

ANALYTIC–NUMERICAL SOLUTIONS WITH *A PRIORI*
BOUNDS FOR MATRIX–VECTOR PARABOLIC
PARTIAL DIFFERENTIAL EQUATIONS

LUCAS JÓDAR¹ AND MARCO MARLETTA²

¹*Departamento de Matemática Aplicada, Universidad Politécnica de Valencia,
PO Box 22.012, 46071 Valencia, Spain*

²*Department of Mathematics and Computer Science, University of Leicester,
University Road, Leicester LE1 7RH, UK*

(Received 23 March 2000)

Abstract We consider a parabolic matrix–vector system in which the diffusion matrix may be time dependent. For the time-independent case we construct approximate solutions with guaranteed error bounds using spectral information from certain matrix–vector Sturm–Liouville problems. For the time-dependent case we employ an approximation procedure which reduces the problem, on each time-step, to the time-independent case. We give an algorithm which may be used *a priori* at each time-step in the time-dependent case to guarantee accuracy to a specified tolerance.

Keywords: Sturm–Liouville problem; eigenfunction; eigenvalue inequality;
computable *a priori* error bounds

AMS 2000 *Mathematics subject classification:* Primary 35C10; 35M10
Secondary 15A

1. Introduction

This paper grows out of the work of Jódar, Navarro and Camacho [6] and Martín, Navarro and Jódar [7]. We shall consider a system of parabolic partial differential equations (PDEs) of the form

$$\left. \begin{aligned} u_t &= Au_{xx}, & 0 < x < 1, \quad t > 0, \\ B_1 u(0, t) + B_2 u_x(0, t) &= \mathbf{0} \in \mathbb{C}^n, & t > 0, \\ C_1 u(1, t) + C_2 u_x(1, t) &= \mathbf{0} \in \mathbb{C}^n, & t > 0, \\ u(x, 0) &= f(x), & 0 \leq x \leq 1. \end{aligned} \right\} \quad (1.1)$$

In this system the matrix A is an $n \times n$ Hermitian positive-definite complex matrix, and may depend on t . We shall treat separately the cases of constant A and time-dependent A . The matrices B_1 , B_2 , C_1 and C_2 appearing in the boundary conditions are complex $n \times n$ matrices, constant in time, such that certain associated boundary-value problems

are self-adjoint: to ensure this we shall assume the Atkinson conditions [2]

$$\left. \begin{aligned} B_1 B_2^* &= B_2 B_1^*, \\ C_1 C_2^* &= C_2 C_1^*, \end{aligned} \right\} \quad (1.2)$$

together with the full-rank hypothesis

$$\text{rank}(B_1^* + iB_2^*) = n = \text{rank}(C_1^* + iC_2^*). \quad (1.3)$$

Equation (1.3) is necessary to ensure that the boundary conditions are linearly independent. The solution u of (1.1) will be an n -vector valued map and, for smoothness, we assume that the initial condition f satisfies the boundary conditions.

Coupled parabolic systems of PDEs arise in a number of applications, the most obvious being those which involve coupled diffusion problems (see, for example, [3, 9, 10]). The equations which arise in quantum mechanical scattering (see, for example, [1, 8]) can also often be expressed in a form similar to (1.1) and solved by an approach similar to the one we discuss in this paper, although it is important to note that quantum mechanical scattering equations arise from the time-dependent Schrödinger equation and therefore involve an anti-Hermitian coefficient matrix A in the term Au_{xx} . This means that the error analysis is somewhat different from the parabolic case which we discuss here.

Compared with [6] and [7], the approach which we take here has a number of advantages. Firstly, we do not need to assume that the kernel of a certain linear combination of the matrices arising in the boundary conditions is invariant under A , or the associated restriction on the initial condition. (On the other hand, we do impose self-adjointness conditions which were not assumed in [6] or in [7].) Secondly, we have a much simpler strategy for dealing with the time dependence of the solution, which does not involve expansion of the solution in a power series in t and does not therefore suffer from the problems associated with a limited radius of convergence. Our method is able to capture correctly the long-time behaviour expected of a parabolic system, although the formal order of accuracy in the time-step is lower than for the methods in [6] and [7], so the method is less suitable for short-time integrations.

