Math. Control Signals Systems (1989) 2: 207–219

Mathematics of Control, Signals, and Systems

A Bisection Method for Computing the $\mathbf{H}_\infty$ Norm of a Transfer Matrix and Related Problems*

S. Boyd,† V. Balakrishnan,† and P. Kabamba‡

Abstract. We establish a correspondence between the singular values of a transfer matrix evaluated along the imaginary axis and the imaginary eigenvalues of a related Hamiltonian matrix. We give a simple linear algebraic proof, and also a more intuitive explanation based on a certain indefinite quadratic optimal control problem. This result yields a simple bisection algorithm to compute the $\mathbf{H}_\infty$ norm of a transfer matrix. The bisection method is far more efficient than algorithms which involve a search over frequencies, and the usual problems associated with such methods (such as determining how fine the search should be) do not arise. The method is readily extended to compute other quantities of system-theoretic interest, for instance, the minimum dissipation of a transfer matrix. A variation of the method can be used to solve the $\mathbf{H}_\infty$ Armijo line-search problem with no more computation than is required to compute a single $\mathbf{H}_\infty$ norm.

Key words. $\mathbf{H}_\infty$ norm, Hamiltonian, Sturm test, Armijo line search.

1. Preliminaries

Throughout this paper A, B, C, D will be real matrices of sizes $n \times n$, $n \times m$, $p \times n$, and $p \times m$, respectively. We refer to the linear dynamical system

$$\begin{aligned} \dot{x} &= Ax + Bu, \\ y &= Cx + Du \end{aligned} \tag{1}$$

as the *system* $\{A, B, C, D\}$. We refer to $H(s) = C(sI - A)^{-1}B + D$ as the *transfer matrix* of the system $\{A, B, C, D\}$.

A is *stable* means that all eigenvalues of A have negative real part. If A is stable, we define the $\mathbf{H}_\infty$ norm of the transfer matrix $H(s)$ to be

$$\|H\|_\infty = \sup_{\operatorname{Re} s > 0} \sigma_{\max}(H(s)) = \sup_{\omega \in \mathbb{R}} \sigma_{\max}(H(j\omega)), \tag{2}$$

where $\sigma_{\max}(\cdot)$ denotes the maximum singular value of a matrix, that is, $\sigma_{\max}(F) =$

---

* Date received: August 28, 1987. Date revised: March 16, 1988.

† Information Systems Laboratory, Electrical Engineering Department, Stanford University, Stanford, California 94305, U.S.A. Research supported in part by NSF under Grant ECS-85-52465, ONR under Grant N00014-86-K-0112, an IBM faculty development award, and Bell Communications Research.

‡ Aerospace Engineering Department, University of Michigan, Ann Arbor, Michigan, U.S.A.

$\lambda_{\max}^{1/2}(F^*F)$. The $\mathbf{H}_\infty$ norm of a transfer matrix arises often in control theory; indeed, there has been considerable recent interest in designing controllers that minimize the $\mathbf{H}_\infty$ norm of some closed-loop transfer matrix of a system (e.g., [ZF] and [V3]). One important interpretation of $\|H\|_\infty$ is as the $\mathbf{L}_2$ or RMS gain of the system (1) (see, e.g., [BD]): whenever x, u, y satisfy (1), $x(0) = 0$, and $T_f > 0$, we have

$$\int_0^{T_f} y(t)^T y(t)\,dt \le \|H\|_\infty^2 \int_0^{T_f} u(t)^T u(t)\,dt. \tag{3}$$

Much research has focused on the numerical computation of these $\mathbf{H}_\infty$ optimal controllers; nevertheless little attention has been paid to the relatively *easier* problem of the numerical computation of $\|H\|_\infty$, given the matrices A, B, C, and D. $\|H\|_\infty$ is usually "computed" by searching for the maximum of $\sigma_{\max}(H(j\omega))$ over $\omega \in \mathbb{R}$. $\sigma_{\max}(H(j\omega))$ is computed for many values of ω; a logarithmic spacing with, say, 20 points per decade over a range spanning five decades is common. Often a plot (referred to as a singular-value or SV plot) is drawn from these computed values. Obvious problems associated with such a method are (a) determining the range and spacing of the frequencies to be checked, and (b) the large number of computations involved (a singular-value decomposition (SVD) is often performed at each frequency point). The problem (a) is particularly evident when A has eigenvalues with small real part, as happens when (1) represents a lightly damped mechanical structure.

