizes $\|(1-gk)^{-1}gk\|_\infty$. Note that $k = 0$ gives $\|(1-gk)^{-1}gk\|_\infty = 0$ and is therefore optimal in this sense provided the plant itself is stable. Thus, when the plant is stable and there are no performance requirements other than stability, the optimal course of action is to use no feedback at all! When $k = 0$ is not allowed because the plant is unstable, the problem is more interesting and the optimal stability margin and the optimal controller are much harder to find. We will return to the analysis of this type of problem in Section 1.4.

In order to lay the groundwork for our analysis of optimal disturbance attenuation and optimal stability robustness, we consider the optimal command response problem. This problem is particularly simple because it contains no feedback. Despite this, it contains many of the essential mathematical features of more difficult (feedback) problems.

1.2 Optimizing the command response

As an introduction to the use of the infinity norm in control system optimization, we analyze the design of reference signal prefilters in command tracking applications. This is our first example of an $\mathcal{H}_\infty$ optimal controller synthesis problem.

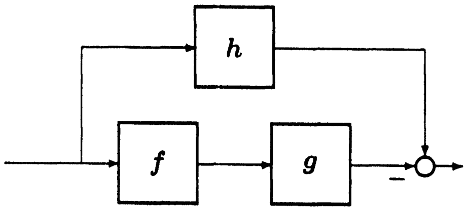

Figure 1.3: Command response optimization.

In the configuration illustrated in Figure 1.3, we suppose that the plant model g is a given stable rational transfer function and that h is a given stable rational transfer function with desired command response properties. The design task is to find a stable rational prefilter with transfer function f such that $\|h - gf\|_\infty$ is minimized. An unstable prefilter is unacceptable in practical applications because it results in unbounded control signals and actuator saturation.

In the case that g has no zeros in the closed-right-half plane, the solution is easy since we may simply set $f = g^{-1}h$. If g has right-half-plane zeros, however, the plant inverse leads to an unstable prefilter unless the right-half-plane poles of g^{-1} happen to be cancelled by zeros of h. Thus, when g has right-half-plane zeros, the requirement that the prefilter be stable forces us to accept some error between gf and h which we denote

$$e = h - gf. \tag{1.2.1}$$

This gives

$$f = g^{-1}(h - e). \tag{1.2.2}$$

If the right-half-plane zeros of g are $z_1, z_2, \ldots, z_m$ and are of multiplicity one, the prefilter will be stable if and only if

$$e(z_i) = h(z_i), \qquad i = 1, 2, \ldots, m. \tag{1.2.3}$$

This is because the unstable poles of g^{-1} will be cancelled by the zeros of $h - e$.

The conditions given in (1.2.3) are called *interpolation constraints.* Any error system e resulting from a stable prefilter must satisfy the conditions (1.2.3) and, conversely, the satisfaction of these constraints ensures that all the right-half-plane poles of g^{-1} will be cancelled by zeros of $h - e$ when forming the prefilter. The optimization problem is to find a stable transfer function e of minimum infinity norm such that the interpolation constraints given in (1.2.3) are satisfied. This is an example of a *Nevanlinna-Pick interpolation* problem. A general solution to problems of this type is complicated and was found early this century. Once the optimal error function is found, f follows by back substitution using (1.2.2). We shall now consolidate these ideas with a numerical example.

Example 1.2.1. Suppose g and h are given by

$$g = \left(\frac{s-1}{s+2}\right), \qquad h = \left(\frac{s+1}{s+3}\right).$$

The transfer function g has a single zero at $s = 1$, so there is a single interpolation constraint given by

$$e(1) = \left(\frac{s+1}{s+3}\right)\bigg|_{s=1} = \frac{1}{2}.$$

Since e is required to be stable, the maximum modulus principle ensures that

$$\begin{aligned} \|e\|_\infty &= \sup_{s=j\omega} |e(s)| \\ &= \sup_{\mathrm{R}_e(s)\geq 0} |e(s)| \\ &\geq |e(1)| = \frac{1}{2}. \end{aligned}$$

The minimum infinity norm interpolating function is therefore the constant function $e = \frac{1}{2}$ and the associated norm is $\|e\|_\infty = \frac{1}{2}$. Back substitution using (1.2.2) yields

$$f = \left(\frac{s+2}{s-1}\right)\left(\frac{s+1}{s+3} - \frac{1}{2}\right) = \frac{1}{2}\left(\frac{s+2}{s+3}\right). \qquad \nabla$$

Interpolating a single data point is particularly simple because the optimal interpolating function is a constant. Our next example, which contains two interpolation constraints, shows that the general interpolation problem is far more complex.