Our approach relies crucially on *a priori* lower bounds for the eigenvalues of the matrix–vector Sturm–Liouville problem

$$\left. \begin{aligned} -\mathbf{y}_{xx} &= \lambda A^{-1} \mathbf{y} \in \mathbb{C}^n, \quad 0 < x < 1, \\ B_1 \mathbf{y}(0) + B_2 \mathbf{y}_x(0) &= \mathbf{0}, \\ C_1 \mathbf{y}(1) + C_2 \mathbf{y}_x(1) &= \mathbf{0}. \end{aligned} \right\} \quad (1.4)$$

We obtain suitable bounds in §2 below by a simple variational argument. Section 3 obtains *a priori* error bounds for a truncated series solution of the PDE in the case where A does not depend on t . In §4 we deal with the case of time-dependent A and obtain an *a priori* error bound which depends only on knowledge of A and dA/dt .

2. A-priori eigenvalue and eigenfunction bounds for a matrix–vector Sturm–Liouville problem

Let $(\lambda_k, \mathbf{y}_k)$, $k = 0, 1, 2, \dots$, denote the eigenvalues and eigenfunctions of the problem (1.4). The eigenfunctions are orthogonal with respect to the inner product

$$\langle f, g \rangle := \int_0^1 f(x)^* A^{-1} g(x) \, dx, \tag{2.1}$$

and of unit length in the associated norm

$$\|f\|_A := \langle f, f \rangle^{1/2}. \tag{2.2}$$

The set of vectors $\mathbf{y}_0(0), \mathbf{y}_1(0), \dots, \mathbf{y}_{k+2n}(0)$ is a set of $k + 2n + 1$ vectors in \mathbb{C}^n , so there exist $k + n + 1$ linearly independent combinations of these vectors which are zero: in other words, there exist $k + n + 1$ linearly independent vectors

$$c_j := (c_{0j}, c_{1j}, \dots, c_{k+2n,j})^T, \quad j = 1, \dots, k + n + 1,$$

such that

$$\sum_{\nu=0}^{k+2n} c_{\nu j} \mathbf{y}_\nu(0) = \mathbf{0}, \quad j = 1, \dots, k + n + 1.$$

We define functions

$$\mathbf{Y}_j(x) = \sum_{\nu=0}^{k+2n} c_{\nu j} \mathbf{y}_\nu(x), \quad j = 1, \dots, k + n + 1.$$

The \mathbf{Y}_j are linearly independent functions and satisfy

$$\mathbf{Y}_j(0) = \mathbf{0}, \quad j = 1, \dots, k + n + 1.$$

Now the vectors $\mathbf{Y}_1(1), \dots, \mathbf{Y}_{k+n+1}(1)$ are $k+n+1$ vectors in \mathbb{C}^n , so there exist $k+1$ linear combinations of these vectors which are zero. Following the same reasoning as before, we can thus construct linearly independent functions $\mathbf{Z}_1, \dots, \mathbf{Z}_{k+1}$ in $\text{span}\{\mathbf{y}_0, \dots, \mathbf{y}_{k+2n}\}$ such that

$$\begin{aligned} \mathbf{0} &= \mathbf{Z}_1(0) = \mathbf{Z}_2(0) = \dots = \mathbf{Z}_{k+1}(0), \\ \mathbf{0} &= \mathbf{Z}_1(1) = \mathbf{Z}_2(1) = \dots = \mathbf{Z}_{k+1}(1). \end{aligned}$$