Several techniques can substantially reduce the number of computations involved in an SV plot. First, the transfer matrix can be evaluated using Laub's method [L1] for efficient computation of $H(s)$ for many values of s. Second, instead of computing an entire SVD of each $H(j\omega_k)$ from scratch, a power method [GV] can be applied to the Hermitian positive semidefinite matrix

$$\begin{bmatrix} 0 & H(j\omega_k) \\ H(j\omega_k)^* & 0 \end{bmatrix}$$

to compute its maximum eigenvalue, which is $\sigma_{\max}^2(H(j\omega_k))$. The previously computed eigenvector of

$$\begin{bmatrix} 0 & H(j\omega_{k-1}) \\ H(j\omega_{k-1})^* & 0 \end{bmatrix}$$

can be used as an initial vector for the power method. Still, computing $\sigma_{\max}(H(j\omega))$ for many values of ω requires considerable computation, and, as mentioned above, there is no good way to know how well $\max_k \sigma_{\max}(H(j\omega_k))$ approximates $\|H\|_\infty$.

We propose instead a bisection method inspired by Byers' bisection method for measuring the distance of a stable matrix to the set of unstable matrices [B2]. The bisection method not only involves less computation, but has the advantage of computing $\|H\|_\infty$ with a guaranteed accuracy.

2. Singular Values of a Transfer Matrix via a Hamiltonian Matrix

We start by establishing a connection between the singular values of the transfer matrix and the imaginary eigenvalues of a certain Hamiltonian matrix. Let $\gamma > 0$, and not a singular value of D. Define

$$M_\gamma = \begin{bmatrix} A & 0 \\ 0 & -A^T \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & -C^T \end{bmatrix} \begin{bmatrix} -D & \gamma I \\ \gamma I & -D^T \end{bmatrix}^{-1} \begin{bmatrix} C & 0 \\ 0 & B^T \end{bmatrix}$$
$$= \begin{bmatrix} A - BR^{-1}D^TC & -\gamma BR^{-1}B^T \\ \gamma C^TS^{-1}C & -A^T + C^TDR^{-1}B^T \end{bmatrix}, \tag{4}$$

where $R = (D^TD - \gamma^2 I)$ and $S = (DD^T - \gamma^2 I)$. M_γ is a *Hamiltonian* matrix, meaning

$$J^{-1}M_\gamma J = -M_\gamma^T \qquad \text{where} \quad J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

The following theorem relates the singular values of $H(j\omega)$ and the imaginary eigenvalues of M_γ.

Theorem 1. *Assume A has no imaginary eigenvalues, $\gamma > 0$ is not a singular value of D, and $\omega_0 \in \mathbb{R}$. Then, γ is a singular value of $H(j\omega_0)$ if and only if $(M_\gamma - j\omega_0 I)$ is singular.*

Proof. Let γ be a singular vaue of $H(j\omega_0)$. Then we have a nonzero u such that

$$\begin{aligned} H(j\omega_0)u &= \gamma v, \\ H(j\omega_0)^* v &= \gamma u, \end{aligned} \tag{5}$$

so that

$$\begin{aligned} (C(j\omega_0 I - A)^{-1}B + D)u &= \gamma v, \\ (B^T(-j\omega_0 I - A^T)^{-1}C^T + D^T)v &= \gamma u. \end{aligned} \tag{6}$$

Define

$$\begin{aligned} r &= (j\omega_0 I - A)^{-1}Bu, \\ s &= (-j\omega_0 I - A^T)^{-1}C^T v. \end{aligned} \tag{7}$$

Now solving for u and v in terms of r and s,

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -D & \gamma I \\ \gamma I & -D^T \end{bmatrix}^{-1} \begin{bmatrix} C & 0 \\ 0 & B^T \end{bmatrix} \begin{bmatrix} r \\ s \end{bmatrix}. \tag{8}$$

Note that (8) guarantees that

$$\begin{bmatrix} r \\ s \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

From (7)

$$\begin{aligned} (j\omega_0 I - A)r &= Bu, \\ (-j\omega_0 I - A^T)s &= C^T v. \end{aligned} \tag{9}$$

From (8) and (9), we obtain

$$\left\{ \begin{bmatrix} A & 0 \\ 0 & -A^T \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & -C^T \end{bmatrix} \begin{bmatrix} -D & \gamma I \\ \gamma I & -D^T \end{bmatrix}^{-1} \begin{bmatrix} C & 0 \\ 0 & B^T \end{bmatrix} \right\} \begin{bmatrix} r \\ s \end{bmatrix} = j\omega_0 \begin{bmatrix} r \\ s \end{bmatrix}. \tag{10}$$

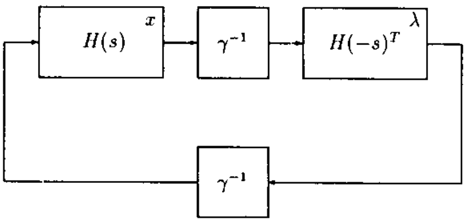

Fig. 1. $\mathbf{H}_\gamma$ gain-normalized closed-loop feedback system.