Example 1.2.2. Consider the command response optimization problem in which

$$g = \frac{(s-1)(s-2)}{(s+3)^2}, \qquad h = \frac{2}{3(s+3)}.$$

The transfer function g has right-half-plane zeros at $z_1 = 1$ and $z_2 = 2$, so we must find a stable transfer function e of minimum norm such that:

$$e(1) = h(1) = \frac{1}{6} = h_1 \tag{1.2.4}$$

and

$$e(2) = h(2) = \frac{2}{15} = h_2. \tag{1.2.5}$$

It follows from the maximum modulus principle that any such e must satisfy

$$\|e\|_\infty \geq \max\left\{\frac{1}{6}, \frac{2}{15}\right\} = \frac{1}{6}.$$

Since we have two values to interpolate, simply setting $e = \frac{1}{6}$ will not do!

The Nevanlinna-Pick interpolation theory says that there is a stable interpolating function e with $\|e\|_\infty \leq \gamma$ if and only if the *Pick matrix* given by

$$\Pi(\gamma) = \begin{bmatrix} \frac{\gamma^2 - h_1^2}{2} & \frac{\gamma^2 - h_1 h_2}{3} \\ \frac{\gamma^2 - h_1 h_2}{3} & \frac{\gamma^2 - h_2^2}{4} \end{bmatrix}$$

is nonnegative definite. Since $\Pi(\gamma_1) \geq \Pi(\gamma_2)$ if $\gamma_1 \geq \gamma_2$, our desired optimal norm is the largest value of γ for which the Pick matrix $\Pi(\gamma)$ is singular. Alternatively, the optimal value of γ (call it γ_{opt}) is the square root of the largest eigenvalue of the symmetric matrix pencil

$$\gamma^2 \begin{bmatrix} \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{4} \end{bmatrix} - \begin{bmatrix} \frac{h_1^2}{2} & \frac{h_1 h_2}{3} \\ \frac{h_1 h_2}{3} & \frac{h_2^2}{4} \end{bmatrix}.$$

Carrying out this calculation gives $\gamma_{opt} \approx 0.207233$. The Nevanlinna-Pick theory also gives the optimal interpolating function as

$$e = \gamma_{opt}\left(\frac{a-s}{a+s}\right),$$

with a given by

$$\begin{aligned} a &= z_i \frac{\gamma_{opt} + h_i}{\gamma_{opt} - h_i} \quad \text{(in which } i \text{ is either 1 or 2)} \\ &\approx 9.21699. \end{aligned}$$

(It is easy to check that this e satisfies the interpolation constraints.) Notice that the optimal interpolating function is a constant multiplied by a stable transfer function with unit magnitude on the imaginary axis, which is a general property of optimal interpolating functions. Since $\|\frac{a-s}{a+s}\|_\infty = 1$, it is clear that $\|e\|_\infty = \gamma_{opt}$. Since $f = g^{-1}(h - e)$, it follows that the optimal prefilter is

$$f = \gamma_{opt}\left(\frac{s+3}{s+a}\right).$$

∇

We conclude from this example that an increase in the number of interpolation constraints makes the evaluation of the interpolating function much harder. Despite this, the error function retains the "constant magnitude on the imaginary axis" property associated with constants. We will not address (or require) a general solution to the Nevanlinna-Pick interpolation problem, although the solution to the $\mathcal{H}_\infty$ optimal control problem we shall develop also provides a solution to the Nevanlinna-Pick interpolation problem. We shall say more about this in Chapter 6.

1.3 Optimal disturbance attenuation

The aim of this section is to solve a simple $\mathcal{H}_\infty$ control problem involving feedback by recasting the optimal disturbance attenuation problem as an optimization problem constrained by interpolation conditions.

In the system illustrated in Figure 1.4, it is assumed that the plant model $\boldsymbol{g}$ is a given stable rational transfer function and that the frequency domain signal d represents some *unknown* disturbance. The aim is to find a compensator $\boldsymbol{k}$ with the following two properties:

1. It must stabilize the loop in a sense to be specified below.

2. It must minimize the infinity norm of the transfer function that maps d to y.

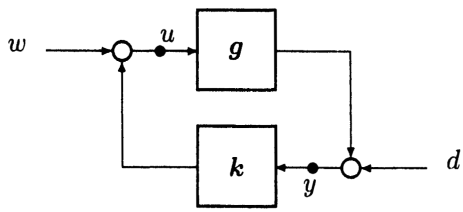

Figure 1.4: The disturbance attenuation problem.

