2

Control System Radii and Robustness Under Approximation

John A. Burns *[1] and Gunther H. Peichl **[2]

[1] Center for Optimal Design and Control
Interdisciplinary Center for Applied Mathematics
Virginia Tech
Blacksburg, VA 24061-0531
burns@icam.vt.edu

[2] Institut für Mathematik und Wissenschaftliches Rechnen
Karl-Franzens-Universität Graz
A-8010 Graz, Austria
gunther.peichl@uni-graz.at

Summary. The purpose of this paper is twofold. First, we provide a short review and summarize results on the robustness of controllability and stabilizability for finite dimensional control problems. We discuss the computation of system radii which provide a measure of robustness. Second, we consider systems which arise as finite difference and finite element approximations to control systems defined by partial differential equations. In particular, we derive controllability criteria for approximations of the controlled heat equation which are easy to check numerically. For a particular example we establish tight theoretical upper and lower bounds on the controllability radii for the finite difference and finite element models and compare these bounds with numerical results. Finally, we present numerical results on stabilizability radii which suggests that conditioning of the LQR control problem may be measured by this radii.

Key words: control system radii, numerical methods, robustness

Introduction

The analysis of mathematical models used in control design and optimization often requires several stages of approximation. Also, in the area of distributed parameter

* This research was supported in part by the Air Force Office of Scientific Research under grant F49620-02-C-0048 and the Special Projects Office at DARPA. Parts of the research were carried out while the author was a visitor at the Institut für Mathematik und Wissenschaftliches Rechnen at Karl-Franzens-Universität Graz, Austria, supported by the Fonds zur Förderung der wissenschaftlichen Forschung, Austria, under project SFB03-18 "Optimierung und Kontrolle".

** This research was supported in part by the Fonds zur Förderung der wissenschaftlichen Forschung under P7522-PHY, P8146-PHY, and SFB03-18 "Optimierung und Kontrolle".

(DP) control, numerical approximation must be introduced at some point in the modelling process. Finite element, Galerkin and finite difference schemes are typically used to "discretize" continuum models, while in the frequency domain one might construct rational approximations of non-rational transfer functions. For computing purposes, state space models offer certain advantages in that there are numerous computational algorithms well suited for the matrix-linear algebra problems that occur in control design. Direct discretization of continuum models usually produce state space models as do frequency domain approximations (followed by realization schemes) and model reduction methods such as proper orthogonal decomposition (POD). It is fair to say that all approaches have advantages and disadvantages and each approach leads to its own characteristic set of problems. However, to achieve robustness in a design based on approximate models one needs to take into account the robustness of the approximate model with respect to system properties.

Given that there are several approaches to constructing finite dimensional state space models, it is reasonable to ask if there is some "measure" that can be used to select the "best" approach for a given system with a specific control design objective. In order to study this problem, it is clear that one must first decide what criteria will be used to determine which state space model is "best" for the particular problem at hand. It is very important to remember that such criteria may change it the system changes, if the control design objective changes or if the numerical method used to solve the corresponding control problem changes. Since the finite dimensional approximate/reduced order model will be used to design and optimize the infinite dimensional system, it is important that the finite dimensional model inherits the essential control system properties and that the finite dimensional control problem is numerically well-conditioned. For example, it has been observed [6; 7] that numerical conditioning of the LQR problem can be negatively influenced when non-uniform meshes are used to approximate a DP system governed by a partial differential equation.

In this paper we investigate these ideas for distributed parameter systems. We shall focus on a specific subset of these problems. Our goal is to illustrate how one can use system measures to aid in the selection of model reduction and discretization algorithms. In particular, we shall use the concept of "system radii" to measure the "quality" of finite dimensional state space models constructed by direct discretization of continuum models. The motivation for this choice lies in the observation that numerical algorithms for control design can be (numerically) unstable if applied to systems that are not controllable (observable, stabilizable, etc.). Moreover, numerical ill-conditioning can result even if the system is controllable and observable but "near" an uncontrollable or unobservable system. This idea is certainly not new and there exist many examples of this type. Demmel [13] has developed a rather nice theory of ill-conditioning and established that numerous problems in numerical linear algebra (matrix inversion, eigenvalue calculations) and control design (pole-placement, robust control) all become ill-conditioned if the state space models used in the calculations are close to an ill-posed problem. Laub and his co-workers have established similar results for the LQR problem [18; 29; 30]. Since one of the often noted "advantages" of state space models is their usefulness for computational purposes, it is reasonable to use the condition number of the problem as one measure to help select a discretization scheme. One can find a nice presentation of these ideas in the recent book [12] by Datta.

Although there are several issues that need to be addressed in the overall approximation process, we shall limit most of our discussion to the study of system radii for systems that typically occur when partial differential equations are discretized by finite element and finite difference schemes. These finite dimensional systems often have nice symmetry properties that can be exploited in the computation of the system radii. The basic problem of preserving system properties under approximation has been addressed by other authors [4; 17; 32] and is crucial to any method. However, we concentrate on the problem of selecting a "good" approximation from the class of all schemes that preserve the appropriate system properties.

The paper is organized as follows. In Section 2.1 we review the basic definitions of system (radii) measures for finite dimensional systems and present examples to illustrate some relationships between these measures and typical control problems. We also summarize a few known results concerning these measures and give some new results on computing these measures. In Section 2.2 we discuss the problem of approximating infinite dimensional systems and use finite element and finite difference approximations of the heat equation to illustrate the ideas. This simple example is rich enough to provide some indication of the difficulties one encounters in developing theoretical and computational results for such problems. In Section 2.3 we provide a case study and compare theoretical bounds to computed values. Finally, we close with a short summary and a simple numerical example to illustrate the potential use of system radii to estimate the numerical condition number of an LQR problem.

2.1 Measures of Robustness

As noted in [1] numerical algorithms which assume a specific system property such as controllability or stabilizability can be expected to be numerically ill-conditioned if the system model is nearly uncontrollable (or nearly unstabilizable). The following simple example illustrates the type of difficulties that one can encounter.

Motivating Example. Consider the control system governed by the second order system

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} \delta \\ \epsilon \end{bmatrix} u(t).$$

Observe that this system is controllable (and stabilizable) if and only if $\delta \neq 0$ and $\epsilon \neq 0$. It becomes uncontrollable if $\epsilon = 0$ and unstabilizable if and only if $\delta = 0$. Moreover, the system becomes "nearly uncontrollable" as $\epsilon \longrightarrow 0$ (and "nearly unstabilizable" as $\delta \longrightarrow 0$) in the sense that a perturbation of order ϵ (δ) in the input matrix $\mathrm{col}(\delta, \epsilon)$ may result in an uncontrollable (unstabilizable) system.

In order to demonstrate the effect of near unstabilizability on control design, we consider the problem of minimizing the quadratic functional

$$J(u) = \int_0^{+\infty} [(x_1(t))^2 + (x_2(t))^2 + u^2(t)]dt.$$

The optimal feedback gain is given by

$$k^*(\delta, \epsilon) = [k_1^*, k_2^*] = -[\delta, \epsilon]\Pi(\delta, \epsilon)$$

where $\Pi(\delta, \epsilon)$ is the solution to the Riccati equation

$$A^*\Pi + \Pi A - \Pi B r^{-1} B^* \Pi + Q = 0,$$

and $Q = I_2$ and $r = 1$.

It is straightforward to show that

$$\Pi(\delta, \epsilon) = \begin{bmatrix} \frac{2+\epsilon^2+2\sqrt{1+\epsilon^2+\delta^2}}{2\delta^2} & \frac{-\epsilon}{2\delta} \\ \frac{-\epsilon}{2\delta} & \frac{1}{2} \end{bmatrix}$$

and hence

$$k^*(\delta, \epsilon) = -\Big[\frac{1+\sqrt{1+\epsilon^2+\delta^2}}{\delta}, 0\Big].$$

Note that as $\delta \to 0$, $\|\Pi(\delta, \epsilon)\| \to +\infty$ and $\|k^*\| \to +\infty$. Here, $\|\cdot\|$ denotes any suitable matrix or vector norm. Thus, as the system approaches an unstabilizable system, the Riccati equation becomes ill-conditioned. As expected, the conditioning of the Riccati equation is not affected by the loss of controllability. However, consider the problem of finding a feedback operator $k_p(\delta, \epsilon) = [k_p^1, k_p^2]$ that places the closed-loop poles at -2 and -4. In particular, if $\delta\epsilon \neq 0$ then the unique solution to this problem exists and is given by

$$k_p = \left[\frac{-15}{2\delta}, \frac{3}{2\epsilon}\right].$$

Observe that as δ or ϵ approach 0, the system becomes nearly uncontrollable and $\|k_p\| \to +\infty$.

The previous example illustrates the need for a device to measure nearness of a system to uncontrollability, respectively unstabilizability. In order to make these ideas precise we introduce the following notation. We identify the control system Σ

$$\dot{x} = Ax + Bu, \qquad (\Sigma)$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $m \leq n$, with the matrix $[A, B] \in \mathbb{R}^{n\times(n+m)}$. For any $\lambda \in \mathbb{C}$ we introduce

$$H(\lambda) = [A - \lambda I, B] \in \mathbb{C}^{n\times(n+m)}.$$

The Hautus - test (see [28]) for controllability is based on embedding $[A, B]$ in the set of complex systems

$$\Gamma = \{[A, B] : A \in \mathbb{C}^{n\times n}, B \in \mathbb{C}^{n\times m}\}.$$

The distance between two systems Σ_1, Σ_2 is defined by

$$\delta(\Sigma_1, \Sigma_2) = \|[A_1 - A_2, B_1 - B_2]\|,$$

where $\|\cdot\|$ denotes any suitable matrix norm on $\mathbb{C}^{n\times(n+m)}$. Given $\Sigma \in \Gamma$ and a subset $S \subset \Gamma$ the distance between Σ and S is defined by

$$d(\Sigma, S) = \inf\{\delta(\Sigma, \Sigma_\alpha) : \Sigma_\alpha \in S\}.$$

Let $N_c \subset \Gamma$ be the set of all complex systems that are not controllable and $N_s \subset \Gamma$ be the set of all complex systems that are not stabilizable, i.e.

$$N_c = \{[A, B] \in \Gamma : [A, B] \text{ is not controllable}\},$$
$$N_s = \{[A, B] \in \Gamma : [A, B] \text{ is not stabilizable}\}.$$

Given $\Sigma \in \Gamma$ one defines the measure of controllability by

$$\gamma_c = d(\Sigma, N_c) \tag{2.1}$$

and the measure of stabilizability by

$$\gamma_c = d(\Sigma, N_s). \tag{2.2}$$

These definitions may be found in [14; 29; 34]. There are several reasons that these measures are useful. First, they can provide an estimate of the condition number for control design algorithms (see [13] and the example on robust pole placement therein). Moreover, they provide numerical bounds on the errors that can be tolerated in the data defining the system matrices A and B.

To obtain more explicit formulae for γ_c and γ_s the norm in $\mathbb{C}^{n\times(n+m)}$ has to be related to the norms in $\mathbb{C}^n$ and $\mathbb{C}^{n+m}$. We shall modify a concept which was used in [21] to calculate the stability radius of a matrix. If we choose the Euclidean norm in $\mathbb{C}^n$ and in $\mathbb{C}^{n+m}$, and the spectral norm in $\mathbb{C}^{n\times(n+m)}$, then γ_c and γ_s are determined by the singular values of $H(\lambda)$. The singular values $\sigma_1, \ldots, \sigma_p$, $p = \min(r, s)$ of a matrix $H \in \mathbb{C}^{r\times s}$ (see [19] for basic definitions) will be ordered in the standard fashion $\sigma_1 \geq \cdots \geq \sigma_p$. We shall also use $\sigma_{min}(H)$ to denote the smallest singular value σ_p of a matrix H.

Definition 1. *Let $\|\cdot\|_n$, $\|\cdot\|_m$ be norms on $\mathbb{C}^n$ and $\mathbb{C}^m$ respectively and let $\|\cdot\|_n^*$ ($\|\cdot\|_m^*$) denote the norm dual to $\|\cdot\|_n$ ($\|\cdot\|_m$). A matrix norm $\|\cdot\|_{n,m}$ on $\mathbb{C}^{n\times m}$ is said to be strongly compatible with $\|\cdot\|_n^*$ and $\|\cdot\|_m^*$ if the following two conditions hold*

(C1) $\|x^*A\|_m^* \leq \|A\|_{n,m}\|x^*\|_n^*$ *for all* $A \in \mathbb{C}^{n\times m}$, $x^* \in (\mathbb{C}^n)^*$.
(C2) For any pair of vectors $x^* \in (\mathbb{C}^n)^*$, $x^* \neq 0$, $y^* \in (\mathbb{C}^m)^*$ *there exists* $H \in \mathbb{C}^{n\times m}$ *satisfying*

$$y^* = x^*H \quad \text{and} \quad \|H\|_{n,m}\|x^*\|_n^* = \|y^*\|_m^*.$$

In the above definition $(\mathbb{C}^n)^*$ denotes the dual space of $\mathbb{C}^n$ which is (algebraically) identified with $\mathbb{C}^{1\times n}$. Easy modifications of the proofs in [21] establish that the operator norms on $\mathbb{C}^{n\times m}$ as well as the Hölder norms are strongly compatible with $\|\cdot\|_n^*$ and $\|\cdot\|_m^*$. In particular, this implies that the spectral norm and the Frobenius norm (see [19]) are strongly compatible with the Euclidean norms in $\mathbb{C}^n$ and $\mathbb{C}^m$.