Thus

$$M_\gamma \begin{bmatrix} r \\ s \end{bmatrix} = j\omega_0 \begin{bmatrix} r \\ s \end{bmatrix}. \tag{11}$$

This proves one direction of Theorem 1.

Now we prove the converse. Suppose that M_γ has eigenvalue $j\omega_0$, that is, (10) holds for some $\begin{bmatrix} r \\ s \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Define u and v by equation (8); clearly, $[u^T \ v^T] \neq 0$. Then from (8) and (10), we conclude (6), which establishes that γ is a singular value of $H(j\omega_0)$. ∎

Remark 1. There are no observability, controllability, or stability conditions on the system $\{A, B, C, D\}$.

Remark 2. Let us give an intuitive explanation of this theorem. Consider the feedback system shown in Fig. 1. If γ is a singular value of $H(j\omega_0)$, this system has a nonzero solution of the form $ue^{j\omega_0 t}$. If $H(s)$ is the transfer matrix of $\{A, B, C, D\}$ then $H(-s)^T$ is the transfer matrix of $\{-A^T, -C^T, B^T, D^T\}$, and a realization of the system above is

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} A - BR^{-1}D^TC & -\gamma BR^{-1}B^T \\ \gamma C^TS^{-1}C & -A^T + C^TDR^{-1}B^T \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = M_\gamma \begin{bmatrix} x \\ \lambda \end{bmatrix}.$$

Thus we see that γ is a singular value of $H(j\omega_0)$ if and only if M_γ has eigenvalue $j\omega_0$.

Remark 3. It is often the case that $D = 0$, in which case M_γ has the simple form

$$M_\gamma = \begin{bmatrix} A & \gamma^{-1}BB^T \\ -\gamma^{-1}C^TC & -A^T \end{bmatrix}. \tag{12}$$

A simple consequence of Theorem 1 is

Theorem 2. *Let A be stable and $\gamma > \sigma_{\max}(D)$. Then $\|H\|_\infty \geq \gamma$ if and only if M_γ has imaginary eigenvalues (i.e., at least one).*

Proof. Since $\gamma > \sigma_{\max}(D) = \lim_{\omega \to \infty} \sigma_{\max}(H(j\omega))$ and $\sigma_{\max}(H(j\omega))$ is a continuous function of ω, we have $\|H\|_\infty \geq \gamma$ if and only if there exists ω_0 such that $\sigma_{\max}(H(j\omega_0)) = \gamma$. Hence Theorem 2 follows immediately from Theorem 1. ∎

Remark 4. There is a connection between the $\mathbf{H}_\infty$ norm of H, the Hamiltonian matrix M_γ, and the following indefinite quadratic optimal control problem:

$$\min \int_0^\infty (\gamma^2 u^T u - y^T y)\, dt. \tag{13}$$

$$\dot{x} = Ax + Bu,$$

$$y = Cx + Du,$$

$$x(0) = 0.$$

First, recall that

$$\|H\|_\infty^2 = \max \int_0^\infty y^T y\, dt \tag{14}$$

subject to $\dot{x} = Ax + Bu$, $y = Cx + Du$, $x(0) = 0$, and $\int_0^\infty u^T u\, dt \leq 1$ (see [DV]). Thus we see that $\|H\|_\infty > \gamma$ implies that the minimum value of (13) is $-\infty$, whereas if the minimum value of (13) is greater than $-\infty$ (and hence zero), we have $\|H\|_\infty \leq \gamma$.

The Hamiltonian matrix associated with the quadratic optimal control problem (13) is simply M_γ (see [W]), and aside from some technical details, the condition that the minimum value of (13) be finite is equivalent to M_γ not having any imaginary eigenvalues.

Indeed, if M_γ has no imaginary eigenvalues then it can be put in the Schur form:

$$Q^* M_\gamma Q = \mathrm{diag}[\lambda_1, \ldots, \lambda_n, -\lambda_1, \ldots, -\lambda_n] + N,$$

where Q is unitary, N is strictly upper triangular, and $\mathrm{Re}\, \lambda_i < 0$, $i = 1, \ldots, n$. If we partition Q as

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$$

(each Q_{ij} is $n \times n$), then it can be shown that $P = Q_{21} Q_{11}^{-1}$ is symmetric and positive definite, and the function $V(x) = \frac{1}{2} x^T P x$ satisfies

$$\frac{d}{dt}(V(x(t))) \leq \gamma^2 u^T u - y^T y$$

whenever $\dot{x} = Ax + Bu$, $y = Cx + Du$. Thus when $x(0) = 0$ and $T_f > 0$ we have

$$\int_0^{T_f} y^T y\, dt \leq \gamma^2 \int_0^{T_f} u^T u\, dt - V(x(T_f)) \leq \gamma^2 \int_0^{T_f} u^T u\, dt.$$

Along with (14), this provides another proof that if M_γ has no imaginary eigenvalues then $\|H\|_\infty \leq \gamma$.