Let V be the space of C^2 functions on $[0, 1]$ which are n -vector valued and satisfy $\mathbf{z}(0) = \mathbf{0} = \mathbf{z}(1)$. By the Rayleigh variational characterization of eigenvalues [4, D2, p. 1543] the k th eigenvalue of the problem

$$\left. \begin{aligned} -\mathbf{y}_{xx} &= \lambda A^{-1} \mathbf{y} \in \mathbb{C}^n, & 0 < x < 1, \\ \mathbf{y}(0) &= \mathbf{0}, \\ \mathbf{y}(1) &= \mathbf{0}, \end{aligned} \right\} \tag{2.3}$$

is given by

$$\lambda_k(\text{Dirichlet}) = \inf_{\substack{V_k \subseteq V \\ \dim V_k = k+1}} \left(\sup_{z \in V_k} \frac{\langle Az, -z'' \rangle}{\langle z, z \rangle} \right),$$

where $\langle \cdot, \cdot \rangle$ is defined by (2.1). In particular, choosing $V_k = \text{span}\{\mathbf{Z}_1, \dots, \mathbf{Z}_{k+1}\} \subseteq V$ gives

$$\lambda_k(\text{Dirichlet}) \leq \sup_{z \in \text{span}\{\mathbf{Z}_1, \dots, \mathbf{Z}_{k+1}\}} \frac{\langle Az, -z'' \rangle}{\langle z, z \rangle}. \tag{2.4}$$

Since $\text{span}\{\mathbf{Z}_1, \dots, \mathbf{Z}_{k+1}\} \subseteq \text{span}\{\mathbf{y}_0, \dots, \mathbf{y}_{k+2n}\}$ and since

$$\frac{\langle Az, -z'' \rangle}{\langle z, z \rangle} \leq \frac{\langle A\mathbf{y}_{k+2n}, -\mathbf{y}_{k+2n}'' \rangle}{\langle \mathbf{y}_{k+2n}, \mathbf{y}_{k+2n} \rangle} = \lambda_{k+2n}, \quad \forall z \in \text{span}\{\mathbf{y}_0, \dots, \mathbf{y}_{k+2n}\}, \tag{2.5}$$

it follows on combining (2.4) and (2.5) that

$$\lambda_k(\text{Dirichlet}) \leq \lambda_{k+2n}, \quad k = 0, 1, 2, \dots, \tag{2.6}$$

or equivalently

$$\lambda_k \geq \lambda_{k-2n}(\text{Dirichlet}), \quad k = 2n, 2n + 1, \dots$$

To complete our lower bound we find lower bounds for the Dirichlet eigenvalues. These eigenvalues are the eigenvalues of (2.3) and are therefore all positive, and at least as big as the eigenvalues of the problem

$$\left. \begin{aligned} -\alpha^2 \mathbf{y}'' &= \lambda \mathbf{y}, & 0 < x < 1, \\ \mathbf{y}(0) = \mathbf{0} &= \mathbf{y}(1), \end{aligned} \right\} \tag{2.7}$$

where

$$\alpha^2 := \text{least eigenvalue of } A, \quad \alpha > 0.$$

By direct substitution into the differential equation and boundary conditions, the reader may verify that the eigenvalues of (2.7) are given by

$$\hat{\lambda}_k = \left(\left[\frac{k}{n} \right] + 1 \right)^2 \pi^2 \alpha^2, \quad k = 0, 1, 2, \dots, \tag{2.8}$$

with corresponding eigenfunctions which may be taken to be*

$$\mathbf{w}_k(x) = \sin\left(\frac{x}{\alpha} \sqrt{\hat{\lambda}_k}\right) \mathbf{e}_\ell, \quad \ell = 1 + k - n \left[\frac{k}{n} \right],$$

in which \mathbf{e}_ℓ denotes the ℓ th standard basis vector of \mathbb{C}^n .

Combining (2.6) with the inequality $\lambda_k(\text{Dirichlet}) \geq \hat{\lambda}_k$ and (2.8), we see that the following has been proved.

* Since each eigenvalue has an eigenspace of dimension n , the eigenfunctions are not uniquely determined.

Lemma 2.1. *The eigenvalues of (1.4) satisfy*

$$\lambda_k \geq \left(\left[\frac{k}{n} \right] - 1 \right)^2 \pi^2 \alpha^2, \quad k = 2n, 2n + 1, \dots, \tag{2.9}$$

where $\alpha^2 = \text{minimum eigenvalue of } A$.