Theorem 1. *Choose any vector norms $\|\cdot\|_n$, $\|\cdot\|_{n+m}$ in $\mathbb{C}^n$ and $\mathbb{C}^{n+m}$ and any matrix norm in $\mathbb{C}^{n\times(n+m)}$ strongly compatible with $\|\cdot\|_n^*$ and $\|\cdot\|_{n+m}^*$. If the system $[A,B]\in\Gamma$ is controllable, then the measure of controllability is given by*

$$\gamma_s = \min_{\lambda\in\mathbb{C}} \min_{\substack{\|x\|_n^*=1\\ x^*\in(\mathbb{C}^n)^*}} \|x^*\,[A-\lambda I, B]\|_{n+m}^*\,. \tag{2.3}$$

Proof. Let α denote the number on the righthand side of (2.3) and let $[\delta A_0,\delta B_0]\in\Gamma$ satisfy $\|[\delta A_0,\delta B_0]\|_{n,n+m}=\gamma_c$ and $[A+\delta A_0, B+\delta B_0]\in N_c$. Hence, there exists $\lambda\in\mathbb{C}$, $x^*\in(\mathbb{C}^n)^*$, $\|x^*\|_n^*=1$ with

$$x^*[A+\delta A_0-\lambda I, B+\delta B_0]=0.$$

Using (C1) this implies

$$\|x^*[A-\lambda I,B]\|_{n+m}^* \le \|x^*\|_n^*\|[\delta A_0,\delta B_0]\|_{n,n+m}$$

and a fortiori

$$\alpha\le\gamma_c.$$

On the other hand one can argue the existence of $\lambda_0\in\mathbb{C}$, $x_0^*\in(\mathbb{C}^n)^*$, $\|x_0^*\|_n^*=1$ with

$$\alpha=\|x_0^*[A-\lambda_0 I,B]\|_{n+m}^*.$$

Condition (C2) applied to $x^*=x_0^*$, $y^*=x_0^*[A-\lambda_0 I,B]$ ensures the existence of $[\delta\hat{A},\delta\hat{B}]\in\Gamma$ with

$$x_0^*[A-\lambda_0 I,B]=x_0^*[\delta\hat{A},\delta\hat{B}] \tag{2.4}$$

$$\|[\delta\hat{A},\delta\hat{B}]\|_{n+m}\cdot\|x_0^*\|_n^*=\|x_0^*[A-\lambda_0 I,B]\|_{n+m}^*. \tag{2.5}$$

The identity (2.4) shows that $[A-\delta\hat{A},B-\delta\hat{B}]\in N_c$, and (2.5) implies that

$$\gamma_c\le\alpha.$$

□

Corollary 1. *Let the norms be chosen as in Theorem 1. If the system $[A,B]\in\Gamma$ is stabilizable, then the measure of stabilizability is given by*

$$\gamma_s = \min_{\substack{\lambda\in\mathbb{C}\\ 0\le Re\lambda}} \min_{\substack{\|x\|_n^*=1\\ x^*\in(\mathbb{C}^n)^*}} \|x^*\,[A-\lambda I, B]\|_{n+m}^*\,. \tag{2.6}$$

Corollary 2. *Choose the Euclidean norms in $\mathbb{C}^n$ and $\mathbb{C}^{n+m}$ and the spectral norm (or Frobenius norm) in $\mathbb{C}^{n\times(n+m)}$.*

i) If $[A, B] \in \Gamma$ is controllable, then

$$\gamma_c = \min_{\lambda\in\mathbb{C}} \sigma_{\min}(H(\lambda)). \tag{2.7}$$

ii) If $[A, B] \in \Gamma$ is stabilizable, then

$$\gamma_s = \min_{Re\lambda\geq 0} \sigma_{\min}(H(\lambda)). \tag{2.8}$$

The characterizations (2.7) and (2.8) were first derived in [14] using a different argument.

Throughout the remaining part of this chapter we shall use the Euclidean norms in $\mathbb{C}^n$ and $\mathbb{C}^{n+m}$ and the Frobenius norm $\|\cdot\|_F$ in $\mathbb{C}^{n\times(n+m)}$. In [15; 44; 45] an alternative solution to (2.1) was given.

Theorem 2. *If $[A, B] \in \Gamma$ is controllable, then*

$$\gamma_c = \min_{\substack{\|q\|=1 \\ q\in\mathbb{C}^n}} \|q^* [A(I - qq^*, B]\| . \tag{2.9}$$

Although (2.9) is equivalent to an optimization problem given in [15], the approach to establish (2.9) given in [44; 45] is entirely different. The minimal perturbation $[\delta A_0, \delta B_0]$ is determined by using a state transformation that produces a certain canonical form. In [44; 45] the identity (2.9) is established by showing its equivalence with (2.7). For completeness we want to give a direct proof based on the above motivation.

Proof of Theorem 2. Let $[\delta A, \delta B] \in \Gamma$ satisfy $[A + \delta A, B + \delta B] \in N_c$. Then there exists $Q = [Q_1, q] \in \mathbb{C}^n$, $Q_1 \in \mathbb{C}^{n,n-1}$, $q \in \mathbb{C}^n$, $Q^*Q = I$ such that in the transformed system

$$Q^*(A+\delta A)Q = \begin{pmatrix} Q_1^*(A+\delta A)Q_1 & Q_1^*(A+\delta A)q \\ q^*(A+\delta A)Q_1 & q^*(A+\delta A)q \end{pmatrix},$$

$$Q^*(B+\delta B) = \begin{pmatrix} Q_1^*(B+\delta B) \\ q^*(B+\delta B) \end{pmatrix},$$

$$q^*(A+\delta A)Q_1 = 0 \quad \text{and} \quad q^*(B+\delta B) = 0 \tag{2.10}$$

hold. Fix $q \in \mathbb{C}^n$. The minimal norm perturbation δB, in (2.10) is given by

$$\delta B_0 = -qq^*B.$$

Since the columns of Q_1 are orthogonal, the first equation in (2.10) implies the existence of $\lambda \in \mathbb{C}$ with

$$q^*(A + \delta A) = \lambda q^*$$

which gives the minimal norm perturbation δA_0

$$\delta A_0 = qq^*\lambda - qq^*A.$$

Also $\|\delta A_0\|_F = \|q\|_2\|q^*\lambda - q^*A\|_2$ minimizing $\|q^*\lambda - q^*A\|_2^2$ with respect to λ so that the minimum is attained at $\lambda_0 = q^*Aq$. Hence, the minimum norm perturbation δA_0 is given by

$$\delta A_0 = qq^*(q^*Aq) - qq^*A = q(q^*Aq)q^* - qq^*A = -qq^*A(I - qq^*).$$

Thus we obtain

$$[\delta A_0, \delta B_0] = -qq^*[A(I - qq^*), B].$$

and therefore

$$\|[\delta A_0, \delta B_0]\|_F = \|q^*[A(I - qq^*), B]\|_2.$$

Minimizing with respect to q, yields the desired result. □

Remark. As a consequence of the above proof we note that if $q_0 \in \mathbb{C}^n$ is optimal in (2.9), then

$$[\delta A_0, \delta B_0] = -q_0 q_0^*[A(I - q_0 q_0^*), B]$$

yields a minimal norm perturbation of $[A, B]$ destroying controllability. It is shown in [44, Theorem 4.3] that if $\lambda_0 \in \mathbb{C}$ is optimal in (2.7) and $U = [u_1, \ldots, u_n] \in \mathbb{C}^{n\times n}$, $V = [v_1, \ldots, v_{n+m}] \in \mathbb{C}^{(n+m)\times(n+m)}$ determine a singular value decomposition of $H(\lambda_0)$, then $q_0 = u_n$ minimizes (2.9) and

$$[\delta A_0, \delta B_0] = \gamma_c u_n v_n^*, \quad \lambda_0 = u_n^* A u_n. \tag{2.11}$$

holds. Conversely if q_0 is a minimizer for (2.9) then $\lambda_0 = q_0^* A q_0$ minimizes (2.7) having q_0 as a left singular vector. In particular (2.11) reveals that for real systems $[A, B] \in \mathbb{R}^{n\times(n+m)}$ the closest uncontrollable (unstabilizable) system will in general be complex.

In order to study the effect of real perturbations, real measures have been introduced (see [1; 15; 45; 18]). In particular, let

$$\Omega = \{\Sigma = [A, B] : A \in \mathbb{R}^{n\times n},\ B \in \mathbb{R}^{n\times m}\}$$

denote the set of real systems. Then one defines the real measure of controllability by

$$\omega_c = d(\Sigma, N_c \cap \Omega)$$

and the real measure of stabilizability by

$$\omega_s = d(\Sigma, N_s \cap \Omega).$$

In general it is clear that

$$\gamma_c \leq \omega_c \quad \text{and} \quad \gamma_s \leq \omega_s$$
$$\gamma_c \leq \gamma_s \quad \text{and} \quad \omega_c \leq \omega_s$$

hold and it can happen that there is a significant difference between the various measures. It is tempting to assume that ω_c could be found by computing the quantity

$$\omega_{c,1} = \min_{\lambda \in \mathbb{R}} \sigma_{\min}(H(\lambda)). \tag{2.12}$$

In general, however, $w_{c,1}$ yields just an upper bound for w_c. This is a consequence of the following theorem which was established in [45]:

Theorem 3. *If* $[A, B] \in \Omega$ *is controllable, then*

$$\omega_c = \min(\omega_{c,1}, \omega_{c,2}),$$

where $\omega_{c,1}$ *is given by* (2.12) *and* $\omega_{c,2}$ *is defined by*

$$\omega_{c,2} = \min_{\substack{Q \in \mathbb{R}^{n \times 2} \\ Q^T Q = I_n}} \left\| Q^T \left[A - QQ^T, Q^T B \right] \right\|_F. \tag{2.13}$$

Corollary 3. *If* $[A, B] \in \Omega$ *is controllable and* $n = 2$*, then*

$$\omega_c = \min(w_{c,1}, \|B\|_F).$$

The proof of this result comes from the observation that $QQ^T = I_2$ holds for $n = 2$. We illustrate the above discussion by means of the following example.

Example 1. Let $\varepsilon < 2$ and define

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \varepsilon \\ 0 \end{pmatrix}.$$

It follows that $\omega_{c,1} = 1$, hence $\omega_c = \varepsilon$ and $[\delta A, \delta B] = [0, -B]$ is a real minimal norm perturbation destroying controllability. A short calculation shows that $\gamma_c = \varepsilon\sqrt{\frac{1}{2} - \frac{\varepsilon^2}{16}}$, the minimum in (2.7) being attained at $\lambda_0 = \pm i\sqrt{1 - \frac{\varepsilon^4}{16}}$. Observe that although $\omega_c = \omega_{c,2}$, it does not necessarily follow that the corresponding minimal norm perturbation destroying controllability $[\delta A, \delta B] = [QQ^T A(I - QQ^T), QQ^T B]$ has rank 2. This should be kept in mind in interpreting the corresponding results in [18]. For a more detailed discussion of ω_c we refer to [18; 45]. We complement these results by a characterization of the equality $\omega_c = \gamma_c$ which is adapted from an analogous result concerning the calculation of stability radii in [33]. (In the next two results we use the spectral norm in $\mathbb{C}^{n \times (n+m)}$.)

Theorem 4. *Let* $[A, B] \in \Omega$ *be controllable and define*

$$\Lambda_0 = \{\lambda^* \in \mathbb{C} : \sigma_{\min}(H(\lambda^*)) = \gamma_c\}.$$

Then $\gamma_c = \omega_c$ *holds if and only if there exist* $\lambda_0 \in \Lambda_0$ *and a singular value decomposition of* $H(\lambda_0)$*,*

$$H(\lambda_0) = \sum_{i=1}^{n} \sigma_i u_i v_i^*,\ u_i \in \mathbb{C}^n,\ v_i \in \mathbb{C}^{n+m},\ i = 1, \dots, n \tag{2.14}$$

satisfying

$$u_n^T u_n = v_n^T v_n. \tag{2.15}$$

Proof. First we show the necessity of (2.15). It suffices to discuss $\Lambda_0 \cap \mathbb{R} = \varnothing$. Assume that $\gamma_c = \omega_c$ holds and that the minimum in (2.7) is attained for some $\lambda_0 \in \mathbb{C} \setminus \mathbb{R}$. Hence, there is a minimal norm (with respect to the spectral norm) real perturbation $E \in \mathbb{R}^{n \times (n+m)}$ of $[A, B]$ satisfying

$$\|E\|_2 = \gamma_c = \sigma_n = \sigma_{\min} H(\lambda_0).$$

Thus, using (2.11) E admits the representation

$$E = -\sigma_n u_n v_n^*. \tag{2.16}$$

It follows that

$$u_n^*([A, B] + E) = u_n^* H(\lambda_0) + u_n^* E + \lambda_0 u_n^* [I, 0] = \lambda_0 u_n^* [I, 0].$$

Multiplying on the right by $\bar{v}_n$ and taking complex conjugates (note that $E \in \mathbb{R}^{n \times (n+m)}$ we arrive at

$$u_n^T([A, B] + E) v_n = \bar{\lambda}_0\, u_n^T [I, 0] v_n. \tag{2.17}$$

Similarly, starting with $([A, B] + E) v_n$ it follows that

$$u_n^T([A, B] + E) v_n = \lambda_0\, u_n^T [I, 0] v_n. \tag{2.18}$$

Comparing (2.17) and (2.18) we obtain

$$u_n^T [I, 0] v_n = 0. \tag{2.19}$$

Although the remaining part of the proof is identical to the one in [33] we present it here for the sake of completeness. The decomposition (2.14) together with (2.19) imply

$$u_n^T [A, B] v_n = \sigma_n u_n^T u_n,$$

while

$$[A, B] - \bar{\lambda}_0 [I, 0] = \sum_{i=1}^{n} \sigma_i \bar{u}_i v_i^T,$$

the complex conjugate of (2.14), combined with (2.19) yields

$$u_n^T[A,B]v_n = \sigma_n v_n^T v_n.$$

This completes the necessity part of the proof. Sufficiency of (2.15) may be shown in exactly the same way as in [33]. □

A consequence of Theorem 4 is the following easily checked sufficient condition for $\gamma_c = \omega_c$ to hold. We present it for the sake of completeness and refer to [33] for the proof.