Remark 5. Theorem 2 may also be derived from certain results of the 1960s concerning spectral factorization, and so is known to several researchers,[1] though

[1] John Doyle, personal communication.

we are unaware of its appearance in the literature. Specifically, we note that $\|H\|_\infty \leq \gamma$ if and only if the transfer matrix $\gamma^2 I - H(-s)^T H(s)$ has a *spectral factorization* [A], [F], that is, if there exists a transfer matrix G such that

$$\gamma^2 I - H(-s)^T H(s) = G(-s)^T G(s). \tag{15}$$

Following [F], it can be shown that in attempting a spectral factorization of $\gamma^2 I - H(-s)^T H(s)$, M_γ arises naturally as a matrix similar to the system matrix of the inverse system $(\gamma^2 I - H(-s)^T H(s))^{-1}$. The necessary and sufficient condition for the spectral factorization [A] turns out to be the nonexistence of imaginary eigenvalues of M_γ.

Remark 6. If A is stable, then

$$\beta(A) = \inf\{\sigma_{\max}(E) \mid E \in \mathbb{C}^{n\times n} \text{ and } A + E \text{ has imaginary eigenvalues}\}$$

is referred to as the *distance to the nearest unstable matrix* in the numerical analysis literature [B2], [V2]. Note that

$$\begin{aligned}\beta(A) &= \inf_{\omega\in\mathbb{R}} \inf\{\sigma_{\max}(E) \mid A + E - j\omega I \text{ is singular}\}\\ &= \inf_{\omega\in\mathbb{R}} \sigma_{\min}(A - j\omega I)\\ &= \left\{\sup_{\omega\in\mathbb{R}} \sigma_{\max}((j\omega I - A)^{-1})\right\}^{-1} = \|(sI - A)^{-1}\|_\infty^{-1}.\end{aligned}$$

Thus $\beta(A)^{-1}$ is simply the $\mathbf{H}_\infty$ norm of the resolvent of A, and Theorem 2 yields Byers' result [B2],

$$\beta(A) = \inf\left\{\alpha \,\middle|\, \begin{bmatrix} A & \alpha I \\ -\alpha I & -A^T \end{bmatrix} \text{has imaginary eigenvalues}\right\}.$$

Remark 7. If A is not stable, but has no imaginary eigenvalues, then Theorem 2 remains true when the $\mathbf{H}_\infty$ norm is replaced with the $\mathbf{L}_\infty$ norm, where

$$\|H\|_{\mathbf{L}_\infty} = \sup_{\omega\in\mathbb{R}} \sigma_{\max}(H(j\omega)).$$

Remark 8. The imaginary eigenvalues of $M_{\|H\|_\infty}$ are exactly the frequencies for which $\sigma_{\max}(H(j\omega)) = \|H\|_\infty$.

Remark 9. M_γ has imaginary eigenvalues if and only if M_γ^2 has real nonpositive eigenvalues.

3. A Bisection Algorithm

Theorem 2 suggests a bisection algorithm for computing $\|H\|_\infty$. Let γ_{lb} and γ_{ub} be some lower and upper bounds, respectively, on $\|H\|_\infty$. For example, we could use the bounds derived by Enns and Glover,

$$\gamma_{\rm lb} = \max\{\sigma_{\max}(D), \sigma_{\rm H1}\}, \qquad \gamma_{\rm ub} = \sigma_{\max}(D) + 2\sum_i \sigma_{{\rm H}i}, \tag{16}$$

where $\sigma_{{\rm H}i}$ are the Hankel singular values of the system $\{A, B, C, D\}$ [E], [G].

The bisection algorithm is as follows:

$\gamma_{\rm L} := \gamma_{\rm lb}$;
$\gamma_{\rm U} := \gamma_{\rm ub}$;
repeat {
  $\gamma := (\gamma_{\rm L} + \gamma_{\rm H})/2$;
  Form M_γ;
  if M_γ *has no imaginary eigenvalues,* $\gamma_{\rm H} := \gamma$,
  else $\gamma_{\rm L} = \gamma$; }
until { $\gamma_{\rm H} - \gamma_{\rm L} \le 2\varepsilon\gamma_{\rm L}$ }.