The second ingredient which we require for our analysis is a bound on the sup-norm of the eigenfunction \mathbf{y}_k . For $\lambda > 0$ let

$$W(\lambda) = \lambda^{1/2} A^{-1/2}. \tag{2.10}$$

Then, if $\lambda_k > 0$, the eigenfunction \mathbf{y}_k has the form

$$\mathbf{y}_k(x) = \{\sin W(\lambda_k)x\}v_1^{(k)} + \{\cos W(\lambda_k)x\}v_2^{(k)} \tag{2.11}$$

for some vectors $v_1^{(k)}$ and $v_2^{(k)}$ in \mathbb{C}^n . For brevity we write this as

$$\mathbf{y}_k(x) = \{\sin Wx\}v_1 + \{\cos Wx\}v_2. \tag{2.12}$$

To obtain a bound on the sup-norm of \mathbf{y}_k we need bounds on $\|v_1\|$ and $\|v_2\|$, to be obtained from the normalization $\|\mathbf{y}_k\|_A = 1$, where $\|\cdot\|_A$ is the norm associated with the inner product (2.1). Now using the fact that A commutes with W , we have

$$\begin{aligned} \mathbf{y}_k(x)^* A^{-1} \mathbf{y}_k(x) &= v_2^* A^{-1} \cos^2(Wx)v_2 + v_1^* A^{-1} \sin^2(Wx)v_1 \\ &\quad + 2 \operatorname{Re}(v_1^* A^{-1} \sin(Wx) \cos(Wx)v_2). \end{aligned}$$

Integrating over $[0, 1]$ with respect to x yields

$$\begin{aligned} \|\mathbf{y}_k\|_A^2 &= \frac{1}{2} v_2^* A^{-1} [I + \frac{1}{2} W^{-1} \sin 2W] v_2 + \frac{1}{2} v_1^* A^{-1} [I - \frac{1}{2} W^{-1} \sin 2W] v_1 \\ &\quad + \frac{1}{2} \operatorname{Re}(v_1^* A^{-1} W^{-1} [I - \cos 2W] v_2). \end{aligned}$$

The left-hand side is equal to 1; bounding the right-hand side from below using the triangle inequality and the inequality $\|I - \cos 2W\| \leq 2$, we obtain

$$\frac{1}{2} (\|A^{-1/2} v_2\|^2 + \|A^{-1/2} v_1\|^2) (1 - \frac{1}{2} \|W^{-1}\|) - \|W^{-1}\| \|A^{-1/2} v_1\| \|A^{-1/2} v_2\| \leq 1. \tag{2.13}$$

Using Young's inequality,

$$2 \|A^{-1/2} v_1\| \|A^{-1/2} v_2\| \leq \|A^{-1/2} v_2\|^2 + \|A^{-1/2} v_1\|^2,$$

to deal with the last term on the left-hand side of (2.13), we obtain, for $\lambda > \frac{9}{4} \|A^{1/2}\|^2$, the inequality

$$\|A^{-1/2} v_1\|^2 + \|A^{-1/2} v_2\|^2 \leq \frac{2}{1 - \frac{3}{2} \lambda^{-1/2} \|A^{1/2}\|},$$

whence

$$\|v_1\|^2 + \|v_2\|^2 \leq \frac{2 \|A\|}{1 - \frac{3}{2} \lambda^{-1/2} \|A^{1/2}\|}, \quad \lambda > \frac{9}{4} \|A^{1/2}\|^2. \tag{2.14}$$

The equation $\mathbf{y}_k(x) = (\sin Wx)v_1 + (\cos Wx)v_2$ implies

$$\|\mathbf{y}_k(x)\| \leq \|v_1\| + \|v_2\| \leq \sqrt{2}\sqrt{\|v_1\|^2 + \|v_2\|^2}. \quad (2.15)$$

Combining (2.14) and (2.15) we have proved the following result.