Proposition 1. *Let $[A,B] \in \Omega$ be controllable and Λ_0 be as defined in Theorem 4. If for some $\lambda_0 \in \Lambda_0$ $\sigma_n = \sigma_{n-1}$ holds in the SVD* (2.14) *of $H(\lambda_0)$, then the real and complex controllability measures coincide, i.e. $\omega_c = \gamma_c$.*

It is apparent from Corollary 2, Theorem 2 and Theorem 3 that computing these measures is a difficult numerical problem and various algorithms have been developed for this purpose. Most of them rely on Corollary 2 (see [4; 44; 18]). For an alternative approach based on Theorems 2 and 3 see [45]. In order to reduce the required computational effort in minimizing (2.7) it is certainly advantageous to have a priori information on the location of the minimizing frequency λ^*.

Theorem 5. *Assume that $[A,B] \in \Gamma$ is controllable.*

(i) If $A = A^*$, then

$$\gamma_c = \min_{\lambda \in \mathbb{R}} \sigma_{\min}(H(\lambda)).$$

(ii) If $A = -A^*$, then

$$\gamma_c = \min_{\lambda \in \mathbb{R}} \sigma_{\min}(H(i\lambda)).$$

(iii) If $[A,B] \in \Omega$ and $A = A^T$, then

$$\gamma_c = \omega_c = \min_{\lambda \in \mathbb{R}} \sigma_{\min}(H(\lambda)).$$

Proof. (i) The square of $\sigma_{\min}[H(\lambda)]$ is equal to the minimum eigenvalue of $H(\lambda)H(\lambda)^*$. If $\lambda = \alpha + i\beta$, then

$$\begin{aligned} H(\lambda)H(\lambda)^* &= [A - \lambda I, B]\begin{bmatrix} A^* - \overline{\lambda} I \\ B^* \end{bmatrix} \\ &= I|\lambda|^2 - 2\mathrm{Re}(\lambda)A + A^2 + BB^* \\ &= \{\alpha^2 I - 2\alpha A + A^2 + BB^*\} + \beta^2 I \\ &= \{(\alpha I - A)^2 + BB^*\} + \beta^2 I. \end{aligned}$$

The Hermitian matrix $G(\alpha) = \{(\alpha I - A)^2 + BB^*\}$ has real eigenvalues $\lambda_1(\alpha)$, $\lambda_2(\alpha), \dots, \lambda_n(\alpha)$ and the spectral theorem [31, p. 312] implies that the eigenvalues of

$$H(\lambda)H^*(\lambda) = G(\alpha) + \beta^2 I$$

are given by $\lambda_i(\alpha)+\beta^2, i = 1, 2, \ldots, n$. Therefore, for each $\lambda = \alpha+i\beta$, the minimum eigenvalue of $H(\lambda)H^*(\lambda)$ occurs at $\beta = 0$ and hence

$$\gamma_c = \min_{\lambda\in\mathbb{C}} \sigma_{\min}[H(\lambda)] = \min_{\lambda\in\mathbb{R}} \sigma_{\min}[H(\lambda)]$$

which completes the proof of (i).

The proof of (ii) follows from (i) by observing that $A = -A^*$ if and only if $[iA] = [iA]^*$.

If $[A, B] \in \Omega$ and $A = A^T$, then part (i) implies that

$$\gamma_c = \min_{\lambda\in\mathbb{R}} \sigma_{\min}[H(\lambda)] = \sigma_{\min}[H(\hat{\lambda})]$$

for some real $\hat{\lambda}$. Since $H(\hat{\lambda})$ is real, the singular vectors of $H(\lambda)$ are real and hence the minimum norm rank reducing perturbation is real (see [19, p. 19]). Therefore, $\gamma_c = \omega_c$, part (iii) is established and this completes the proof. □

We note that Theorem 5 could be deduced exploiting (2.11). However, the proof presented above also establishes the following result.

Theorem 6. *Assume that $[A, B] \in \Gamma$ is stabilizable.*

(i) If $A = A^*$, then

$$\gamma_s = \min_{\lambda\geq 0} \sigma_{\min}(H(\lambda)).$$

(ii) If $A = -A^*$, then

$$\gamma_s = \min_{\lambda\in\mathbb{R}} \sigma_{\min}(H(i\lambda)).$$

(iii)If $[A, B] \in \Omega$ and $A = A^T$, then

$$\gamma_s = \omega_s = \min_{\lambda\geq 0} \sigma_{\min}(H(\lambda)).$$

Corollary 4. *If $[A, B] \in \Omega$ is stabilizable, $A = A^T$ and $x^T Ax \leq 0$ for all $x \in \mathbb{R}^n$, then*

$$\omega_s = \gamma_s = \sigma_{\min}(H(0)) = \sigma_{\min}([A, B]).$$

Proof. It follows from Theorem 6 that

$$\omega_s = \gamma_s = \min_{\lambda\in\mathbf{R}^+} \sigma_{\min}[H(\lambda)]$$

so we need only show that $\lambda^* = 0$ provides such a minimum. Since $(\lambda I - A)^2 + BB^T$ is positive semi-definite and $\sigma^2_{\min}[H(\lambda)]$ is the minimum eigenvalue $\hat{\sigma}(\lambda)$ of

$$H(\lambda)H^T(\lambda) = (A - \lambda I)^2 + BB^T,$$

it follows that

$$\begin{aligned}\hat{\sigma}(\lambda) &= \min_{\|x\|=1}[x^T(A-\lambda I)^2x + x^TBB^Tx] \\ &\geq \min_{\|x\|=1}[x^TA^2x + x^TBB^Tx] + \min_{\|x\|=1}[\lambda^2\|x\|^2 - 2\lambda x^TAx].\end{aligned}$$

The last term in this inequality is non-negative so that

$$\hat{\sigma}(\lambda) \geq \min_{\|x\|=1}[x^T(A^2+BB^T)x] = \hat{\sigma}(0)$$

and this completes the proof. □

We conclude this section with an example demonstrating that symmetry of A is sufficient for $\gamma_c = \min_{\lambda\in\mathbb{R}}\sigma_{\min}(H(\lambda))$ to hold, but the symmetry of A is not necessary.

Example 2. Let $\varepsilon \leq \frac{1}{2}$ and

$$A = \begin{pmatrix} 1 & 0 \\ \varepsilon & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

It follows that

$$\sigma_{\min}^2(H(\lambda)) = |1-\lambda|^2 + \frac{1}{2}(1+\varepsilon^2) - [\varepsilon^2|1-\lambda|^2 + \frac{1}{4}(1-\varepsilon^2)^2]^{\frac{1}{2}}.$$

Therefore, $\sigma_{\min}(H(\lambda))$ attains its minimum at $\lambda = 1$ which implies $\gamma_c = \omega_c = \varepsilon$. Note that in this case we also have $\gamma_c = \gamma_s$.

2.2 Infinite Dimensional Systems and Approximations

In this section we consider the control system

$$\dot{z}(t) = \mathcal{A}z(t) + \mathcal{B}u(t), \quad z(0) = z_0 \tag{2.20}$$

with output

$$y(t) = \mathcal{C}z(t). \tag{2.21}$$

We assume $\mathcal{A}$ generates a C_0-semigroup $\mathcal{S}(t)$ on the Hilbert space Z, $\mathcal{B} : U \to Z$, $\mathcal{C} : Z \to Y$ are bounded linear operators and U, Y are Hilbert spaces. Solutions of (2.20) will be mild solutions defined by

$$z(t) = \mathcal{S}(t)z_0 + \int_0^t \mathcal{S}(t-s)\mathcal{B}u(s)ds. \tag{2.22}$$

For $t > 0$ the reachable set at time t is given by

$$\mathcal{R}(t) = \left\{ \int_0^t \mathcal{S}(t-s)\mathcal{B}u(s)ds \middle| u \in L_2(0,t;U) \right\}.$$

System (2.20) is said to be exactly controllable in time t, if $\mathcal{R}(t) = Z$ and exactly controllable if $\cup_{t>0}\mathcal{R}(t) = Z$. Also (2.20) is called approximately controllable in time t, if $\overline{\mathcal{R}(t)} = Z$ and approximately controllable if $\overline{\cup_{t>0}\mathcal{R}(t)} = Z$. For analytic semigroups it is known (see [16]) that $\overline{\cup_{t>0}\mathcal{R}(t)} = Z$ if and only if there is a finite time $\hat{t}$ such that $\overline{\mathcal{R}(\hat{t})} = Z$. This is also true for semigroups generated by finite delay differential equations (see [2; 41]).

System (2.20) is said to be (exponentially) stabilizable if there is a bounded linear operator $\mathcal{F} : \mathcal{Z} \to \mathcal{U}$ such that the closed-loop operator $\mathcal{A}_c = \mathcal{A} + \mathcal{B}\mathcal{F}$ generates a C_0-semigroup $S(t)$ satisfying

$$\|\mathcal{S}(t)\| \le Me^{-\beta t}$$

for some $M > 0$ and $\beta > 0$. There are analogous definitions of observability and detectability and various other types of controllability (i.e. null controllability). A good summary of these definitions and topics may be found in [11]. However, we shall concentrate primarily on controllability questions and make some comments about stabilizability. It will be clear that dual results will exist for observability and detectability.

The first problem one faces when trying to define system measures for infinite dimensional systems is that in general most of the system properties are not generic. Consider the following simple example.

Example 3. Let $U = Z = \ell_2$ and define $\mathcal{A} = I$ and $\mathcal{B} : \ell_2 \to \ell_2$ by $\mathcal{B}(u_1, u_2, u_3, \dots) = (u_1, u_2/2, u_3/3, \dots, u_i/i, \dots)$. The operators $\mathcal{A}$ and $\mathcal{B}$ are bounded and the system $(\mathcal{A}, \mathcal{B})$ is approximately controllable. Define the perturbed systems $\mathcal{A}^N = \mathcal{A}$ and $\mathcal{B}^N(u_1, u_2, \dots) = (u_1, u_2/2, \dots, u_{N-1}/(N-1), 0, u_{N+1}/(N+1), \dots)$. Observe that $\|\mathcal{A}^N - \mathcal{A}\| = 0$ and $\|\mathcal{B}^N - \mathcal{B}\| \le 1/N$ and yet the system $(\mathcal{A}^N, \mathcal{B}^N)$ is not controllable for all $N \ge 1$. If we choose for U the Banach space $\ell_2^{-1} = \{(\xi_i) \big| \sum_{i=1}^\infty i^{-2}|\xi_i|^2 < \infty\}$, then $(\mathcal{A}, \mathcal{B})$ as defined above is exactly controllable and stabilizable and yet $(\mathcal{A}^N, \mathcal{B}^N)$ is neither controllable nor stabilizable for all $N \ge 1$.

Example 4. Let $Z = \ell_2$ and define $\mathcal{A}$ and $\mathcal{A}^N$ by $\mathcal{A} = -I$ and $\mathcal{A}^N x = \mathcal{A}^N(x_1, x_2, \dots) = (-x_1, -x_2, \dots, -x_{N-1}, 2x_N, -x_{N+1}, \dots)$. Observe that $\|e^{\mathcal{A}t}\| = e^{-t}$ and $\|\mathcal{A}^N x - \mathcal{A}x\| = 3|x_N| \to 0$. Moreover, $\|\mathcal{A}^N\| = 2$ so that $\mathcal{A}^N$ is a numerically stable and consistent approximation of $\mathcal{A}$. If e^N denotes the unit vector $e^N = (0, 0, \dots, 1, 0, 0, \dots) \in \ell_2$, then $\mathcal{A}^N e^N = 2e^N$. $\mathcal{A}^N$ has an unstable eigenvalue $\lambda = 2$ for all $N \ge 1$.

These examples indicate that there are no such things as stability, controllability or stabilizability measures for general infinite dimensional systems. For a more detailed discussion see also [37]. In order to define a reasonable measure it is essential to limit the set of allowable perturbations to a specific (structured) set. Initial results on structured perturbations that preserve stability have been established in [40]. In [8] it was shown that, for certain delay systems, approximate controllability is preserved under small perturbations of the system coefficients (including the delay).

From a design point of view it is worthwhile to think of finite dimensional approximating systems as "structured" perturbations of infinite dimensional systems. In particular, one can use finite element and finite difference schemes to construct very special (perturbed) approximating control systems for distributed parameter models governed by partial and functional differential equations. Therefore, one question of interest is that of determining those numerical schemes that preserve the various system properties and then finding among such schemes the ones that maximize the measures of the finite dimensional models.