If $w = 0$, it is immediate from Figure 1.4 that

$$\begin{aligned} y &= (1 - \boldsymbol{gk})^{-1} d \\ &= (1 + \boldsymbol{gk}(1 - \boldsymbol{gk})^{-1}) d, \end{aligned}$$

and we note that the closed-loop transfer function is a nonlinear function of $\boldsymbol{k}$. To restore an affine parametrization of the type given in (1.2.1), we set

$$\boldsymbol{q} = \boldsymbol{k}(1 - \boldsymbol{gk})^{-1}, \tag{1.3.1}$$

which is the transfer function between the disturbance d and the plant input u. The closed-loop mapping d to y may now be written as

$$y = (1 + \boldsymbol{gq}) d, \tag{1.3.2}$$

which is affine in the unknown parameter $\boldsymbol{q}$. Before continuing, we need to introduce the notion of internal stability and discover the properties required of $\boldsymbol{q}$ in order that the resulting controller be internally stabilizing.

1.3.1 Internal stability theory for stable plants

Definition 1.3.1 *The feedback system given in Figure 1.4 is called internally stable if each of the four transfer functions mapping w and d to u and y are stable.*

If the feedback system in Figure 1.4 is internally stable, we say that $\boldsymbol{k}$ is an internally-stabilizing controller for $\boldsymbol{g}$.[2]

Internal stability is a more stringent stability requirement than the simple input-output stability of closed-loop transfer functions, because it also bans all right-half-plane pole-zero cancellations between cascaded subsystems within the feedback loop.

Example 1.3.1. The transfer functions $\boldsymbol{g} = \left(\frac{-s}{s+1}\right)$ and $\boldsymbol{k} = \left(\frac{s+3}{s}\right)$ produce the stable transfer function $(1 - \boldsymbol{gk})^{-1} = \left(\frac{s+1}{2(s+2)}\right)$ mapping d to y. However, the closed-loop transfer function between d and u is $\boldsymbol{k}(1-\boldsymbol{gk})^{-1} = \left(\frac{(s+1)(s+3)}{2s(s+2)}\right)$, which is unstable due to the closed-loop pole at the origin. We therefore conclude that the system in Figure 1.4 is not internally stable for this particular plant and controller combination, although it is input-output stable. ∇

We will now prove our first result on internal stability.

Lemma 1.3.1 *The feedback loop in Figure 1.4 is internally stable if and only if*

$$\begin{bmatrix} 1 & -\boldsymbol{k} \\ -\boldsymbol{g} & 1 \end{bmatrix}^{-1} \tag{1.3.3}$$

is stable.

Proof. It is immediate from Figure 1.4 that

$$\begin{aligned} u &= ky + w \\ y &= gu + d, \end{aligned}$$

or equivalently

$$\begin{bmatrix} w \\ d \end{bmatrix} = \begin{bmatrix} 1 & -\boldsymbol{k} \\ -\boldsymbol{g} & 1 \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix}.$$

This gives

$$\begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} 1 & -\boldsymbol{k} \\ -\boldsymbol{g} & 1 \end{bmatrix}^{-1} \begin{bmatrix} w \\ d \end{bmatrix}$$

and the result follows from Definition 1.3.1. ∎

[2] The terms internally-stabilizing controller and stabilizing controller are synonymous in this book—internally-stabilizing controller is used to draw special attention to the requirement of internal stability.

We will now discover the properties required of the q-parameter defined in (1.3.1) for internal stability in the stable plant case. Since

$$\begin{bmatrix} 1 & -k \\ -g & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -g & 1 \end{bmatrix} \begin{bmatrix} 1 & -k \\ 0 & 1-gk \end{bmatrix},$$

we get

$$\begin{aligned} \begin{bmatrix} 1 & -k \\ -g & 1 \end{bmatrix}^{-1} &= \begin{bmatrix} 1 & k(1-gk)^{-1} \\ 0 & (1-gk)^{-1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ g & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & q \\ 0 & 1+gq \end{bmatrix} \begin{bmatrix} 1 & 0 \\ g & 1 \end{bmatrix} \end{aligned}$$

on substituting from (1.3.1). Since g is assumed stable, it is apparent that

$$\begin{bmatrix} 1 & -k \\ -g & 1 \end{bmatrix}^{-1}$$

is stable if and only if q is stable. This gives the following result:

Lemma 1.3.2 *Suppose g is stable. Then k is an internally-stabilizing controller for the feedback loop in Figure 1.4 if and only if $q = k(1-gk)^{-1}$ is stable. Equivalently, k is an internally-stabilizing controller if and only if $k = q(1+qg)^{-1}$ for some stable q.*

1.3.2 Solution of the disturbance attenuation problem

We may now return to the disturbance attenuation problem given in (1.3.2). Since the transfer functions that maps d to y is given by

$$h = 1 + gq, \tag{1.3.4}$$

one obtains

$$q = g^{-1}(h-1).$$

For the loop to be internally stable, we need to ensure that q is stable.

When g^{-1} is stable we could, in principle, set $q = -g^{-1}$, since this results in $h = 0$ and perfect disturbance attenuation. Unfortunately, such a q is not achievable by a realizable controller since k has infinite gain. We may, however, use $q = -(1-\epsilon)g^{-1}$ for an arbitrarily small ϵ. This gives $h = \epsilon$ and

$$k = -\left(\frac{1-\epsilon}{\epsilon}\right)g^{-1}.$$

The controller is simply the negative of the inverse of the plant together with an arbitrarily high gain factor. This is not a surprising conclusion, because high gain improves disturbance attenuation and we know from classical root locus theory that a plant will be closed-loop stable for arbitrarily high gain if all the plant zeros are in the open-left-half plane.

In the case that $\boldsymbol{g}^{-1}$ is not stable, $\boldsymbol{q}$ will be stable if and only if

$$\boldsymbol{h}(z_i) = 1, \qquad i = 1, 2, \ldots, m, \tag{1.3.5}$$

for each zero, z_i, of $\boldsymbol{g}$ such that $\mathrm{R}_e(z_i) \geq 0$ (provided each of the zeros z_i is of multiplicity one). The optimal disturbance attenuation problem therefore requires us to find a stable closed-loop transfer function $\boldsymbol{h}$, of minimum infinity norm, which satisfies the interpolation constraints given in (1.3.5). It follows from (1.3.4) that the corresponding optimal $\boldsymbol{q}$ may be interpreted as the best stable approximate inverse of $-\boldsymbol{g}$, in the infinity norm sense.

It follows from the maximum modulus principle that the constraints $\boldsymbol{h}(z_i) = 1$ make it impossible to achieve $\|\boldsymbol{h}\|_\infty < 1$ when the plant has a right-half-plane zero. Since the plant is stable, we can set $\boldsymbol{k} = 0$ to achieve $y = d$, which is optimal in this case. The presence of a right-half-plane zero makes broadband disturbance attenuation impossible.

If some spectral information is available about the disturbance d, one may be able to improve the situation by introducing frequency response weighting. If d is bandlimited, we could seek to minimize $\|\boldsymbol{wh}\|_\infty$ in which $\boldsymbol{w}$ is some low-pass stable and minimum phase weighting function. If $\|\boldsymbol{wh}\|_\infty < 1$, it follows that $|\boldsymbol{h}(j\omega)| < |\boldsymbol{w}^{-1}(j\omega)|$ for all real ω. Since $|\boldsymbol{w}^{-1}(j\omega)|$ is small at low frequency due to the low pass nature of $\boldsymbol{w}$, it follows that $|\boldsymbol{h}(j\omega)|$ will also be small there. The idea is that $|\boldsymbol{h}(j\omega)|$ should be small over the range of frequencies for which $|d(j\omega)|$ is large. If we set $\widehat{\boldsymbol{h}} = \boldsymbol{wh}$, one obtains

$$\widehat{\boldsymbol{h}} = \boldsymbol{w} + \boldsymbol{wgq}$$

and consequently that

$$\boldsymbol{q} = \boldsymbol{g}^{-1}\boldsymbol{w}^{-1}(\widehat{\boldsymbol{h}} - \boldsymbol{w}).$$