Lemma 2.2. *Let k be sufficiently large to ensure that $\lambda_k > \frac{9}{4}\|A^{1/2}\|^2$. Then the eigenfunction \mathbf{y}_k normalized by the condition $\|\mathbf{y}_k\|_A = 1$ satisfies*

$$\sup_{0 \leq x \leq 1} \|\mathbf{y}_k(x)\| \leq \frac{2\|A\|^{1/2}}{\sqrt{1 - \frac{3}{2}\lambda_k^{-1/2}\|A^{1/2}\|}}. \quad (2.16)$$

3. Analytic–numerical solution with error bounds: the time-independent case

The solution of (1.1) admits a series representation in terms of the eigenfunctions of (1.4),

$$u(x, t) = \sum_{k=0}^{\infty} c_k e^{-\lambda_k t} \mathbf{y}_k(x), \quad (3.1)$$

where c_k is determined from the eigenfunction expansion of the initial condition: in terms of the inner product (2.1),

$$\langle f, \mathbf{y}_k \rangle = c_k. \quad (3.2)$$

We first examine the effects of truncating the series in (3.1): we let

$$u_N(x, t) = \sum_{k=0}^N c_k e^{-\lambda_k t} \mathbf{y}_k(x) \quad (3.3)$$

and examine the error $\|u(x, t) - u_N(x, t)\|$. Evidently

$$u(x, t) - u_N(x, t) = \sum_{k=N+1}^{\infty} c_k e^{-\lambda_k t} \mathbf{y}_k(x),$$

and so, upon taking (Euclidean) norms,

$$\|u(x, t) - u_N(x, t)\| \leq \sum_{k=N+1}^{\infty} |c_k| e^{-\lambda_k t} \|\mathbf{y}_k(x)\|. \quad (3.4)$$

We now substitute (2.16) into (3.4) and obtain

$$\sup_{0 \leq x \leq 1} \|u(x, t) - u_N(x, t)\| \leq \sum_{k=N+1}^{\infty} |c_k| e^{-\lambda_k t} \frac{2\|A\|^{1/2}}{\sqrt{1 - \frac{3}{2}\lambda_k^{-1/2}\|A^{1/2}\|}}, \quad (3.5)$$

provided $\lambda_{N+1} > \frac{9}{4}\|A^{1/2}\|^2$.

Let $\epsilon > 0$ and $T > 0$ be fixed. We look for an error bound valid over the interval $t \in [T, \infty)$. Suppose that N may be chosen to ensure that

$$\frac{3}{2}\lambda_{N+1}^{-1/2}\|A^{1/2}\| \leq \frac{3}{4}, \quad \sum_{k=N+1}^{\infty} e^{-\lambda_k T} < \frac{\epsilon}{4\|f\|_A\|A\|^{1/2}}. \tag{3.6}$$

Using the inequality $|c_k| \leq \|f\|_A$ which follows from (3.2), we obtain from (3.5) and (3.6) the bound

$$\sup_{0 \leq x \leq 1} \|u(x, t) - u_N(x, t)\| < \epsilon. \tag{3.7}$$

We now use the lower bound on λ_k given by Lemma 2.1 to determine a value for N such that (3.6) will hold. Simple algebra, together with the inequality of Lemma 2.1, shows that the first of the two inequalities holds if

$$N \geq n \left(2 + \left\lceil \frac{2\|A^{1/2}\|}{\pi\alpha} \right\rceil \right). \tag{3.8}$$

For the second inequality we assume that $N + 1 > 2n$, and we use the inequality

$$\lambda_k \geq \left(\frac{k}{n} - 2 \right)^2 \pi^2 \alpha^2.$$

Purely for convenience we also assume that N is large enough to ensure that

$$\left(\frac{N+1}{n} - 2 \right)^2 \pi^2 \alpha^2 T \geq 1. \tag{3.9}$$

Then we have

$$\begin{aligned} \sum_{k=N+1}^{\infty} e^{-\lambda_k T} &\leq \sum_{k=N+1}^{\infty} e^{-((k/n)-2)^2 \pi^2 \alpha^2 T} \\ &\leq \sum_{k=N+1}^{\infty} e^{-((k/n)-2)\pi\alpha\sqrt{T}} \quad \text{by (3.9)} \\ &= \frac{e^{2\pi\alpha\sqrt{T}}}{1 - e^{-\pi\alpha\sqrt{T}/n}} e^{-(N+1)\pi\alpha\sqrt{T}/n}. \end{aligned}$$