Note that we always have $\gamma_{\rm L} \le \|H\|_\infty \le \gamma_{\rm H}$; moreover, after M iterations, $\gamma_{\rm H} - \gamma_{\rm L} = 2^{-M}(\gamma_{\rm ub} - \gamma_{\rm lb})$. On exit, $(\gamma_{\rm L} + \gamma_{\rm H})/2$ is guaranteed to approximate $\|H\|_\infty$ within a relative accuracy of ε, that is,

$$|(\gamma_{\rm L} + \gamma_{\rm H})/2 - \|H\|_\infty| \le \varepsilon\|H\|_\infty.$$

Let us briefly discuss the computation involved in this algorithm. Enns' and Glover's bounds can be computed by solving the observability and controllability grammian Lyapunov equations [GNV],

$$\begin{aligned} A^T W_{\rm o} + W_{\rm o} A + C^T C &= 0, \\ A W_{\rm c} + W_{\rm c} A^T + B B^T &= 0 \end{aligned} \tag{17}$$

and then computing the eigenvalues of $W_{\rm c} W_{\rm o}$, which are the squares of the Hankel singular values. An alternative which avoids this eigenvalue computation is to set

$$\gamma_{\rm lb} = \max\{\sigma_{\max}(D), \sqrt{{\rm Tr}\ W_{\rm c} W_{\rm o}/n}\}, \qquad \gamma_{\rm ub} = \sigma_{\max}(D) + 2\sqrt{n\ {\rm Tr}\ W_{\rm c} W_{\rm o}}.$$

The work in computing R^{-1} and S^{-1} while forming M_γ in the $D \ne 0$ case is reduced to a minimum by the following initial transformation on $\{A, B, C, D\}$. Let $D = U\Sigma V^T$ be an SVD of D, where U is $p \times p$, V is $m \times m$ and orthogonal. Of course,

$$\|H\|_\infty = \|U^T H(s) V\|_\infty = \|U^T C(sI - A)^{-1} BV + \Sigma\|_\infty.$$

The bisection algorithm is then applied to the transformed system $\{A, BV, U^T C, \Sigma\}$. For this system, R and S are diagonal, so computing R^{-1} and S^{-1} is fast.

One method for determining whether M_γ has imaginary eigenvalues is simply to compute the eigenvalues and check. Special methods which exploit the Hamiltonian structure can be used; see [V1] and [B1]. Computing the eigenvalues by these methods require roughly as much computation as one SVD (several tens of n^3 flops), so that each iteration of the bisection is comparable to computing the SVD of $H(j\omega)$ at a *single* frequency. Thus the computational savings of the bisection method over an SV-plot method is very roughly the number of frequency samples in the plot divided by the number of bisection iterations, which is typically under 20.

4. A Sturm/Routh Test for Imaginary Eigenvalues

Needless to say, the eigenvalues of M_γ cannot be exactly computed in a finite number of steps, even with exact arithmetic computations. Nevertheless, it *is* possible to determine whether M_γ has imaginary eigenvalues in a *finite* number of steps, without actually computing the eigenvalues.

Let $r(s)$ be the characteristic polynomial of M_γ, that is, $r(s) = \det(sI - M_\gamma)$. Since M_γ is Hamiltonian, $r(s)$ is a polynomial in s^2, so $r(s) = p(-s^2)$ for some degree n polynomial p. p could be computed from M_γ by the Leverrier–Faddeeva algorithm [K, pp. 657–658]. M_γ has imaginary eigenvalues if and only if p has real nonnegative roots.

A *Sturm* method can be used to test whether p has real nonnegative roots (see, e.g., [H]). We assume that $p(0)$ is nonzero, since otherwise p has the nonnegative real root 0. First we apply the Euclid algorithm to p and p': let $p_n = p$, $p_{n-1} = p'$ and recursively divide:

$$p_n = q_1 p_{n-1} - p_{n-2},$$

$$p_{n-1} = q_2 p_{n-2} - p_{n-3},$$

$$\vdots$$

$$p_0 \text{ is constant.}$$

Then $p_n, p_{n-1}, \ldots, p_0$ is a *Sturm sequence*. Let v_∞ be the number of sign changes in the *leading* coefficients of $p_n, \ldots, p_0$, and let v_0 be the number of sign changes in the *constant* coefficients of $p_n, \ldots, p_0$.

Fact 1. *p has exactly $v_\infty - v_0$ distinct real nonnegative roots. In particular, M_γ has no imaginary eigenvalues if and only if $v_\infty = v_0$.*

This Sturm test can be implemented using a *Routh array* in which the divisions required in the Euclid algorithm are avoided (see [BS]).

5. Some Extensions

The results presented so far, and various generalizations, have many applications in computing other quantitites of system-theoretic interest; we mention three in this section. We show how a similar method can be used to compute the minimum dissipation of a transfer matrix, or the $\mathbf{H}_\infty$ norm of a transfer matrix over a restricted frequency range. Our third extension concerns numerical optimization—we establish a bisection method for a fast line search.