Under these conditions the $\boldsymbol{q}$-parameter will be stable if and only if the interpolation constraints

$$\widehat{\boldsymbol{h}}(z_i) = \boldsymbol{w}(z_i), \qquad i = 1, 2, \ldots, m,$$

are satisfied. If the right-half-plane plant zeros occur beyond the bandwidth of the weighting function, the $\boldsymbol{w}(z_i)$'s will be small and it is at least possible that an $\widehat{\boldsymbol{h}}$ can be found such that $\|\widehat{\boldsymbol{h}}\|_\infty < 1$. Since $\|\widehat{\boldsymbol{h}}\|_\infty < 1 \ \Rightarrow \ |\boldsymbol{h}(j\omega)| < |\boldsymbol{w}^{-1}(j\omega)|$ for all ω, we conclude that $|\boldsymbol{h}(j\omega)| < \epsilon$ whenever $|\boldsymbol{w}(j\omega)| \geq 1/\epsilon$. Consequently, by designing $\boldsymbol{w}$, one can guarantee an appropriate level of disturbance attenuation provided a controller exists such that $\|\widehat{\boldsymbol{h}}\|_\infty < 1$. Conversely, if $\boldsymbol{w}(z_i) > 1$ for at least one z_i, we must have $\|\widehat{\boldsymbol{h}}\|_\infty > 1$ and $|\boldsymbol{w}(j\omega)| \geq 1/\epsilon$ no longer ensures $|\boldsymbol{h}(j\omega)| < \epsilon$.

Main points of the section

1. The optimal disturbance attenuation problem is a feedback problem and it is possible to replace the nonlinear parametrization of h in terms of stabilizing controllers k, by an affine parametrization of h in terms of stable functions q. So far we have only established this fact for the stable plant case, but it is true in general.
2. The optimization problem requires us to find a stable transfer function h of minimum norm that satisfies the interpolation constraints given in (1.3.5). This is a classical Nevanlinna-Pick interpolation problem and satisfaction of the interpolation constraints guarantees the internal stability of the feedback system. We note that minimizing $\|h\|_\infty$ is equivalent to finding a stable approximate inverse of the plant.
3. If the plant has a right-half-plane zero, the constraint $h(z_i) = 1$ makes it impossible to achieve $\|h\|_\infty < 1$ thereby attenuating unknown disturbances. In this case the best one can do is set $k = 0$, since this will give $y = d$. If some spectral information about the disturbance is available, the situation may be improved if the right-half-plane zero is outside the bandwidth in which there is significant disturbance energy.

1.4 A robust stability problem

When a design team is faced with the problem of designing a controller to meet certain closed-loop performance specifications, they will hardly ever have a perfect model of the plant. As a consequence, the design process is complicated by the fact that the controller has to be designed to operate satisfactorily for all plants in some model set. The most fundamental of all design requirements is that of finding a controller to stabilize all plants in some class; we call this the robust stabilization problem. To set this problem up in a mathematical optimization framework, we need to decide on some representation of the model error. If the nominal plant model is g, we can use an additive representation of the model error by describing the plant as $g + \delta$ in which the stable transfer function δ represents the unknown dynamics; this is an alternative to the multiplicative description of model error given in Section 1.1.

Let us consider the robust stabilization problem in which some nominal plant model g is given, and we seek a stabilizing controller for all plants of the form $g + \delta$ in which the allowable $\|\delta\|_\infty$ is maximized. A controller that maximizes $\|\delta\|_\infty$ is optimally robust in the sense that it stabilizes the largest ball of plants with center g. A block diagram of the set-up under consideration is given in Figure 1.5 and

$$z = (1 - kg)^{-1} kw.$$

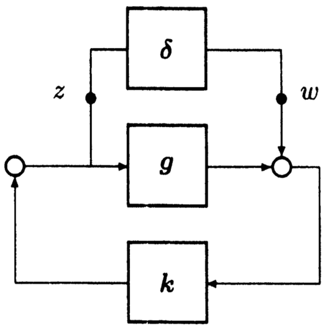

Figure 1.5: A robust stability problem.

If δ and the nominal closed-loop system are stable, it follows from an earlier "small gain" argument based on the Nyquist criterion that the perturbed closed loop will also be stable provided

$$\|\delta\|_\infty \|(1 - kg)^{-1} k\|_\infty < 1.$$

The optimal robustness problem therefore requires a stabilizing controller that minimizes $\|(1 - kg)^{-1} k\|_\infty$.

As before, in the case that the plant is stable, the solution is trivially obtained by setting $k = 0$; note, however, that $k = 0$ offers no protection against unstable perturbations however small! Before substituting

$$q = (1 - kg)^{-1} k,$$

we need the conditions on q that lead to a stable nominal closed-loop system. The mere stability of q is not enough in the unstable plant case. Since

$$\begin{bmatrix} 1 & -k \\ -g & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 + qg & q \\ (1 + qg)g & 1 + gq \end{bmatrix},$$

it is clear that the nominal closed loop will be stable if and only if

1. q is stable,
2. gq is stable, and
3. $(1 + qg)g$ is stable.