Thus the second inequality in (3.6) will hold provided N is sufficiently large to ensure that $N + 1 > 2n$, (3.8) holds, and

$$\frac{e^{2\pi\alpha\sqrt{T}}}{1 - e^{-\pi\alpha\sqrt{T}/n}} e^{-(N+1)\pi\alpha\sqrt{T}/n} < \frac{\epsilon}{4\|f\|_A\|A\|^{1/2}}. \tag{3.10}$$

Solving this inequality and bearing in mind (3.9), we find that a sufficient condition on N is

$$N > \max \left(\frac{n}{\pi\alpha\sqrt{T}} |\log \hat{\epsilon}| - 1, n \left(2 + \frac{1}{\pi\alpha\sqrt{T}} \right) - 1, n \left(2 + \left\lceil \frac{2\|A^{1/2}\|}{\pi\alpha} \right\rceil \right) \right), \tag{3.11}$$

where

$$\hat{\epsilon} = \frac{\epsilon(1 - e^{-\pi\alpha\sqrt{T}/n})}{4\|f\|_A\|A\|^{1/2}e^{2\pi\alpha\sqrt{T}}}. \quad (3.12)$$

We have thus proved the following result.

Theorem 3.1. *Let $T > 0$ and $\epsilon > 0$ be given. Suppose that the solution u of (1.1) given by (3.1) is approximated for $t \in [T, \infty)$ by the function u_N given by (3.3). Suppose, moreover, that N is chosen to satisfy (3.11). Then*

$$\sup_{0 \leq x \leq 1} \|u(x, t) - u_N(x, t)\| < \epsilon, \quad \forall t \in [T, \infty). \quad (3.13)$$

4. Analytic–numerical solution with error bounds: the time-dependent case

4.1. Preliminaries

In this section we shall find it convenient to use the ‘logarithmic seminorm’ $\mu(M)$ of a square matrix M , defined by

$$\mu(M) = \limsup_{\epsilon \searrow 0} \frac{\|I + \epsilon M\| - 1}{\epsilon}. \quad (4.1)$$

We take $\|\cdot\|$ on the right-hand side of (4.1) to be the usual matrix 2-norm; noting that for each $\epsilon > 0$ and vector v ,

$$v^* M v = (1/\epsilon)\{v^*(I + \epsilon M)v - v^*v\},$$

it is then easy to see that

$$v^* M v \leq \mu(M)v^*v. \quad (4.2)$$

4.2. Coefficient approximation and its effect

We approach the time-dependent case by approximating the coefficient $A(t)$ in the system of PDEs by a piecewise-constant matrix $\hat{A}(t)$: so we assume that there exists a sequence

$$0 = t_0 < t_1 < t_2 < \cdots < t_k < t_{k+1} < \cdots,$$

such that $\hat{A}(t) = A_k$ (constant) for $t \in [t_k, t_{k+1})$. The problem with the piecewise-constant coefficient \hat{A} can be solved using the approach of §3. We need to compare the solutions of the following problems:

$$\mathcal{P} : \begin{cases} u_t = A(t)u_{xx}, & 0 < x < 1, \quad t > 0, \\ B_1 u(0, t) + B_2 u_x(0, t) = 0, & t > 0, \\ C_1 u(1, t) + C_2 u_x(1, t) = 0, & t > 0, \\ u(x, 0) = f(x), & 0 \leq x \leq 1, \end{cases} \quad (4.3)$$

$$\hat{\mathcal{P}} : \begin{cases} \hat{u}_t = \hat{A}(t)\hat{u}_{xx}, & 0 < x < 1, \quad t > 0, \\ B_1\hat{u}(0, t) + B_2\hat{u}_x(0, t) = 0, & t > 0, \\ C_1\hat{u}(1, t) + C_2\hat{u}_x(1, t) = 0, & t > 0, \\ \hat{u}(x, 0) = f(x), & 0 \leq x \leq 1. \end{cases} \quad (4.4)$$