This problem was considered for approximations of control systems with delays in [8]. For single-input systems it was possible to give sufficient conditions for an approximation scheme to preserve controllability. Moreover, it was shown in [8] that several of the standard numerical schemes for delay equations satisfy these conditions and hence preserve controllability under approximation. The situation becomes much more complex when the system is governed by parabolic and hyperbolic partial differential equations.

Two specific numerical schemes for approximating differential operators are the finite difference and finite element methods. Both approaches have certain advantages and limitations. In many cases (not always) these schemes lead to system matrices with a very special structure. For example, it is typical that such schemes produce matrices that are symmetric and sparse (banded, block tridiagonal, etc.). When such methods are used to approximate control systems governed by partial differential equations it is possible to use this structure in the design and analysis of the control problem. Moreover, since there are often several methods for approximating a particular control problem it is important to identify those schemes that produce models that are robust. More precisely, we are interested in determining the schemes that have good system measures γ_c, γ_s, etc. We shall consider this problem for a one dimensional heat equation. We focus on the standard finite (central) difference and (piecewise linear) finite element schemes.

Consider the system governed by the heat equation

$$y_t(t,x) = y_{xx}(t,x) + b(x)u(t), \quad 0 \le x \le 1, \quad t > 0, \tag{2.23}$$

with Dirichlet boundary conditions $y(t,0) = y(t,1) = 0$. Here we assume that $b(\cdot) \in L_2(0,1)$. Let $\mathcal{A}$ be the operator defined on $L_2(0,1)$ by

$$\mathcal{D}(\mathcal{A}) = \left\{\phi \in L_2(0,1) \middle| \phi \in H^2(0,1),\, \phi(0) = \phi(1) = 0\right\}, \tag{2.24}$$

and for $\phi \in \mathcal{D}(\mathcal{A})$

$$\mathcal{A}\phi = \frac{d^2}{dx^2}\phi. \tag{2.25}$$

The operator $\mathcal{A}$ generates a C_0-semigroup $\mathcal{S}(t)$ on $L_2(0,1)$ given by

$$[\mathcal{S}(t)\phi](x) = \sum_{k=1}^{\infty} e^{-\lambda_k t}\langle\phi, \phi_k\rangle\phi_k(x), \tag{2.26}$$

where $\lambda_k = k^2\pi^2$, $\phi_k(x) = \sqrt{2}\,\sin(k\pi x)$ and $\langle\cdot,\cdot\rangle$ denotes the standard inner product on $L_2(0,1)$. We define $\mathcal{B} : \mathbb{R} \to L_2(0,1)$ by

$$[\mathcal{B}u](x) = b(x)u, \tag{2.27}$$

and consider the equation (2.23) as a control problem in $L_2(0,1)$ governed by

$$\dot{z}(t) = \mathcal{A}z(t) + \mathcal{B}u(t). \tag{2.28}$$

We denote by $\Sigma^H = (\mathcal{A},\mathcal{B})$ the system operators defined by (2.24) - (2.25) and (2.27). Recall that (2.28) is approximately controllable in $L_2(0,1)$ if and only if

$$\langle\phi_k, b\rangle \neq 0 \quad \text{for all} \quad k = 1,2,\ldots. \tag{2.29}$$

We note that many of the results below can be extended to problems in more than one space variable and to some general parabolic systems. However, this introduces so many technical details that many of the basic ideas get lost. Also, it will become clear that even this "simple" one dimensional heat equation leads to difficult problems. We refer the reader to [25] for a discussion of controllability for more general problems.

The system (2.28) will be approximated by using finite difference and finite element schemes for (2.23). Divide the interval $(0,1)$ into $N+1$ equal subintervals $[x_i, x_{i+1}]$ where $x_i = \frac{i}{N+1}$, $i = 0,1,\ldots,N+1$. Assuming that $b(\cdot) \in H^1(0,1)$, then applying the central difference approximation of $\frac{d^2}{dx^2}$ leads to the N dimensional system

$$\dot{z}^N(t) = A_D^N z^N(t) + B_D^N u(t), \tag{2.30}$$

where

$$A_D^N = (N+1)^2 \begin{bmatrix} -2 & 1 & 0 & & & \\ 1 & -2 & 1 & & & 0 \\ & & \ddots & \ddots & \ddots & \\ 0 & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix} = (N+1)^2 \tilde{A}_D^N, \tag{2.31}$$

$$B_D^N = \mathrm{col}\,\big(b(x_1), b(x_2), \ldots, b(x_N)\big), \tag{2.32}$$

and $z^N(t)$ is identified with the vector

$$z^N(t) = \mathrm{col}\,\big(y(t,x_1), y(t,x_2), \ldots, y(t,x_N)\big). \tag{2.33}$$

The system $\Sigma_D^N = (A_D^N, B_D^N)$ is called the finite difference model.

We turn now to the finite element scheme. For each $i = 1,2,\ldots,N$ let $h_i^N(x)$ denote the hat function

$$h_i^N(x) = \begin{cases} (N+1)(x - x_{i-1}) & x_{i-1} \le x \le x_i, \\ -(N+1)(x - x_{i+1}) & x_i \le x \le x_{i+1}, \\ 0, \quad \text{elsewhere.} & \end{cases} \tag{2.34}$$

If $y(t,x)$ is approximated by

$$y^N(t,x) = \sum_{i=1}^{N} z_i^N(t) h_i^N(x), \tag{2.35}$$

then a standard Galerkin procedure leads to the finite element approximation

$$E_E^N \dot{z}^N(t) = F_E^N z^N(t) + G_E^N u(t). \tag{2.36}$$

Moreover

$$E_E^N = [\langle h_i^N, h_j^N \rangle] = \frac{1}{6(N+1)} \begin{bmatrix} 4 & 1 & 0 & & & \\ 1 & 4 & 1 & & 0 & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & 1 & 4 & 1 \\ & & & & 1 & 4 \end{bmatrix}, \tag{2.37}$$

$$F_E^N = -[\langle \dot{h}_i^N, \dot{h}_j^N \rangle] = (N+1)\tilde{A}_D^N, \tag{2.38}$$

and

$$G_E^N = \mathrm{col}\left(\langle b, h_1^N \rangle, \langle b, h_2^N \rangle, \ldots, \langle b, h_N^N \rangle\right). \tag{2.39}$$

Let

$$A_E^N = [E_E^N]^{-1} F_E^N, \quad B_E^N = [E_E^N]^{-1} G_E^N,$$

and define the finite element model $\Sigma_E^N = (A_E^N, B_E^N)$ by

$$\dot{z}^N(t) = A_E^N z^N(t) + B_E^N u(t). \tag{2.40}$$

It is obvious for this simple case that both schemes preserve stabilizability under approximation uniformly in N (i.e. have property (POES) as defined in [4]). In fact, it is shown in [3] that the same is true for finite element schemes applied to more general parabolic problems in several space dimensions. Their approach can also be extended to certain (but not all) finite difference schemes for such systems. It is *not* obvious that these schemes preserve controllability (even for the particular model considered here) and in fact this problem is not yet resolved. Therefore, it is worthwhile to have some conditions on the system that can be used to determine the controllability properties of the models Σ_D^N and Σ_E^N.

First consider the finite difference model Σ_D^N. The tridiagonal matrix A_D^N has eigenvalues (see [42])

$$\lambda_{D,k}^N = -4(N+1)^2 \sin^2 \frac{k\pi}{2(N+1)}, \quad k = 1, 2, \ldots, N, \tag{2.41}$$

and associated eigenvectors

$$z_{D,k}^N = \mathrm{col}\left(\sin(k\alpha_N), \sin(2k\alpha_N), \ldots, \sin(Nk\alpha_N)\right), \tag{2.42}$$

where $\alpha_N = \pi/(N+1)$. Consequently, it follows (see the identity 1.351 in [20]) that

$$\|z_{D,k}^N\|^2 = \sum_{j=1}^{N} \sin^2(jk\alpha_N) = \frac{N}{2} - (-1)^k \frac{\sin(Nk\pi/(N+1))}{2\sin(k\pi/(N+1))} = \frac{1}{2}(N+1). \tag{2.43}$$

In view of (2.41) - (2.43), it is clear that

$$\Phi^N = \sqrt{\frac{2}{(N+1)}} \begin{bmatrix} \sin\alpha_N & \sin 2\alpha_N & \dots & \sin N\alpha_N \\ \sin 2\alpha_N & \sin 4\alpha_N & \dots & \sin 2N\alpha_N \\ \vdots & & & \\ \sin N\alpha_N & \sin 2N\alpha_N & \dots & \sin N^2\alpha_N \end{bmatrix} \tag{2.44}$$

is the orthogonal transformation that diagonalizes A_D^N, i.e. $[\Phi^N]^T = [\Phi^N]^{-1}$ and

$$[\Phi^N]^T A_D^N \Phi^N = \Lambda_D^N = \mathrm{diag}(\lambda_{D,1}^N, \lambda_{D,2}^N, \dots, \lambda_{D,N}^N). \tag{2.45}$$

Observe that Φ^N is also symmetric so that $[\Phi^N]^T = \Phi^N = [\Phi^N]^{-1}$. Let ϕ_k^N denote the k-th column of Φ^N,

$$\phi_k^N = \sqrt{\frac{2}{(N+1)}}\, \mathrm{col}\,(\sin(k\alpha_N), \sin(2k\alpha_N), \dots, \sin(Nk\alpha_N)). \tag{2.46}$$

Lemma 1. *The finite difference model $\Sigma_D^N = (A_D^N, B_D^N)$ is controllable if and only if*

$$\langle \phi_k^N, B_D^N \rangle \neq 0 \quad \textit{for all} \quad k = 1, 2, \dots, N. \tag{2.47}$$

Proof. Let $\xi_k = \langle \phi_k^N, B_D^N \rangle$, $k = 1, 2, \dots, N$ and note that

$$\Phi^N B_D^N = \mathrm{col}(\xi_1^N, \xi_2^N, \dots, \xi_N^N).$$

The system Σ_D^N is controllable if and only if the controllability matrix

$$\mathcal{K}_\mathcal{D}^\mathcal{N} = [\mathcal{B}_\mathcal{D}^\mathcal{N}, \mathcal{A}_\mathcal{D}^\mathcal{N} \mathcal{B}_\mathcal{D}^\mathcal{N}, \dots, [\mathcal{A}_\mathcal{D}^\mathcal{N}]^{N-1} \mathcal{B}_\mathcal{D}^\mathcal{N}]$$

has maximal rank N. However,

$$\mathrm{rank}\, \mathcal{K}_\mathcal{D}^\mathcal{N} = \mathrm{rank} \begin{bmatrix} \xi_1^N & & & \\ & \xi_2^N & 0 & \\ & & \ddots & \\ & 0 & & \ddots & \\ & & & & \xi_N^N \end{bmatrix} \begin{bmatrix} 1 & \lambda_{D,1}^N & \cdots & (\lambda_{D,1}^N)^{N-1} \\ 1 & \lambda_{D,2}^N & \cdots & (\lambda_{D,2}^N)^{N-1} \\ & \vdots & & \vdots \\ & \vdots & & \vdots \\ 1 & \lambda_{D,N}^N & \cdots & (\lambda_{D,N}^N)^{N-1} \end{bmatrix}$$

Since the eigenvalues of A_N^D are all distinct, the Vandermonde matrix is non-singular. Hence, $\mathcal{K}_\mathcal{D}^\mathcal{N}$ has rank N if and only if (2.47) holds. $\square$

If $b(\cdot) \in H^1(0,1)$, let $g(x) = b(x + \frac{1}{2})$ for $-\frac{1}{2} \leq x \leq \frac{1}{2}$. In the cases where g is odd or even, (2.29) does not hold. If $g(x) = g(-x)$, then

$$B_D^N = \text{col}\,\big(b(x_1), \ldots, b(x_\ell), b(x_\ell), \ldots, b(x_1)\big), \quad \text{if } N = 2\ell$$

and

$$B_D^N = \text{col}\,\big(b(x_1), \ldots, b(x_\ell), b(x_{\ell+1}), b(x_\ell), \ldots, b(x_1)\big), \quad \text{if } N = 2\ell + 1.$$

If $g(x) = -g(-x)$, then

$$B_D^N = \text{col}\,\big(b(x_1), \ldots, b(x_\ell), -b(x_\ell), \ldots, -b(x_1)\big), \quad \text{if } N = 2\ell$$

and

$$B_D^N = \text{col}\,\big(b(x_1), \ldots, b(x_\ell), 0, -b(x_\ell), \ldots, -b(x_1)\big), \quad \text{if } N = 2\ell + 1.$$

Proposition 2. *If $g(x)$ is odd or even, then Σ^H and Σ_D^N are not controllable for all $N \geq 1$.*

Proof. As noted above, Σ^H is not controllable since (2.29) fails. Let $\xi_k^N = \langle \phi_k^N, B_D^N \rangle$. If $N = 2\ell$, then a direct calculation yields

$$\xi_k^N = \sqrt{\frac{2}{N+1}} \sum_{i=1}^{\ell} b(x_i)(1 \mp (-1)^k) \sin \frac{ik\pi}{2\ell + 1}.$$

If $N = 2\ell + 1$, then it follows just as above that

$$\xi_k^N = \sqrt{\frac{2}{N+1}} \sum_{i=1}^{\ell} b(x_i)(1 \mp (-1)^k) \sin \frac{ik\pi}{2(\ell + 1)} + b(x_{\ell+1}) \sin \frac{k\pi}{2}.$$

Above, the minus sign is valid if g is even, the plus sign if g is odd. Hence we conclude that the even (odd) numbered coordinates of ξ^N vanish if g is symmetric (skew symmetric). □

A close look at the above proof yields a clear relationship between (2.29) and (2.47) in the special cases above where (2.29) fails because of the special form of $g(\cdot)$. This form is also present in B_D^N and it is precisely this form that causes (2.47) to fail. As we shall see below, the same structure is preserved by the finite element scheme. In particular, let

$$\begin{aligned}
\tilde{E}_E^N &= 6(N+1)E_E^N, \qquad \tilde{F}_E^N = (N+1)^{-1}F_E^N, \\
\tilde{A}_E^N &= [\tilde{E}_E^N]^{-1}\tilde{F}_E^N = \frac{1}{6(N+1)^2}A_E^N, \\
\tilde{B}_E^N &= [\tilde{E}_E^N]^{-1}G_E^N = \frac{1}{6(N+1)}B_E^N \quad \text{and} \\
\tilde{A}_D^N &= (N+1)^{-2}A_D^N.
\end{aligned}$$

Clearly, $\tilde{\Sigma}_E^N = (\tilde{A}_E^N, \tilde{B}_E^N)$ is controllable if and only if $\Sigma_E^N = (A_E^N, B_E^N)$ is controllable.