If g is stable and Condition 1 is satisfied, Conditions 2 and 3 follow automatically. If $(p_1, p_2, \ldots, p_m)$ are the right-half-plane poles of g, it follows from Condition 2 that internal stability requires satisfaction of the interpolation constraints

2′. $q(p_i) = 0$, for $i = 1, 2, \ldots, m$, while Condition 3 demands

3′. $(1 + gq)(p_i) = 0$, for $i = 1, 2, \ldots, m$.

To keep things simple, we will assume for the present that each unstable pole has multiplicity one and that $\mathrm{R}_e(p_i) > 0$.

Since the closed-loop transfer function of interest is q, the solution of the robust stabilization problem requires a stable q of minimum infinity norm that satisfies the interpolation constraints of Conditions 2′ and 3′.

As we will now show, it is possible to reformulate the problem so that there is one, rather than two, interpolation constraints per right-half-plane pole. To effect the reformulation, we introduce the completely unstable function[3]

$$a = \prod_{i=1}^{m} \left(\frac{\bar{p}_i + s}{p_i - s} \right) \tag{1.4.1}$$

which has the property that $|a(j\omega)| = 1$ for all real ω. If we define $\tilde{q} := aq$ it follows that:

1. $\|q\|_\infty = \|\tilde{q}\|_\infty$.
2. If $\tilde{q}$ is stable, so is q.
3. If $\tilde{q}$ is stable, $q(p_i) = 0$, because $q = \tilde{q} \prod_{i=1}^{m} \left(\frac{p_i - s}{\bar{p}_i + s} \right)$.
4. $\tilde{q}(p_i) = -(ag^{-1})(p_i) \Rightarrow (1 + qg)(p_i) = 0$.

In its new form, the robust stabilization problem is one of finding a stable $\tilde{q}$ of minimum infinity norm such that

$$\tilde{q}(p_i) = -(ag^{-1})(p_i) \qquad i = 1, 2, \ldots, m, \tag{1.4.2}$$

which is yet another Nevanlinna-Pick interpolation problem . The corresponding (optimal) controller may be found by back substitution as

$$k = (a + \tilde{q}g)^{-1}\tilde{q}. \tag{1.4.3}$$

Example 1.4.1. Suppose the plant is given by

$$g = \frac{s + 2}{(s + 1)(s - 1)}.$$

Since there is a single right-half-plane pole at +1, it follows that the allpass function given in equation (1.4.1) is

$$a = \left(\frac{1 + s}{1 - s} \right)$$

[3]Such functions are sometimes known as Blaschke products.

in this particular case. As a consequence

$$-ag^{-1} = \frac{(s+1)^2}{(s+2)},$$

and the interpolation condition follows from (1.4.2) as

$$\tilde{q}(1) = -ag^{-1}\Big|_{s=1} = \frac{4}{3}.$$

It is now immediate from the maximum modulus principle that $\|\tilde{q}\|_\infty \geq 4/3$, so that $\tilde{q} = 4/3$ is optimal. Substitution into (1.4.3) yields

$$k = -\frac{4(s+1)}{(3s+5)}$$

as the optimal controller that will stabilize the closed-loop system for all stable δ such that $\|\delta\|_\infty < 3/4$. ∇

Our second robust stabilization example shows that it is impossible to robustly stabilize a plant with a right-half-plane pole-zero pair that almost cancel. We expect such a robust stability problem to be hard, because problems of this type have an unstable mode that is almost uncontrollable.

Example 1.4.2. Consider the unstable plant

$$g = \left(\frac{s-\alpha}{s-1}\right), \quad \alpha \neq 1,$$

which has a zero at α. As with the previous example, we require

$$a = \left(\frac{1+s}{1-s}\right)$$

which gives

$$-ag^{-1} = \left(\frac{s+1}{s-\alpha}\right).$$

The only interpolation constraint is therefore

$$\tilde{q}(1) = -ag^{-1}\Big|_{s=1} = \frac{2}{1-\alpha}.$$

Invoking the maximum modulus principle yields $\tilde{q} = 2/(1-\alpha)$ as the optimal interpolating function. Substitution into (1.4.3) gives

$$k = \frac{2}{1+\alpha}$$

as the optimal controller. The closed loop will therefore be stable for all stable δ such that $\|\delta\|_\infty < |(1-\alpha)/2|$. From this we conclude that the stability margin measured by the maximum allowable $\|\delta\|_\infty$ vanishes as $\alpha \to 1$. ∇

Our final example considers the robust stabilization of an integrator.