To this end we let $v := u - \hat{u}$ and $g = (A - \hat{A})u_{xx}$. Elementary manipulations show that for $t \in [t_k, t_{k+1})$, v is the solution of the problem

$$\mathcal{Q} : \begin{cases} v_t = A_k v_{xx} + g(x, t), & 0 < x < 1, \quad t \in [t_k, t_{k+1}), \\ B_1 v(0, t) + B_2 v_x(0, t) = 0, & t \in [t_k, t_{k+1}), \\ C_1 v(1, t) + C_2 v_x(1, t) = 0, & t \in [t_k, t_{k+1}), \\ v(x, t_k) = v_k(x) := u(x, t_k) - \hat{u}(x, t_k), & 0 \leq x \leq 1. \end{cases} \quad (4.5)$$

To start the error analysis, for $t \in [t_k, t_{k+1})$ we expand v and g in terms of the eigenfunctions of the following eigenproblem:

$$\left. \begin{aligned} -\mathbf{y}'' &= \lambda A_k^{-1} \mathbf{y}, & 0 < x < 1, \\ B_1 \mathbf{y}(0) + B_2 \mathbf{y}'(0) &= 0, \\ C_1 \mathbf{y}(1) + C_2 \mathbf{y}'(1) &= 0. \end{aligned} \right\} \quad (4.6)$$

We let $(\lambda_n, \mathbf{y}_n)$ ($n = 0, 1, \dots$) be the eigenpairs of this problem, normalized so that

$$\langle \mathbf{y}_n, A_k^{-1} \mathbf{y}_m \rangle_k := \int_0^1 \mathbf{y}_n^* A_k^{-1} \mathbf{y}_m \, dx = \delta_{nm}. \quad (4.7)$$

The following assumption, which is essentially a dissipativity assumption, is also quite important at various points in the analysis:

Assumption \mathcal{A} : the eigenvalues λ_j ($j \geq 0$) of (4.6) are all > 0 .

Remark 4.1. Problem (4.6) depends on A_k , which depends on the time-interval $[t_k, t_{k+1})$ under scrutiny. However, since $A(t)$ and $\hat{A}(t)$ are positive-definite Hermitian for all t , it is easy to show by an elementary variational argument that if the assumption holds for some k , then it holds for all k .

We also let

$$v(x, t) = \sum_{j=0}^{\infty} c_j(t) \mathbf{y}_j(x), \quad g(x, t) = \sum_{j=0}^{\infty} \phi_j(t) \mathbf{y}_j(x). \quad (4.8)$$

Substituting these expansions into problem (4.5) we get $c_j'(t) = -\lambda_j c_j(t) + \phi_j(t)$. Multiplying both sides of this equation by $2c_j(t)$ gives

$$\frac{d}{dt}(|c_j|^2) = -2\lambda_j |c_j|^2 + 2 \operatorname{Re}(c_j \phi_j) \leq -2\lambda_j |c_j|^2 + 2 \operatorname{Re} \left(\epsilon |c_j|^2 + \frac{1}{4\epsilon} |\phi_j|^2 \right),$$

for any $\epsilon > 0$. Choosing $\epsilon = \frac{1}{2}\lambda_j$ gives

$$\frac{d}{dt}(|c_j|^2) \leq -\lambda_j|c_j|^2 + \frac{1}{\lambda_j}|\phi_j|^2,$$

which we can write as

$$\frac{d}{dt}(e^{\lambda_j(t-t_k)}|c_j|^2) \leq \frac{1}{\lambda_j}e^{\lambda_j(t-t_k)}|\phi_j(t)|^2.$$