Observe that the "stiffness" matrix $(N+1)^2 \tilde{F}_E^N$ is the system matrix for the finite difference equation, i.e.

$$\tilde{F}_E^N = \tilde{A}_D^N,$$

and that the "mass" matrix $\tilde{E}_E^N$ can be written as

$$\tilde{E}_E^N = \tilde{F}_E^N + 6I^N = \tilde{A}_D^N + 6I^N, \tag{2.48}$$

where I^N is the $N \times N$ identity matrix.

Lemma 2. *Let $\tilde{A}_E^N$ and $\tilde{A}_D^N$ be as given above. Then $\lambda \in \mathbb{C}$ is an eigenvalue of $\tilde{A}_E^N$ with eigenvector z_λ^N if and only if $6\lambda/(1-\lambda)$ is an eigenvalue of $\tilde{A}_D^N$ with eigenvector z_λ^N.*

The proof of Lemma 2 follows immediately from (2.48). Moreover, as a consequence of Lemma 2 and (2.48) it follows that Φ^N defined by (2.44) diagonalizes $\tilde{A}_D^N$ and $\tilde{A}_E^N$. We use these observations to establish the following result.

Proposition 3. *The finite element model $\Sigma_E^N = (A_E^N, B_E^N)$ is controllable if and only if*

$$\langle \phi_k^N, G_E^N \rangle \neq 0 \quad \textit{for all} \quad k = 1, 2, \ldots, N. \tag{2.49}$$

Proof. The system Σ_E^N is controllable if and only if $\tilde{\Sigma}_E^N$ is controllable, i.e. if and only if

$$\tilde{\mathcal{K}}_E^N = \left[\tilde{B}_E^N, \tilde{A}_E^N \tilde{B}_E^N, \ldots, [\tilde{A}_E^N]^{N-1} \tilde{B}_E^N\right]$$

has rank N. However, $\tilde{E}_E^N \tilde{F}_E^N = \tilde{F}_E^N \tilde{E}_E^N$ so that

$$\begin{aligned}
\operatorname{rank} \tilde{\mathcal{K}}_E^N &= \operatorname{rank}\big[[\tilde{E}_E^N]^{-1} G_E^N, [\tilde{E}_E^N]^{-1}\big([\tilde{E}_E^N]^{-1} \tilde{F}_E^N\big) G_E^N, \\
&\qquad \ldots, [\tilde{E}_E^N]^{-1}\big([\tilde{E}_E^N]^{-1} \tilde{F}_E^N\big)^{N-1} G_E^N\big] \\
&= \operatorname{rank}[\Phi^N G_E^N, \tilde{\Lambda}^N \Phi^N G_E^N, \ldots, (\tilde{\Lambda}^N)^{N-1} \Phi^N G_E^N],
\end{aligned}$$

where $\tilde{\Lambda}^N$ is a non-singular diagonal matrix. Hence, $N = \operatorname{rank} \tilde{\mathcal{K}}_E^N$ if and only if (2.49) holds and this completes the proof. □

Proposition 4. *If $g(x)$ is odd or even, then Σ^H and Σ_E^N are not controllable for all $N \geq 1$.*

As seen above there is a nice relationship between the system matrices for Σ_D^N and Σ_E^N. Moreover, one has the following

Proposition 5. *Assume that there exists a non-singular transformation T^N such that*

(1) $\tilde{A}_D^N = T^N \tilde{A}_D^N (T^N)^{-1}$
(2) $G_E^N = T^N B_D^N$.

Then, Σ_D^N is controllable if and only if Σ_E^N is controllable.

Proof. Since scalar factors do not affect the controllability, it suffices to consider the

$$\begin{aligned}&\operatorname{rank}\big[[\tilde{E}_E^N]^{-1}\tilde{F}_E^N - \lambda I^N, [\tilde{E}_E^N]^{-1}G_E^N\big] = \operatorname{rank}[\tilde{F}_E^N - \lambda\tilde{E}_E^N, G_E^N]\\ &= \operatorname{rank}\left[\tilde{A}_D^N - \frac{6\lambda}{1-\lambda}I^N, G_E^N\right] = \operatorname{rank}\left[T^N\tilde{A}_D^N(T^N)^{-1} - \frac{6\lambda}{1-\lambda}I^N, T^N B_D^N\right].\end{aligned}$$

The conclusion follows from the Hautus test and Lemma 2. □

Example 5. Let $b(x) = x$, $0 \le x \le 1$. It follows from (2.29) that Σ^H is controllable and $\langle\phi_k^N, B_D^N\rangle$ can be calculated. In particular, $B_D^N = \frac{1}{N+1}\operatorname{col}(1, 2, \ldots, N)$ and

$$\xi_k^N = \langle\phi_k^N, B_D^N\rangle = \sqrt{\frac{2}{N+1}}\sum_{i=1}^{N}\frac{i}{N+1}\sin ki\alpha_N.$$

Applying the identity 1.352 (i) in [20] to the expression

$$\sqrt{\frac{2}{N+1}}\frac{1}{N+1}\left[\frac{\sin[(N+1)k\alpha_N]}{4\sin^2\frac{k\alpha_N}{2}} - \frac{N+1}{2}\frac{\cos[\frac{2N+1}{2}k\alpha_N]}{\sin\frac{k\alpha_N}{2}}\right],$$

and performing a few elementary manipulations it follows that

$$\xi_k^N = (-1)^{k+1}\sqrt{\frac{1}{2(N+1)}}\cot\frac{k\pi}{2(N+1)}, \qquad k = 1, 2, \ldots, N. \tag{2.50}$$

Since

$$0 < |\xi_N^N| < |\xi_{N-1}^N| < \cdots < |\xi_1^N|, \tag{2.51}$$

it follows that the finite difference model Σ_D^N is controllable for all $N \ge 1$. If the finite element scheme is applied, then

$$G_E^N = \frac{1}{(N+1)^2}\operatorname{col}(1, 2, \ldots, N), \tag{2.52}$$

and $G_E^N = T^N B_D^N$ where $T^N = (N+1)I^N$. Therefore, it follows from Proposition 2.8 that the finite element scheme Σ_E^N is also controllable for all $N \ge 1$.

Remark. It is important to note that checking for controllability of finite dimensional systems is in general a subtle numerical task [36; 30]. In contrast, verification of (2.49) is numerically accurate and straightforward. In the next section we shall concentrate on the robustness of the system measures for this example.

2.3 System Measures and Case Studies

In this section we consider the system measures γ_c and γ_s for each of the two schemes applied to Example 5. Also, we use the simple model problem defined in Example 5 to present a comparison of the condition number for the Riccati equation that comes from an LQR problem and the stabilizability radii.

Approximations for the Heat equation

In particular, we consider the problem governed by

$$y_t(t,x) = y_{xx}(t,x) + xu(t), \quad 0 \le x \le 1, \quad t > 0 \tag{2.53}$$

with Dirichlet boundary conditions $y(t,0) = y(t,1) = 0$. The finite difference system $\Sigma_D^N = (A_D^N, B_D^N)$ and the finite element system $\Sigma_E^N = (A_E^N, B_E^N)$ are defined as above, where for this problem

$$B_D^N = \frac{1}{N+1} \text{ col}\,(1,2,\ldots,N) \tag{2.54}$$

and

$$B_E^N = [E_E^N]^{-1} G_E^N = [E_E^N]^{-1} \frac{1}{(N+1)^2} \text{ col}\,(1,2,\ldots,N). \tag{2.55}$$

Let γ_{DC}^N (γ_{DS}^N) and γ_{EC}^N (γ_{ES}^N) denote the controllability (stabilizability) measures of Σ_D^N and Σ_E^N, respectively. From the previous section we know that Σ_D^N and Σ_E^N are controllable. Moreover, since $A_D^N = [A_D^N]^T$ and $A_E^N = [A_E^N]^T$ we have from Theorem 5 that

$$\gamma_{DC}^N = \omega_{DC}^N = \min_{\lambda \in \mathbf{R}} \sigma_{\min}[\lambda I^N - A_D^N, B_D^N], \tag{2.56}$$

and

$$\gamma_{EC}^N = \omega_{EC}^N = \min_{\lambda \in \mathbf{R}} \sigma_{\min}[\lambda I^N - A_E^N, B_E^N]. \tag{2.57}$$

Corollary 4 implies that

$$\gamma_{DS}^N = \omega_{DS}^N = \sigma_{\min}[A_D^N, B_D^N], \tag{2.58}$$

and

$$\gamma_{ES}^N = \omega_{ES}^N = \sigma_{\min}[A_E^N, B_E^N]. \tag{2.59}$$

These formulas can be used to compute the various measures. Moreover, it is possible to obtain explicit upper and lower bounds on the controllability measures. The following theorems provide such bounds. The proofs are given in the appendix.

Theorem 7. *Let $\gamma_{DC}^N = \omega_{DC}^N$ denote the controllability measure of the finite difference approximation of* (2.53). *If $N \ge 8$, then*

$$\delta_D^N - \epsilon_D^N < (\gamma_{DC}^N)^2 \le \delta_D^N, \tag{2.60}$$

where

$$\delta_D^N = |\xi_N^N|^2 = \frac{1}{2(N+1)} \cot^2 \frac{N\pi}{2(N+1)}, \tag{2.61}$$

and

$$\epsilon_D^N = \frac{1}{\beta_D} \cdot \frac{\pi^2 (3 \ln \frac{2(N+1)}{3} + \pi)}{24N^2(N^2-1)}, \tag{2.62}$$

and

$$\beta_D = 726.$$

Theorem 8. *Let $\Sigma_E^N = (A_E^N, B_E^N)$ denote the finite element approximation of* (2.53). *If $\gamma_{EC}^N = \omega_{EC}^N$ is the controllability measure for Σ_E^N, then for $N \geq 8$*

$$\delta_E^N - \epsilon_E^N < (\gamma_{EC}^N)^2 \leq \delta_E^N, \tag{2.63}$$

where

$$\delta_E^N = \frac{9}{[3 - 2\sin^2 \frac{N\pi}{2(N+1)}]^2} \cdot \delta_D^N, \tag{2.64}$$

and

$$\epsilon_E^N = \frac{1}{\beta_E} \cdot \frac{\pi^2(3 \ln \frac{2(N+1)}{3} + \pi)}{24N^2(N^2-1)}, \tag{2.65}$$

and

$$\beta_E = 550.$$

Corollary 5. *If γ_{DC}^N and γ_{EC}^N are as above, then*

$$\lim_{N\to+\infty} \gamma_{DC}^N = \lim_{N\to\infty} \gamma_{EC}^N = 0.$$

The proof of Corollary 5 is an easy consequence of Theorems 7 and 8. Note that the asymptotic rates of δ_D^N and δ_E^N are the same. However, for "small" values of N, there are considerable differences. Moreover, the rate at which δ_D^N (and δ_E^N) converges to zero decreases significantly for $N \geq 10$. If one had no theoretical estimates (such as (2.61) and (2.64)) then numerically calculated values for $N \leq 10$ might lead one to believe that the decay of controllability occurs so rapidly that the high order models would be numerically uncontrollable on most digital computers.

In Tables 2.1, 2.2 we compare computed values of $(\gamma_{DC}^N)^2$ and $(\gamma_{EC}^N)^2$ with the theoretical upper and lower bounds given in Theorem 7 and Theorem 8. In our computations we used Algorithm II in [44]. The eigenvalues of A_D^N (A_E^N) served as initial guesses for the location of the minimum in (2.7). The proofs of Theorems 7 and 8 show that in the particular example one can choose as initial guess the eigenvalue with largest modulus which significantly reduces the computational effort. We also report the observation supported by numerous examples [35] that the minimum in (2.7) is usually attained close to an eigenvalue of A_D^N (A_E^N).

The most important point to make is that the finite element model is roughly one order of magnitude more robust than the corresponding finite difference model. Both systems provide order $\mathcal{O}(h^2)$, $h = 1/(N+1)$ approximation of the generator $\mathcal{A}$ and for fixed N, Σ_D^N and Σ_E^N are of the same dimension.