5.1. Minimum Dissipation

First we consider the *minimum dissipation*, $\operatorname{diss}(H)$, of a transfer matrix defined by

$$\operatorname{diss}(H) = \inf_{\operatorname{Re} s > 0} \lambda_{\min}((H(s) + H(s)^*)/2).$$

The one-parameter Hamiltonian matrix for the dissipation problem is

$$N_\delta = \begin{bmatrix} A & 0 \\ 0 & -A^T \end{bmatrix} + \begin{bmatrix} B \\ -C^T \end{bmatrix} (2\delta I - (D + D^T))^{-1} [C \quad B^T], \tag{18}$$

defined if and only if δ is not an eigenvalue of $(D + D^T)/2$.

The following theorem relates the eigenvalues of $(H(j\omega) + H(j\omega)^*)/2$ and the imaginary eigenvalues of N_δ.

Theorem 3. *Assume A has no imaginary eigenvalues, δ is not an eigenvalue of $(D + D^T)/2$, and $\omega_0 \in \mathbb{R}$. Then, δ is an eigenvalue of $(H(j\omega_0) + H(j\omega_0)^*)/2$ if and only if $(N_\delta - j\omega_0 I)$ is singular.*

The proof follows the pattern of the proof of Theorem 1 and is left to the reader. The analog of Theorem 2 is simply:

Theorem 4. *Let A be stable and $\delta < \lambda_{\min}((D + D^T)/2)$. Then $\operatorname{diss}(H) \le \delta$ if and only if N_δ has imaginary eigenvalues.*

A bisection algorithm for computing $\operatorname{diss}(H)$ is readily designed.

5.2. $\mathbf{H}_\infty$ *Norm Over an Interval*

The second extension we discuss concerns computing the maximum of the maximum singular value of the transfer matrix evaluated between two frequency limits. Suppose it is required to compute

$$\|H\|_{[\alpha,\beta]} = \sup_{\alpha \le \omega \le \beta} \sigma_{\max}(H(j\omega)),$$

where $0 \le \alpha < \beta$. The bisection algorithm presented in Section 3 is modified as follows: first, the *a priori* lower bound is changed to $\gamma_{\mathrm{lb}} = \max\{\sigma_{\max}(H(j\alpha)), \sigma_{\max}(H(j\beta))\}$; and, second, the eigenvalue test is modified to:

if M_γ has no imaginary eigenvalues between $j\alpha$ and $j\beta$, ...

The Sturm test presented in Section 4 (Fact 1) is readily modified to check whether a Hamiltonian matrix has any imaginary eigenvalues between $j\alpha$ and $j\beta$. Using the notation of Section 4, let $v_{\sqrt{\alpha}}$ denote the number of sign changes in the sequence of real numbers $p_n(\sqrt{\alpha}), p_{n-1}(\sqrt{\alpha}), \ldots, p_0$, and similarly let $v_{\sqrt{\beta}}$ be the number of sign changes in the sequence $p_n(\sqrt{\beta}), p_{n-1}(\sqrt{\beta}), \ldots, p_0$. Then we have:

Fact 2. *p has exactly $v_{\sqrt{\beta}} - v_{\sqrt{\alpha}}$ distinct real roots between $\sqrt{\alpha}$ and $\sqrt{\beta}$. In particular, M_γ has no imaginary eigenvalues between $j\alpha$ and $j\beta$ if and only if $v_{\sqrt{\beta}} = v_{\sqrt{\alpha}}$.*

We note that this result concerning $\|H\|_{[\alpha,\beta]}$ is not readily apparent from the spectral factorization formulation mentioned above in Remark 5.

5.3. *Fast Line Search*

Consider the problem of minimizing (or, at least, finding a local minimum of) the $\mathbf{H}_\infty$ norm of a system $\{A(a), B(a), C(a), D(a)\}$ which depends (differentiably) on a parameter $a \in \mathbb{R}^k$. In a standard *descent algorithm* [L2, Chapter 7], each iteration consists of two substeps: first, a *descent direction* $\delta a \in \mathbb{R}^k$ is determined, and, second, an appropriate *stepsize* $h \in \mathbb{R}$ is determined. The parameter a is then replaced by $a + h\delta a$, and the process is repeated. The process of choosing the stepsize is often called a line search, since it is equivalent to choosing the next parameter value somewhere along the line through a in the direction δa.

By a descent direction, we mean δa such that the function

$$\varphi(h) = \|(H(a + h\delta a, s))\|_\infty$$

has a right derivative at $h = 0$ which is negative: $\varphi'_+(0) < 0$.

One widely used method for determining an appropriate stepsize is the *Armijo rule* [L2, p. 212]. Usually, the line search involves many function evaluations, since $\varphi(h)$ is evaluated for many candidate stepsizes. We will demonstrate a simple bisection method which simultaneously computes an Armijo stepsize h^* and computes $\varphi(h^*)$ with no more computation than computing $\varphi(h)$ for a given h. Thus the $\mathbf{H}_\infty$ Armijo line-search problem can be solved with no more effort than a single function evaluation (i.e., an $\mathbf{H}_\infty$ norm computation).