Example 1.4.3. Consider the case of

$$g = \frac{1}{s}.$$

At first sight this appears to be an awkward problem because the interpolation constraint occurs at $s = 0$, and the allpass function in (1.4.1) degenerates to 1. Suppose we ignore this difficulty for the moment and restrict our attention to constant controllers given by $k \leq 0$. This gives

$$q = (1 - kg)^{-1}k = \frac{ks}{s - k}$$

with

$$\begin{aligned} \|(1 - kg)^{-1}k\|_\infty &= \left|\frac{sk}{s - k}\right|\Bigg|_{s=\infty} \\ &= |k|. \end{aligned}$$

To solve the problem we observe that if we want to stabilize the closed loop for any stable δ such that $\|\delta\|_\infty < 1/\epsilon$, we simply set $k = -\epsilon$; ϵ may be arbitrarily small! In problems such as this one, which has an interpolation constraint on the imaginary axis, it is not possible to achieve the infimal value of the norm. For any positive number, we can achieve a closed-loop with that number as its infinity norm, but we cannot achieve a closed-loop infinity norm of zero. ▽

1.5 Concluding comments and references

We will now conclude this introductory chapter with a few remarks about the things we have already learned and the things we still hope to achieve.

1. $\mathcal{H}_\infty$ control problems can be cast as constrained minimization problems. The constraints come from an internal stability requirement and the object we seek to minimize is the infinity norm of some closed-loop transfer function. The constraints appear as interpolation constraints and stable closed-loop transfer functions that satisfy the interpolation data may be found using the classical Nevanlinna-Schur algorithm. This approach to control problems is due to Zames [227] and is developed in Zames and Francis [228] and Kimura [118]. In our examples we have exploited the fact that there is no need for the Nevanlinna algorithm when there is only one interpolation constraint.

2. We will not be discussing the classical Nevanlinna-Pick-Schur theory on analytic interpolation in this book. The interested reader may find this material in several places such as Garnett [69] and Walsh [207] for a purely function theoretic point of view, and [53, 43, 44, 129, 221, 227, 228], for various applications of analytic interpolation to system theory.

3. The reader may be puzzled as to why the interpolation theory approach to $\mathcal{H}_\infty$ control problems is being abandoned at this early stage of our book. There are several reasons for this:

   (a) Interpolation theoretic methods become awkward and unwieldy in the multivariable case and in situations where interpolation with multiplicities is required; if there are several interpolation constraints associated with a single right-half-plane frequency point, we say that the problem involves interpolation with multiplicities.

   (b) It is our opinion that interpolation theoretic methods are computationally inferior to the state-space methods we will develop in later chapters of the book. Computational issues become important in realistic design problems in which one is forced to deal with systems of high order.

   (c) Frequency domain methods (such as interpolation theory) are restricted to time-invariant problems. The state-space methods we will develop are capable of treating linear time varying problems.

   (d) It is not easy to treat multitarget problems in an interpolation based framework. To see this we cite one of many possible problems involving robust stabilization with performance. Take the case of disturbance attenuation with robust stability, in which we require a characterization of the set

$$\arg\min_{k\in\mathcal{S}} \left\| \begin{bmatrix} (1-gk)^{-1} \\ k(1-gk)^{-1} \end{bmatrix} \right\|_\infty$$

   with $\mathcal{S}$ denoting the set of all stabilizing controllers. If the plant is stable, we may introduce the q-parameter to obtain

$$\arg\min_{q\in\mathcal{H}_\infty} \left\| \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} g \\ 1 \end{bmatrix} q \right\|_\infty .$$

   Problems of this type are not directly addressable via interpolation due to the nonsquare nature of $\begin{bmatrix} g \\ 1 \end{bmatrix}$; we will not pursue this point at this stage.

4. Solving each $\mathcal{H}_\infty$ control problem from scratch, as we have done so far, is a practice we will now dispense with. This approach is both effort intensive and an intellectually clumsy way to proceed. Rather, we will develop a single solution framework that captures many $\mathcal{H}_\infty$ optimization problems of general interest as special cases. A large part of the remainder of the book will be devoted to the development of a comprehensive theory for multivariable, multitarget problems.