Condition Number and Stabilizability Radius

Here we discuss the relationship between the condition number for the algebraic Riccati equation and the stabilizability radius γ_s. We compute the actual condition number K_{RIC} as defined on page 533 (Theorem 13.4.2) in Datta's book [12] and compare this to the inverse of γ_s. To illustrate the point we return to the first example given as the motivating problem. In particular, we consider the

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} \delta \\ \epsilon \end{bmatrix} u(t),$$

with $\epsilon = 1$. In this case the system becomes unstabilizable as δ approaches zero. In Table 2.5 below we see that the Riccati equation condition number K_{RIC} increases as δ approaches zero. In addition, ratio $K_{RIC}{*}\gamma_s$ approaches a constant value 1.0797 which shows that the inverse of the stabilizability radius γ_s provides a good estimate of the conditioning number K_{RIC}. Therefore, it might also be possible to use system radii to help construct "well-conditioned" approximating control systems. This topic needs further study.

2.4 Concluding Remarks

In this paper we discussed a measure of controllability (stabilizability) quantifying the distance of a controllable (stabilizable) system to the set of uncontrollable (unstabilizable) systems. This measure was applied to finite dimensional systems which arise by approximating an infinite dimensional system. This process leads to several important questions that are often not fully addressed. In particular one should consider the following issues:

1. If the original distributed parameter system is controllable (stabilizable) are the finite dimensional approximation also controllable (stabilizable)? For the heat equation this leads to a very specific problem:
   (P1) Find sufficient conditions for $b \in H^1(0,1)$ which ensure that $\langle \varphi_k, b\rangle_{L^2} \neq 0$, $k = 1, 2, \ldots$, implies that $\langle \phi_k^N, B_D^N\rangle \neq 0$ for all $k = 1, \ldots, N$ (where ϕ_k^N, B_D^N are defined in (2.46) and (2.32)).
2. If the finite dimensional approximate model is controllable (stabilizable) how robust is the model with respect to the measure γ_c (γ_s)?
3. How does γ_c (γ_s) vary as the approximation scheme is refined?

It is not unexpected that the controllability margin in the example discussed in this paper deteriorates as the approximation is refined. On the one hand the finite dimensional systems converge to the distributed parameter system for which (approximate) controllability is not robust with respect to bounded perturbations [37; 10]. On the other hand, the dimension of the finite dimensional approximating systems grows as the discretization is refined making it more likely that small perturbations will destroy controllability [1]. In addition, the measures considered here do not take into account the typical structure of the finite dimensional approximating systems. In general therefore, it will not be possible to interpret the closest uncontrollable system as an approximation of a perturbed distributed parameter system.

The problem of preserving stabilizability under approximation has been considered by serval authors [8; 24; 27]. This problem can be rather complex or quite simple depending on the particular distributed parameter system and the choice of approximating scheme. "Standard finite element" schemes applied to hyperbolic systems often preserve stabilizability while spline based schemes for delay equations are much more difficult to analyze [24], [27]. The problem of preserving controllability is much more complex and only a few results exist for this problem (see [8], [26]).

In addition to the system dynamics (Σ) we consider an output equation

$$y(t) = Cx(t)$$

and discuss the observability of the pair $[A, C]$. By duality one can define observability measures γ_o, ω_o analogous to γ_c, ω_c. It is known that none of these measures is invariant under an arbitrary change of basis. The effect of similarity transformations in the state space has been discussed in [38] and [39]. In particular, it was shown that for a generic controllable system, γ_c (γ_s) may attain any value in $(0, \infty)$ by choosing an appropriate basis. Furthermore, in [18] the authors suggest a sequence of similarity transformations based on the successive solution of a Riccati equation which tends to increase significantly the margin of stabilizability (controllability).

Let the output operator for the heat equation be $\mathcal{C} = \mathcal{I}$, the identity operator on $L_2(0,1)$. If one uses the finite element scheme then the approximate output matrix is given by the mass matrix

$$C_E^N = E_E^N.$$

It is straightforward to use the ideas above to calculate detectability and observability measures for this problem. Also, if one balances the system (see [35]) one has another realization of the form

$$\tilde{A}_E^N = TA_E^N T^{-1}, \quad \tilde{B}_E^N = TB_E^N, \quad \tilde{C}_E^N = TC_E^N.$$

Observe that this system is a single input / multiple output system. In Table 2.3 we see that the stabilizability and detectability measures for the finite element model and the balanced model are bounded away from zero. Also, for the balanced system $\tilde{\gamma}_{Es}^N = \tilde{\gamma}_{Eo}^N$. However, as illustrated by Table 2.4, balancing increases the controllability measure and decreases the observability measure. This was observed in [35] for finite element approximations of various parabolic, hyperbolic and mixed distributed parameter systems. In closing we list some open problems that if solved in a satisfactory way could be useful in constructing approximate models for control design and in estimating condition numbers of numerical algorithms used in control. In particular, since in general $\gamma_o \neq \gamma_c$ and (even balancing) state transformations that increase γ_c can decrease γ_o, we pose the following problems:

(P2) Given that $[A, B]$ is controllable and $[A, C]$ is observable, find a non-singular $T^* \in \mathbf{R}^{n \times n}$ that maximizes

$$J_{co}(T) = \min\{\tilde{\gamma}_c(T), \tilde{\gamma}_o(T)\}$$

over the set of non-singular matrices $T \in \mathbb{R}^{n\times n}$, where

$$\tilde{\gamma}_c(T) = \tilde{\gamma}_c(TAT^{-1}, TB) \quad \text{and} \quad \tilde{\gamma}_o(T) = \gamma_o(TAT^{-1}, CT^{-1}).$$

(P3) Find T^* maximizing $J_{co}(T)$ subject to the additional constraint that T is "well conditioned", i.e. that $J_{co}(T)$ is maximized over a set of the form

$$K(\delta) = \{T \in \mathbb{R}^{n\times m} : \sigma_{\min}(T) \geq \delta\} \quad \text{for some } \delta > 0.$$

These problems could also be stated for the stabilizability and detectability measures.

Appendix

In this section we present the proofs for Theorems 7 and 8. It is easy to see that the minimization problem defined by (2.9) is equivalent to the problem

$$\gamma_c^2 = \min_{\substack{\|z\|=1 \\ z\in C^n}} (z^* AA^T z - |z^* Az|^2 + \|z^* B\|^2).$$

Since

$$z^* AA^T z - |z^* Az|^2 \geq 0 \quad \text{for} \quad \|z\| = 1,$$

with equality holding if and only if z^* is a left eigenvector of A, we deduce the following upper bound for γ_c^2

$$\gamma_c^2 \leq \|z^* B\|_2^2, \tag{2.66}$$

where z^* is any normalized left eigenvector of A. (2.66) together with (2.50), (2.51) gives the upper bound for $[\gamma_{DC}^N]^2$ as presented in (2.61).

In order to get *lower* bounds for $[\gamma_{DC}^N]^2$ we intensively exploit the rich structure of Example 5. Since the measure of controllability is invariant under unitary transformations we may equivalently discuss the dependence of the eigenvalues of

$$(\lambda I^N - \Lambda_D^N)^2 + \Phi^N B_D^N (\Phi^N B_D^N)^T \tag{2.67}$$

on the real parameter λ, where Λ_D^N and Φ^N have been defined in (2.44) and (2.45). However, any eigenvalue of (2.67) is contained in one of the Gershgorin discs (see [43])

$$\mathcal{D}_i^N = \{\mu \in \mid |\mu - ((\lambda - \lambda_{D,i}^N)^2 + [\xi_i^N]^2)| \leq r_i^N\}$$
$$r_i^N = \sum_{\substack{j=1 \\ j\neq i}}^{N} |\xi_i^N \xi_j^N|, \quad i = 1, \ldots, N,$$
$$\Phi^N B_D^N = \text{col}\,(\xi_1^N, \ldots, \xi_N^N).$$

Note that λ affects only the centers of the Gershgorin discs but not their radii. Therefore, one of the Gershgorin discs is shifted as close as possible to the origin if we choose $\lambda = \lambda^N_{D,i_0}$, where i_0 is the index satisfying

$$[\xi^n_{i_0}]^2 = \min_{i=1,\dots,N} [\xi^N_i]^2.$$

In view of (2.51) $i_0 = N$ for Example 2.28.

In the following we first establish tight bounds on r^N_k and $[\xi^N_N]^2$ and show in a second step that $\mathcal{D}^N_N \cap \mathcal{D}^N_k = \varnothing$, $k = N-1,\dots,1$. As a consequence we infer that $\mathcal{D}^N_N$ contains exactly one eigenvalue of (2.67) [43, p. 71]. In order to enhance the accuracy of the estimates we scale the system by multiplying the last column of (2.67) by $\alpha_N > 0$ and accordingly the last row by $\frac{1}{\alpha_N}$. α_N will be appropriately chosen in the course of the proof. Consequently, the radii of the Gershgorin discs are given by

$$\begin{aligned}
r^N_N &= \frac{1}{\alpha_N}\frac{1}{2(N+1)}\cot\frac{N\pi}{2(N+1)}\sum_{i=1}^{N-1}\cot\frac{i\pi}{2(N+1)},\\
r^N_k &= \frac{1}{2(N+1)}\cot\frac{k\pi}{2(N+1)}\sum_{i=1}^{N-1}\cot\frac{i\pi}{2(N+1)}\\
&\quad -\frac{1}{2(N+1)}\cot^2\frac{k\pi}{2(N+1)}+\frac{\alpha_N}{2(N+1)}\cot\frac{k\pi}{2(N+1)}\cot\frac{N\pi}{2(N+1)},\\
&\quad k=1,\dots,N-1.
\end{aligned} \tag{2.68}$$

We will also make frequent use of the estimates

$$\cot\frac{\pi}{2(N+1)} \le \frac{2}{3}(N+1), \qquad N \ge 2, \tag{2.69}$$

$$\cot\frac{k\pi}{2(N+1)} \le \frac{\pi}{2}\frac{N-k+1}{k}, \quad k=1,\dots,N, \tag{2.70}$$

$$\cot\frac{k\pi}{2(N+1)} \ge \frac{2}{\pi}\frac{N+1-k}{k}, \quad k=1,\dots,N. \tag{2.71}$$

With regard to r^N_N we have the estimate

$$\begin{aligned}
\sum_{i=1}^{N-1}\cot\frac{i\pi}{2(N+1)} &\le \int_1^{N-1}\cot\frac{\pi x}{2(N+1)}dx + \cot\frac{\pi}{2(N+1)}\\
&= \frac{2}{\pi}(N+1)\left(\ln\sin\frac{N-1}{N+1}\frac{\pi}{2} - \ln\sin\frac{\pi}{2(N+1)}\right) + \cot\frac{\pi}{2(N+1)}.
\end{aligned}$$

Since $\sin x \ge \frac{3}{\pi}x$ for $x \in [0, \frac{\pi}{6}]$, it follows from (2.69) that

$$\sum_{i=1}^{N-1} \cot \frac{i\pi}{2(N+1)} \leq \frac{2}{\pi}(N+1)\ln\frac{2(N+1)}{3} + \frac{2}{3}(N+1).$$

This inequality combined with (2.70) gives

$$r_N^N \leq \frac{1}{\alpha_N}\left[\frac{1}{2N}\ln\frac{2(N+1)}{3} + \frac{\pi}{6N}\right], \quad N = 2, 3, \ldots \quad . \tag{2.72}$$

A similar argument shows that

$$r_{N-1}^N \leq \frac{1}{N-1}\ln\frac{2(N+1)}{3} + \frac{\pi}{3(N-1)} + \frac{\pi^2}{4N(N^2-1)}\alpha_N, \tag{2.73}$$

$N = 2, 3, \ldots$, and using (2.69) - (2.71) again yields

$$\begin{aligned} r_k^N \leq\ & \frac{N-k+1}{k}\left[\frac{1}{2}\ln\frac{2(N+1)}{3} + \frac{\pi}{6}\right] - \frac{1}{N+1}\frac{2}{\pi^2}\left(\frac{N+1-k}{k}\right)^2 \\ & + \alpha_N\frac{\pi^2}{8}\frac{(N-k+1)}{N(N+1)k} \quad k = 1, \ldots, N-2. \end{aligned} \tag{2.74}$$

Next we estimate the gap between $\lambda_{D,i}^N$ and $\lambda_{D,N}^N$, $i = 1, \ldots, N-1$. This is based on the identity

$$|\lambda_{D,N}^N - \lambda_{D,i}^N| = 4(N+1)^2 \cdot \sin\frac{(N+i)\pi}{2(N+1)}\sin\frac{(N-i)\pi}{2(N+1)}. \tag{2.75}$$

In particular, we have

$$|\lambda_{D,N}^N - \lambda_{D,N-1}^N| = 4(N+1)^2 \cdot \sin\frac{3\pi}{2(N+1)}\sin\frac{\pi}{2(N+1)},$$

and estimating the sine as above we find that

$$|\lambda_{D,N}^N - \lambda_{D,N-1}^N| \geq 27, \quad N \geq 8. \tag{2.76}$$

The estimate (2.76) is rather tight. This is clear when it is compared to the exact limit

$$\lim_{N\to\infty} |\lambda_{D,N}^N - \lambda_{D,N-1}^N| = 3\pi^2.$$

Employing the estimates

$$\begin{aligned} \sin x &\geq \frac{2\sqrt{2}}{\pi}x, & x &\in \left[0, \frac{\pi}{4}\right] \quad \text{and} \\ \sin x &\geq 2\sqrt{2}\left(1 - \frac{x}{\pi}\right), & x &\in \left[\frac{3\pi}{4}, \pi\right], \end{aligned}$$

we obtain the bound

$$|\lambda_{D,N}^N - \lambda_{D,N-2}^N| \geq 64, \quad N \geq 8, \tag{2.77}$$

and making a simple linear estimate of $\sin x$ we finally have

$$|\lambda_{D,N}^N - \lambda_{D,k}^N| \geq 4(N-k)(N-k+2), \quad k = 1, \ldots, N-2. \tag{2.78}$$