Let us describe the Armijo rule. Consider the straight line defined by

$$\psi(h) = \varphi(0) + \varepsilon\varphi'_+(0)h, \tag{19}$$

where ε is a parameter chosen in $(0, 1)$ ($\varepsilon = 0.2$ is typical). Then h is an (exact) Armijo step length if $h > 0$ and $\varphi(h) = \psi(h)$, that is, it is a positive stepsize corresponding to one of the points of intersection of the line described by (19) and the function $\varphi(h)$. Thus, in Fig. 2, the exact Armijo step lengths are h_1, h_2, and h_3.

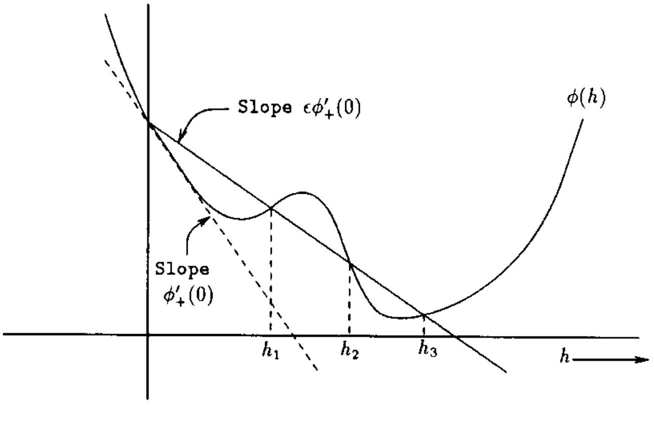

Fig. 2. Armijo rule for line search: exact Armijo step lengths.

We now present a bisection algorithm which determines an exact Armijo step length and the corresponding function value. The following is an immediate consequence of Theorem 2: if $M_{\psi(h)}$ has no imaginary eigenvalues, then $\psi(h) > \varphi(h)$; if $M_{\psi(h)}$ has imaginary eigenvalues, then $\psi(h) < \varphi(h)$.

Let

$$h_{\text{lb}} = 0, \qquad h_{\text{ub}} = -\varphi(0)/(\varepsilon\varphi'_+(0)).$$

Of course, $\psi(h_{\text{ub}}) < \varphi(h_{\text{ub}})$ and $h_{\text{lb}} = 0$ is not a valid Armijo stepsize.

The bisection algorithm is then:

$h_{\text{L}} := h_{\text{lb}}$;
$h_{\text{U}} := h_{\text{ub}}$;
repeat {
  $h := (h_{\text{L}} + h_{\text{H}})/2$;
  Form $M_{\psi(h)}$;
  if $M_{\psi(h)}$ *has no imaginary eigenvalues*, $h_{\text{L}} := h$,
  else $h_{\text{H}} = h$; }
until { $h_{\text{H}} - h_{\text{L}} \leq 2\varepsilon h_{\text{L}}$ }.

After k iterations we have $\psi(h_{\text{U}}) < \varphi(h_{\text{U}})$, $\psi(h_{\text{L}}) \geq \varphi(h_{\text{L}})$, and $h_{\text{U}} - h_{\text{L}} = 2^{-k}\varphi(0)/\varepsilon|\varphi'_+(0)|$. Thus there is always at least one exact Armijo stepsize between h_{L} and h_{U}. This algorithm computes an exact Armijo step length h^* with guaranteed accuracy ε. Of course we have also computed $\varphi(h^*)$, since $\varphi(h^*) = \psi(h^*) = \varphi(0) + \varepsilon\varphi'_+(0)h^*$.

If this bisection algorithm is applied to the φ shown above, we will compute $h^* = h_1$.

In this discussion we have ignored the requirement that $A(h)$ be stable, but that is readily incorporated into the bisection algorithm above.

6. A Simple Example

We present a very simple example where we compute the $\mathbf{H}_\infty$ norm of the system described by

$$A = \begin{bmatrix} -0.08 & 0.83 & 0 & 0 \\ -0.83 & -0.08 & 0 & 0 \\ 0 & 0 & -0.7 & 9 \\ 0 & 0 & -9 & -0.7 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & -1 \\ 0 & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 0.4 & 0 & 0.4 & 0 \\ 0.6 & 0 & 1 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 0.3 & 0 \\ 0 & -0.15 \end{bmatrix}.$$

An SV plot for this system (25 points per decade) is shown in Fig. 3. Enns' and Glover's bounds are $3.2022 \leq \|H\|_\infty \leq 15.1537$, and 17 bisections are required to compute $\|H\|_\infty \doteq 6.4405$ with a relative accuracy exceeding 10^{-5} (i.e., all digits correct). The maximum of the maximum singular values for the SV plot is 6.2126,