5. The solutions to the problems we have considered so far have a common theme. With the exception of the robust stabilization of an integrator, the

magnitudes of the optimal closed-loop transfer functions are a constant function of frequency. It turns out that this is a general property of the solutions of all single-input, single-output problems that are free of imaginary axis interpolation constraints. In each case, the optimal closed-loop transfer function is a scalar multiple of a rational inner function. Inner functions are stable allpass functions, and rational allpass functions have the form

$$a = \prod_{i=1}^{m} \left(\frac{\bar{p}_i + s}{p_i - s} \right)$$

which we have already encountered. Since the poles and zeros of allpass functions are symmetrically located about the imaginary axis, it is not hard to see that they have the property $|a(j\omega)| = 1$ for all real ω. The "flat frequency response" property of optimal closed-loop transfer functions is fundamental in the design of frequency weighting functions.

1.6 Problems

Problem 1.1. Prove that $\|\cdot\|_\infty$ is a norm and that $\|gh\|_\infty \le \|g\|_\infty \|h\|_\infty$.

Problem 1.2. Consider the frequency weighted disturbance attenuation problem of finding a stabilizing controller that minimizes $\|w(1-gk)^{-1}\|_\infty$. If

$$g = \left(\frac{s-\alpha}{s+2} \right), \qquad w = \left(\frac{s+4}{2(s+1)} \right),$$

in which α is real, show that when $0 \le \alpha \le 2$ there is no stabilizing controller such that

$$|(1-gk)^{-1}(j\omega)| < |w^{-1}(j\omega)|, \qquad \text{for all } \omega.$$

Problem 1.3. Consider the command tracking problem in which

$$g = \left(\frac{(s-1)^2}{(s+2)(s+3)} \right), \qquad h = \frac{1}{s+4}.$$

Show that the error $e = h - gf$ must satisfy the interpolation constraints

$$e(1) = \frac{1}{5}, \qquad \frac{de}{ds}(1) = \frac{-1}{25}.$$

The construction of such an e requires the solution of an interpolation problem with derivative constraints.

Problem 1.4. Suppose an uncertain plant is described by $g(1+\delta)$ in which g is a given unstable transfer function and δ is a stable but otherwise unknown linear perturbation bounded in magnitude by $\|\delta\|_\infty < \alpha$.

1. Give an interpolation theoretic procedure for finding the optimal controller that stabilizes every $g(1+\delta)$ of the type described and with α maximized. (Hint: you need to introduce the stable minimum phase spectral factor m that satisfies $gg^{\sim} = mm^{\sim}$.)
2. Give two reasons why α must always be strictly less than one.
3. Suppose $g = \left(\frac{s-2}{s-1}\right)$. Show that the largest achievable value of α is $\alpha_{max} = \frac{1}{3}$, and that the corresponding controller is $k = \frac{3}{4}$.

Problem 1.5. Suppose an uncertain plant is described by $g+\delta$ in which g is a given unstable transfer function and δ is a stable but otherwise unknown linear perturbation such that $|\delta(j\omega)| < |w(j\omega)|$ for all ω. The function w is a stable and minimum phase frequency weight.

1. Show that k will stabilize all $g+\delta$ with δ in the above class provided it stabilizes g and $\|wk(1-gk)^{-1}\|_\infty \leq 1$.
2. Explain how to find a stabilizing controller that minimizes $\|wk(1-gk)^{-1}\|_\infty$.
3. If $g = \left(\frac{s+1}{s-2}\right)$ and $w = \left(\frac{s+1}{s+4}\right)$, find a controller (if one exists) that will stabilize every $g+\delta$ in which δ is stable with $|\delta(j\omega)| < |w(j\omega)|$ for all ω.

Problem 1.6. Consider the multivariable command response optimization problem in which the stable transfer function matrices G and H are given and a stable prefilter F is required such that $E = H - GF$ is small in some sense.

1. If G is nonsingular for almost all s and F is to be stable, show that $H - E$ must have a zero at each right-half-plane zero of G, taking multiplicities into account.
2. If all the right-half-plane zeros z_i, $i = 1, 2, \ldots, m$, of G are of multiplicity one, show that F is stable if and only if there exist vectors $w_i \neq 0$ such that

$$w_i^* \begin{bmatrix} H(z_i) - E(z_i) & G(z_i) \end{bmatrix} = 0.$$

   Conclude from this that multivariable problems have vector valued interpolation constraints. What are they?

The relationship between vector interpolation and $\mathcal{H}_\infty$ control is studied in detail in Limebeer and Anderson [129] and Kimura [119].