It is easily checked that

$$r_N^N + [\xi_N^N]^2 < 1, \quad N \geq 8,$$

which is an upper bound for $\mathcal{D}_N^N$. In the following step we choose α_N as large as possible so that $\mathcal{D}_N^N \cap \mathcal{D}_{N-1}^N = \varnothing$. This is possible since

$$[\xi_{N-1}^N]^2 + |\lambda_{D,N}^N - \lambda_{D,N-1}^N|^2 - r_{N-1}^N > |\lambda_{D,N}^N - \lambda_{D,N-1}^N|^2 - r_{N-1}^N \geq 1.$$

If we insert (2.73) and (2.76) into the above inequality, it is clear that we may choose α_N so that

$$\alpha_N \leq \frac{4}{\pi^2} N(N^2-1) \cdot 726. \tag{2.79}$$

A similar argument using (2.74) and (2.77) shows that $\mathcal{D}_N^N \cap \mathcal{D}_{N-2}^N = \varnothing$. Hence it remains to show that $\mathcal{D}_N^N \cap \mathcal{D}_k^N = \varnothing$, $k = 1, \ldots, N-3$, $N \geq 8$. Observe that in view of (2.78), (2.74) and (2.79)

$$|\lambda_{D,N}^N - \lambda_{D,k}^N|^2 + [\xi_k^N]^2 - r_k^N \geq 1$$

is a consequence of the inequality

$$16(N-k)^2(N-k+2)^2 - \frac{N-k+1}{k}\left[\frac{1}{3}e^{-1}(N+1) + \frac{\pi}{6}\right]$$
$$- \frac{1}{2} \cdot 726(N-1)\frac{N-k+1}{k} \geq 1.$$

Therefore it suffices to establish the estimate

$$\begin{aligned} &16(N-k)^2(N-k+2)k - \frac{1}{3}e^{-1}(N+1) - \frac{\pi}{6} \\ &> \frac{k}{N-k+1} + 363(N-1), \quad k = 1, \ldots, N-3, \quad N \geq 8. \end{aligned} \tag{2.80}$$

The estimate

$$\frac{k}{N-k+1} \leq \frac{N-3}{4}, \qquad k = 1, \ldots, N-3$$

provides an upper bound for the right hand side of (2.80). Combined with $\frac{4}{3}e^{-1} < .5$ and $\frac{2\pi}{3} < 2.5$ one can show that (2.80) follows from the stronger inequality

$$\begin{aligned} &64(N-k)^2(N-k+2)k - \left(\frac{3}{2} + 726 \cdot 2\right) N + 726 \cdot 2 > 0, \\ &\quad k = 1, \ldots, N-3, \quad N \geq 8. \end{aligned} \tag{2.81}$$

We now establish (2.81). Define for fixed $N \geq 8$

$$G(N,x) = (N-x)^2(N-x+2)x, \quad x \in [0,N].$$

Since $G(N,x)$ is nonnegative and possesses a unique maximum on $[0,N]$ it is obvious that

$$G(N,k) \geq \min\big(G(N,1), G(N,N-3)\big).$$

Since $G(N,1) \geq G(N,N-3)$, $N \geq 8$ it suffices to verify (2.81) for $k = N-3$. This completes the proof of Theorem 7.

The proof of Theorem 8 is very similar. Therefore we restrict ourselves to a brief sketch and leave the details to the reader.

Let $\tilde{\lambda}^N_{D,k}$, $k = 1,\ldots,N$ denote the eigenvalues of $\tilde{A}^N_D$ and

$$\tilde{\Lambda}^N_D = \text{diag}\,(\tilde{\lambda}^N_{D,1},\ldots,\tilde{\lambda}^N_{D,N}). \tag{2.82}$$

Defining

$$\tilde{\lambda}^N_{E,k} = \frac{\tilde{\lambda}^N_{D,k}}{6+\tilde{\lambda}^N_{D,k}}, \quad k = 1,\ldots,N,$$

Lemma 2.5 implies that $\tilde{\lambda}^N_{E,k}$ is an eigenvalue of $\tilde{A}^N_E$. We collect them in

$$\begin{aligned} \tilde{\Lambda}^N_E &= \text{diag}(\tilde{\lambda}^N_{E,1},\ldots,\tilde{\lambda}^N_{E,N}), \\ \Lambda^N_E &= 6(N+1)^2 \cdot \tilde{\Lambda}^N_E = \text{diag}(\lambda^N_{E,1},\ldots,\lambda^N_{E,N}). \end{aligned} \tag{2.83}$$

As a consequence of (2.48) we note

$$\Phi^N \tilde{E}^N_E \Phi^N = \tilde{\Lambda}^N_D + 6I^N. \tag{2.84}$$

Applying the orthogonal transformation Φ^n to Σ^N_E, it follows that

$$[w^N_{EC}] = \min_{\lambda\in\mathbf{R}} \sigma_{\min}\big([\lambda I - \Lambda^N_D, \phi^N[E^N_E]^{-1}G^N_E]\big).$$

Taking into account (2.84) and (2.52) we conclude

$$\hat{B}^N := \Phi^N[E^N_E]^{-1}G^N_E = 6(\tilde{\Lambda}^N_E + 6I^N)^{-1}\Phi^N B^N_D. \tag{2.85}$$

Consequently, analogous to (2.67) we consider

$$(\lambda I - \Lambda^N_E)^2 + \hat{B}^N[\hat{B}^N]^T.$$

Since the k-th coordinate of $\hat{B}^N$ satisfies

$$|(\hat{B}^N)_k| = \frac{6|\xi_k|}{6-4\sin^2\frac{k\pi}{2(N+1)}} = \frac{6}{6-4\sin^2\frac{k\pi}{2(N+1)}}\sqrt{\frac{1}{2(N+1)}}\cot\frac{k\pi}{2(N+1)},$$

and since $f(x) = (6 - 4\sin^2 x)^{-1} \cot x$ is strictly decreasing on $(0, \frac{\pi}{2})$ we obtain the estimate

$$|(\hat{B}^N)_N| < |(\hat{B}^N)_k|, \quad k = 1, \dots, N-1.$$

This establishes the upper bound for $[w_{EC}^N]^2$ given in Theorem 8 and also shows that the choice $\lambda = \lambda_{E,N}^N$ shifts the configuration of Gershgorin discs as far as possible to the left. Analogous to the finite difference approximation the centers of the Gershgorin discs $\mathcal{E}_k^N$ for the finite element scheme are given by

$$(\lambda_{E,N}^N - \lambda_{E,k}^N)^2 + (\hat{B}^N)_k^2,$$

and their radii by

$$R_N^N = \frac{1}{\tilde{\alpha}_N} \sum_{i=1}^{N-1} |(\hat{B}^N)_N||(\hat{B}^N)_k|$$

and

$$R_k^N = |(\hat{B}^N)_k| \sum_{i=1}^{N-1} |(\hat{B}^N)_i| - |(\hat{B}^N)_k|^2 + \tilde{\alpha}_N |(\hat{B}^N)_k (\hat{B}^N)_N|.$$

In view of the result for the finite difference approximation we set

$$\tilde{\alpha}_N = \frac{4}{\pi^2} N(N^2 - 1)\beta. \tag{2.86}$$

Therefore we get by (2.68) and (2.72)

$$R_N^N \leq \frac{9}{\tilde{\alpha}_N} \left[\frac{1}{2N} \ln \frac{2(N+1)}{3} + \frac{\pi}{6N} \right] \tag{2.87}$$

where we also used the simple estimate

$$1 \leq \frac{6}{6 + \tilde{\lambda}_{D,k}^N} \leq 3. \tag{2.88}$$

Analogously, one may derive the bounds

$$R_{N-1}^N \leq \frac{18}{(6 + \tilde{\lambda}_{D,N-1}^N)} \left[\frac{1}{N-1} \ln \frac{2(N+1)}{3} + \frac{\pi}{3(N-1)} \right] + \frac{36\beta}{(6 + \tilde{\lambda}_{D,N}^N)(6 + \tilde{\lambda}_{D,N-1}^N)}$$

and

$$R_k^N \leq 9 \frac{N-k+1}{k} \left[\frac{1}{2} \ln \frac{2(N+1)}{3} + \frac{\pi}{6} \right] - \frac{1}{N+1} \frac{2}{\pi^2} \left(\frac{N+1-k}{k} \right)^2 + \frac{N-k+1}{k} (N-1) \cdot \frac{18\beta}{(6 + \tilde{\lambda}_{D,N}^N)(6 + \tilde{\lambda}_{D,k}^N)}.$$

We also need estimates of $|\lambda_{E,N}^N - \lambda_{E,k}^N|$. These will follow from

$$|\lambda_{E,N}^N - \lambda_{E,k}^N| = 36(N+1)^2 \frac{|\tilde{\lambda}_{D,N}^N - \tilde{\lambda}_{D,k}^N|}{(6+\tilde{\lambda}_{D,N}^N)(6+\tilde{\lambda}_{D,k}^N)}. \tag{2.89}$$

Therefore, (2.78) implies

$$|\lambda_{E,N}^N - \lambda_{E,k}^N| \geq 36 \cdot 4 \cdot \frac{(N-k)(N-k+2)}{(6+\tilde{\lambda}_{D,N}^N)(6+\tilde{\lambda}_{D,k}^N)}, \tag{2.90}$$

for $k = 1, \ldots, N-2$. Next we establish

$$|(\hat{B}^N)_N|^2 + R_N^N < 2, \quad N \geq 8.$$

In view of (2.87), (2.88) and the definition of $\hat{B}^N$ one obtains

$$|(\hat{B}^N)_N|^2 + R_N^N < 9 \left(\frac{\pi^2}{4\beta N(N^2-1)} \left(\frac{1}{2N} \ln \frac{2(N+1)}{3} + \frac{\pi}{6N} \right) + \frac{1}{2(N+1)} \frac{\pi^2}{4} \frac{1}{N^2} \right),$$

which using $\frac{1}{x} \ln x \leq e^{-1}$ for $x \geq 1$, $N \geq 8$ and $\beta > 1$ yields the desired bound. At this point we want to choose β so that

$$|\lambda_{E,N}^N - \lambda_{E,k}^N|^2 + (\hat{B}^N)_k^2 - R_k^N > 2,$$

for $k = 1, \ldots, N-4$. An argument similar to the one employed for the finite difference approximation shows that this can be done if one can choose β satisfying

$$36^2 \cdot 16 \frac{(N-k)^2(N-k+2)k}{(6+\tilde{\lambda}_{D,N}^N)(6+\tilde{\lambda}_{D,k}^N)} - \left(3(N+1)e^{-1} + \frac{2\pi}{3} \right) (6+\tilde{\lambda}_{D,N}^N)(6+\tilde{\lambda}_{D,k}^N)$$
$$-18\beta(N-1) > \frac{2}{5}(N-4)(6+\tilde{\lambda}_{D,N}^N)(6+\tilde{\lambda}_{D,k}^N),$$

for $k = 1, \ldots, N-4$. It is easy to see that the right hand side is bounded above by

$$6(N-4) \quad \text{for} \quad N \geq 8,$$

and the left hand side is bounded below by

$$6 \cdot 36 \cdot 8 \left(1 + \frac{\pi^2}{2 \cdot 81} \right)^{-1} \cdot (N-k)^2(N-k+2)k$$
$$- 9 \left(\frac{e^{-1}}{3}(N+1) + \frac{\pi}{6} \right) \cdot 12 \cdot (1 + \frac{\pi^2}{2 \cdot 81}) - 18\beta(N-1).$$

Combining these two estimates and simplifying we find that it is sufficient to show that

$$\begin{aligned} &1628(N-k)^2(N-k+2)k - 21N - 51 - 18\beta N \\ &+ 18\beta > 0, \quad \text{for} \quad k = 1, \dots, N-4, \quad N \geq 8. \end{aligned} \tag{2.91}$$

An argument similar to the one used for the finite difference approximation yields that the left hand side of (2.91) achieves its minimum for $k = N - 4$. Therefore, (2.91) is satisfied if we choose β such that

$$1628 \cdot 96 \cdot (N-4) - 21N - 51 - 18\beta N - 18\beta > 0, \quad N \geq 8.$$

Since the left hand side is increasing in N (if β is not too big) we infer that an appropriate choice of β is given by

$$\beta = 4950.$$

Up to now we have shown that $\mathcal{E}_N^N \cap \mathcal{E}_k^N = \varnothing, k = 1, \dots, N-4$. It is somewhat tedious but straightforward (using the estimates derived so far) to show that $\mathcal{E}_N^N \cap \mathcal{E}_k^N = \varnothing$ for $k = N-3, N-2$ and $N-1$. This completes the proof of Theorem 8.

Table 2.1. Robustness measure of controllability for $y_t = y_{xx} + x \cdot u(t)$. Finite difference approximation.