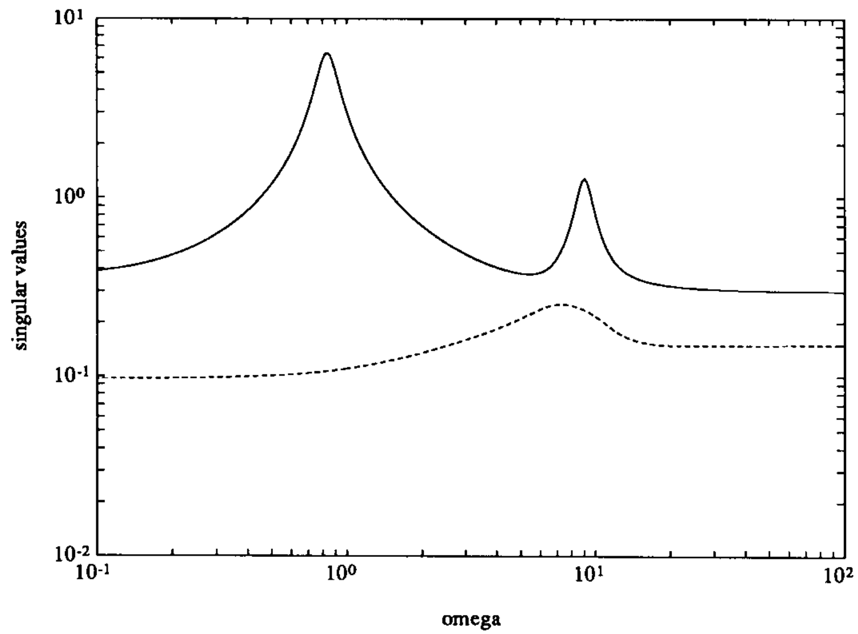

Fig. 3. An SV plot for the example system: the solid line is the maximum singular value and the dotted line is the minimum (second) singular value.

so the relative accuracy is only about 3.5%, even though this system does not have particularly high Q resonances.

References

[A] B. D. O. Anderson, An algebraic solution to the spectral factorization problem, *IEEE Trans. Automat. Control*, **12** (1967), 410–414.

[BS] S. Barnett and D. D. Siljak, Routh's algorithm: a centennial survey, *SIAM Rev.*, **19** (1977), 472–489.

[BD] S. Boyd and J. Doyle, Comparison of peak and RMS gains for discrete-time systems, *Systems Control Lett.*, **9** (1987), 1–6.

[B1] R. Byers, A Hamiltonian QR algorithm, *SIAM J. Sci. Statist. Comput.*, **7** (1986), 212–229.

[B2] R. Byers, A Bisection Algorithm for Measuring the Distance of a Stable Matrix to the Unstable Matrices, Technical Report, North Carolina State University at Raleigh, 1987; to appear in *SIAM J. Sci. Statist. Comput.*

[DV] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input–Output Properties*, Academic Press, New York, 1975.

[E] D. F. Enns, Model reduction with balanced realizations: an error bound and a frequency weighted generalization, *Proceedings of the 23rd IEEE Conference on Decision and Control*, Las Vegas, NV, 1984, pp. 127–132.

[F] B. A. Francis, *A Course in H_∞ Control Theory*, Lecture Notes in Control and Information Sciences, Vol. 88, Springer-Verlag, New York, 1987.

[G] K. Glover, All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds, *Internat. J. Control*, **39** (1984), 1115–1193.

[GNV] G. H. Golub, S. Nash, and C. Van Loan, A Hessenberg–Schur method for the matrix problem $AX + XB = C$, *IEEE Trans. Automat. Control*, **24** (1979), 909–913.

[GV] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 1983.

[H] P. Henrici, *Applied and Computational Complex Analysis*, Vol. 1, Wiley-Interscience, New York, 1978.

[K] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[L1] A. J. Laub, Efficient calculation of frequency response matrices from state space models, *ACM Trans. Math. Software*, **12** (1986), 26–33.

[L2] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd edn., Addison-Wesley, Reading, MA, 1984.

[V1] C. F. Van Loan, A symplectic method for approximating all eigenvalues of a Hamiltonian matrix, *Linear Algebra Appl.*, **61** (1984), 233–251.

[V2] C. F. Van Loan, How near is a stable matrix to an unstable one? *Contemp. Math.*, **47** (1985), 465–478.

[V3] M. Vidyasagar, *Control System Synthesis: A Factorization Approach*, MIT Press, Cambridge, MA, 1985.

[W] J. C. Willems, Least squares stationary optimal control and the algebraic Riccati equation, *IEEE Trans. Automat. Control*, **16** (1971), 621–634.

[ZF] G. Zames and B. A. Francis, Feedback, minimax sensitivity, and optimal robustness, *IEEE Trans. Automat. Control*, **28** (1983), 585–601.