N	lower bound	upper bound	$\left(\gamma_{Dc}^N\right)^2$ calculated
2	0.55470×10^{-1}	0.55556×10^{-1}	0.55470×10^{-1}
3	0.21428×10^{-1}	0.21447×10^{-1}	0.21434×10^{-1}
4	0.10551×10^{-1}	0.10557×10^{-1}	0.10554×10^{-1}
5	0.59801×10^{-2}	0.59831×10^{-2}	0.59821×10^{-2}
6	0.37195×10^{-2}	0.37211×10^{-2}	0.37207×10^{-2}
7	0.24720×10^{-2}	0.24729×10^{-2}	0.24727×10^{-2}
8	0.17261×10^{-2}	0.17273×10^{-2}	0.17272×10^{-2}
9	0.12535×10^{-2}	0.12543×10^{-2}	0.12542×10^{-2}
10	0.93912×10^{-3}	0.93965×10^{-3}	0.93962×10^{-3}
20	0.13367×10^{-3}	0.13371×10^{-3}	0.13371×10^{-3}
30	0.41474×10^{-4}	0.41483×10^{-4}	0.41483×10^{-4}
40	0.17915×10^{-4}	0.17918×10^{-4}	0.17918×10^{-4}
50	0.93050×10^{-5}	0.93062×10^{-5}	0.93062×10^{-5}
60	0.54370×10^{-5}	0.54377×10^{-5}	0.54377×10^{-5}
70	0.34477×10^{-5}	0.34481×10^{-5}	0.34481×10^{-5}
80	0.23218×10^{-5}	0.23220×10^{-5}	0.23220×10^{-5}
90	0.16373×10^{-5}	0.16375×10^{-5}	0.16375×10^{-5}
100	0.11975×10^{-5}	0.11976×10^{-5}	0.11976×10^{-5}

Table 2.2. Robustness measure of controllability for $y_t = y_{xx} + x \cdot u(t)$. Finite difference approximation.

N	lower bound	upper bound	$\left(\gamma_{Dc}^N\right)^2$ calculated
2	0.222136×10^0	0.222222×10^0	0.222137×10^0
3	0.115457×10^0	0.115472×10^0	0.115459×10^0
4	0.066981×10^0	0.066986×10^0	0.066983×10^0
5	0.418736×10^{-1}	0.418755×10^{-1}	0.418745×10^{-1}
6	0.277254×10^{-1}	0.277263×10^{-1}	0.277259×10^{-1}
7	0.192182×10^{-1}	0.192187×10^{-1}	0.192186×10^{-1}
8	0.138272×10^{-1}	0.138275×10^{-1}	0.138274×10^{-1}
9	0.102595×10^{-1}	0.102597×10^{-1}	0.102596×10^{-1}
10	0.078111×10^{-1}	0.078112×10^{-1}	0.078112×10^{-1}
20	0.117693×10^{-2}	0.117698×10^{-2}	0.117698×10^{-2}
30	0.369533×10^{-3}	0.369544×10^{-3}	0.369544×10^{-3}
40	0.160314×10^{-3}	0.160318×10^{-3}	0.160317×10^{-3}
50	0.834376×10^{-4}	0.834393×10^{-4}	0.834393×10^{-4}
60	0.488086×10^{-4}	0.488094×10^{-4}	0.488094×10^{-4}
70	0.309715×10^{-4}	0.309720×10^{-4}	0.309720×10^{-4}
80	0.208664×10^{-4}	0.208667×10^{-4}	0.208667×10^{-4}
90	0.147195×10^{-4}	0.147196×10^{-4}	0.147196×10^{-4}
100	0.107680×10^{-4}	0.107681×10^{-4}	0.107681×10^{-4}

Table 2.3. Stabilizability and detectability radii.

N	γ_{Es}^N	$\tilde{\gamma}_{Es}^N$	γ_{Ed}^N	$\tilde{\gamma}_{Ed}^N$
2	10.8333	10.6325	10.8036	10.6325
4	10.2512	9.9300	10.2001	9.9300
8	10.0631	9.6515	9.9708	9.6515
10	10.0499	9.6098	9.9373	9.6098
12	10.0509	9.5857	9.9180	9.5857
14	10.0590	9.5705	9.9060	9.5705
16	10.0711	–	9.8979	–

Table 2.4. Controllability and observability radii.

N	γ_{Ec}^N	$\tilde{\gamma}_{Ec}^N$	γ_{Eo}^N	$\tilde{\gamma}_{Eo}^N$
2	0.47131	0.33469	0.16667	0.22899
4	0.25881	0.22885	0.07940	0.06555
8	0.11759	0.15311	0.03927	0.00663
10	0.08838	0.13509	0.03153	0.00196
12	0.06942	0.12229	0.02639	0.00060
14	0.05634	0.11264	0.02271	0.00020
16	0.04688	–	0.01994	–

Table 2.5. Stabilizability radii and Riccati condition number. Model Problem.

δ	γ_s	$1/\gamma_s$	K_{RIC}	$K_{RIC} * \gamma_s$
$1e^{0}$	$8.6603e^{-1}$	$1.1547e^{0}$	$2.2788e^{0}$	1.9735
$1e^{-1}$	$8.9425e^{-2}$	$1.1183e^{+1}$	$1.2716e^{+1}$	1.1371
$1e^{-2}$	$8.9443e^{-3}$	$1.1180e^{+2}$	$1.2130e^{+2}$	1.0850
$1e^{-3}$	$8.9443e^{-4}$	$1.1180e^{+3}$	$1.2072e^{+3}$	1.0802
$1e^{-4}$	$8.9443e^{-5}$	$1.1180e^{+4}$	$1.2071e^{+4}$	1.0797
$1e^{-5}$	$8.9433e^{-6}$	$1.1180e^{+5}$	$1.2071e^{+5}$	1.0797
$1e^{-6}$	$8.9433e^{-7}$	$1.1180e^{+6}$	$1.2071e^{+6}$	1.0797

References

[1] Arnold, W. F. and Laub, A. J., *Generalized algorithms and software for algebraic Riccati equations*, Proc. IEEE **12** (1984), 1746–1754.

[2] Banks, H. T., Jacobs, M. Q. and Langenhop, C. E., Characterization of the controlled states in $W_2^{(1)}$ of linear hereditary systems, SIAM J. Control Opt. **13** (1975), 611–649.

[3] Banks, H. T. and Kunisch, K., *The linear regulator problem for parabolic systems*, SIAM J. Control Opt. **22** (1984), 684–697.

[4] Boley, D. L., *Computing rank-deficiency of rectangular matrix pencils*, Systems & Control Letters **9** (1987), 207–214.

[5] Boley, D. L. and Wu-Sheng Lu, *Measuring how far a controllable system is from an uncontrollable one*, IEEE Trans. Automat. Control **AC-31** (1986), 249–251.

[6] Borggaard, J., Burns, J. A. and Zietsman, L., *Computational challenges in control of partial differential equations*, Proc. High Performance Computing (2004), 51–56.

[7] Burns, J. A., King, B. B., Krueger, D. and Zietsman, L., *Computation of feedback operators for distributed parameter systems with non-normal linearizations*, 2003 American Control Conference (2003), Paper TA04-2.

[8] Burns, J. A. and Peichl, G. H., *Preservation of controllability under approximation*, Differential and Integral equations **2** (1989), 439–452.

[9] Burns, J. A. and Peichl, G. H., *A note on the asymptotic behavior of controllability radii for a scalar hereditary system*, Proceedings of the 4th International Conference on Control of Distributed Parameter Systems, (Kappel F., Kunisch K., Schappacher W., eds.), Birkhäuser Verlag, 1989, 27–39.

[10] Burns, J. A. and Peichl, G. H., *On robustness of controllability for finite dimensional approximations of distributed parameter systems*, Proceedings of the IMACS/IFAC International Symposion on modeling and simulation of distributed parameter systems, (Sunahara, T., Tzafestas, S.G. and Futagami T., eds.), Hiroshima Institute of Technology, 1987, 491–496.

[11] Curtain, R. C. and Pritchard, A. J., *Infinite dimensional linear systems theory*, Lecture Notes in Control and Information Sciences 7, Springer, Berlin, 1978.

[12] Datta, B., *Numerical Methods for Linear Control Systems Design and Analysis*, Elsevier Academic Press, 2003.

[13] Demmel, J. W., *On condition numbers and the distance to the nearest ill posed problem*, Numer. Math. **51** (1987), 251–289.

[14] Eising, R., *Between controllable and uncontrollable*, Systems & Control Lett. **4** (1984), 263–264.

[15] Eising, R., *The distance between a system and the set of uncontrollable systems*, Lecture Notes in Control and Information Sciences, Proc. MTNS-83, Beer Sheva, vol. 58, Springer, New York, 1984, 303–314.

[16] Fattorini, H. O., *On complete controllability of linear systems*, JDE **3** (1967), 391–462.

[17] Gibson, J. S., *Linear quadratic optimal control of hereditary differential systems: Infinite dimensional Riccati equations and numerical approximations*, SIAM J. Control Opt. **21** (1983), 95–139.

[18] Gahinet, P. and Laub A. J., *Algebraic Riccati equations and the distance to the nearest uncontrollable pair*, SIAM J. Control Opt. **30** (1992), 765–786.

[19] Golub, G. H. and van Loan, C. F., *Matrix Computations*, The John Hopkins University Press, Baltimore, 1984.

[20] Gradstein, I. S. and Ryshik, I. R., *Tables of Series, Products and Integrals*, Vol I Verlag Harri Deutsch, Frankfurt, 1981.

[21] Hinrichsen, D. and Motscha, M., *Optimization problems in the robustness analysis of linear state space systems*, Approximation and Optimization, (Gomez, A., Guerra, F., Jimenez, M. A. and Lopez, G., eds.), Lecture Notes in Mathematics, 1354 Springer, 1988, 54–78.

[22] Hinrichsen, D. and Pritchard, A. J., *Stability radii for linear systems*, Systems & Control Letters **7** (1986), 1–10.

[23] Hinrichsen, D. and Pritchard, A. J., *Stability radiius for structured perturbations and the algebraic Riccati equation*, Systems & Control Letters **8** (1986), 105–113.

[24] Ito, K. and Kappel, F., *A uniformly differentiable approximation scheme for delay systems using splines*, J. Appl. Math. Opt. **23** (1991) 217–262.

[25] El Jai, A. and Pritchard, A. J., *Sensors and Controls in the Analysis of Distributed Systems*, J. Wiley, 1988.

[26] Kappel, F. and Peichl, G. H., *Preservation of controllability under approximation for delay systems*, Matematica Aplicada e Computational **8** (1989), 23–47.

[27] Kappel, F. and Salamon, D., *On the stability properties of spline approximations for retarded systems*, SIAM J. Control Opt. **27** (1989), 407–431.

[28] Kailath, T., *Linear Systems*, Englewood–Cliffs, Prentice Hall, 1980.

[29] Kenney, C. and Laub, A. J., *Controllability and stability radii for companion form systems*, Math. Control Signal Systems **1** (1988), 239–256.

[30] Laub, A. J., *Numerical linear algebra aspects of control design computations*, IEEE Trans. Automat. Control **AC-30** (1985), 97–108.

[31] Lancaster, P. and Tismenetsky, M., *The theory of Matrices*, Academic Press, Orlando, 1985.

[32] Lasiecka, I. and Triggiani, R., *Differential and algebraic Riccati equations with application to boundary /point control problems: continuous theory and approximation theory*, Lecture Notes in Control and Information Sciences 164, Springer, Berlin, 1978.

[33] Lewkowicz, I., *When are the complex and the real stability radii equal*, IEEE Transactions on Automatic Control **AC-37** (1992), 880–883.

[34] Miminis, G., *Numerical algorithms for controllability and eigenvalue computation*, Masters thesis, School of Computer Science, McGill Univ., 1981.

[35] Oates, K., *A study of control system radii for approximations of infinite dimensional systems*, MS thesis, Department of Mathematics, VPI–SU, Blacksburg, VA, 1991, ICAM Report 91-08-01.

[36] Paige, C., *Properties of numerical algorithms related to computing controllability*, IEEE Trans. Automat. Control **AC-26** (1981), 130–139.

[37] Pandolfi, L., *Controllability properties of perturbed distributed parameter systems*, Lin. Alg. Appl. **122** (1989), 525–538.

[38] Peichl, G. H. and Wang, C.: *On the uniform stabilizability and the margin of stabilizability of the finite dimensional approximations of distributed parameter systems*, J. Math. Systems Estimation and Control **7** (1997), 277–304.

[39] Peichl, G. H. and Wang, C.: *Asymptotic Analysis of stabilizability of a control system for a discretized damped wave equation*, Numerical Functional Analysis **19** (1998), 91-113.

[40] Pritchard, A. J. and Townley, S., *A stability radius for infinite dimensional systems*, Distributed parameter systems, (F.Kappel, K.Kunisch, W.Schappacher, eds.), Lecture Notes in Control and Information Sciences 102, 1986, 272–291.

[41] Salamon, D., *Controllability and observability of time delay systems*, IEEE Trans. Automat. Control **AC-29** (1984), 432–439.

[42] Smith, G. D., *Numerical solution of partial differential equations*, Clarendon Press, Oxford, 1976.

[43] Wilkinson, J. H., *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

[44] Wicks, M. and DeCarlo, R., *Computing the distance to an uncontrollable system*, Proceedings of the 28th Conference on Decision and Control, Tampa, Fl, 1989, 2361–2366.

[45] Wicks, M. and DeCarlo, R., *Computing the distance to an uncontrollable system*, IEEE Transactions on Automatic Control **AC-36** (1991), 39–49.

http://www.springer.com/978-0-387-28263-3

Robust Optimization-Directed Design
Kurdila, A.J.; Pardalos, P.; Zabarankin, M. (Eds.)
2006, IX, 275 p., Hardcover
ISBN: 978-0-387-28263-3