Integrating both sides yields

$$|c_j(t)|^2 \leq |c_j(t_k)|^2 e^{-\lambda_j(t-t_k)} + \frac{1}{\lambda_j^2}(1 - e^{-\lambda_j(t-t_k)}) \sup_{s \in [t_k, t_{k+1})} |\phi_j(s)|^2.$$

Multiplying by λ_j and summing,

$$\begin{aligned} \sum_{j=0}^{\infty} \lambda_j |c_j(t)|^2 &\leq \sum_{j=0}^{\infty} e^{-\lambda_j(t-t_k)} \lambda_j |c_j(t_k)|^2 \\ &\quad + \left(\sum_{j=0}^{\infty} \lambda_j^{-1} (1 - e^{-\lambda_j(t-t_k)}) \right) \sup_{j \geq 0} \sup_{s \in [t_k, t_{k+1})} |\phi_j(s)|^2. \end{aligned} \quad (4.9)$$

Since ϕ_j is the j th generalized Fourier coefficient of g ,

$$|\phi_j(t)| = \left| \int_0^1 \mathbf{y}_j^*(x) A_k^{-1} g(x, t) \, dx \right| \leq \left(\int_0^1 g(x, t)^* A_k^{-1} g(x, t) \, dx \right)^{1/2};$$

substituting $g(x, t) = (A(t) - A_k)u_{xx}(x, t)$ we get

$$\begin{aligned} |\phi_j(t)|^2 &\leq \int_0^1 u_{xx}^*(A(t) - A_k) A_k^{-1} (A(t) - A_k) u_{xx} \, dx \\ &\leq \|A^{-1/2}(A - A_k) A_k^{-1} (A - A_k) A^{-1/2}\| \int_0^1 \|A^{1/2} u_{xx}\|^2 \, dx \\ &= \|A^{-1/2}(A - A_k) A_k^{-1} (A - A_k) A^{-1/2}\| \int_0^1 u_{xx}^* A u_{xx} \, dx. \end{aligned} \quad (4.10)$$

Define

$$E(k, A) := \sup_{t \in [t_k, t_{k+1})} \|A^{-1/2}(A - A_k) A_k^{-1} (A - A_k) A^{-1/2}\|^2. \quad (4.11)$$

Then (4.10) gives

$$|\phi_j(t)|^2 \leq E(k, A)^2 \int_0^1 u_{xx}^* A u_{xx} \, dx. \quad (4.12)$$

Substituting back into (4.9) we obtain

$$\begin{aligned} \sum_{j=0}^{\infty} \lambda_j |c_j(t)|^2 &\leq e^{-\lambda_0(t-t_k)} \sum_{j=0}^{\infty} \lambda_j |c_j(t_k)|^2 \\ &+ \left(\sum_{j=0}^{\infty} \lambda_j^{-1} (1 - e^{-\lambda_j(t-t_k)}) \right) E(k, A)^2 \\ &\times \sup_{t \in [t_k, t_{k+1})} \int_0^1 u_{xx}(x, t)^* A(t) u_{xx}(x, t) \, dx. \end{aligned} \tag{4.13}$$

To continue with our analysis we require an *a priori* bound on the quantity

$$\int_0^1 u_{xx}(x, t)^* A(t) u_{xx}(x, t) \, dx, \tag{4.14}$$

which appears on the right-hand side of (4.13). We define

$$w(x, t) := A(t) u_{xx}(x, t) = u_t,$$

and observe that since $w = u_t$, provided u is sufficiently smooth w will satisfy the same boundary conditions as u :

$$B_1 w(0, t) + B_2 w_x(0, t) = 0, \quad C_1 w(1, t) + C_2 w_x(1, t) = 0. \tag{4.15}$$

Also, we observe that

$$\begin{aligned} \int_0^1 u_{xx}(x, t)^* A(t) u_{xx}(x, t) \, dx &= \int_0^1 w(x, t)^* A(t)^{-1} w(x, t) \, dx \\ &= \langle w, w \rangle, \end{aligned} \tag{4.16}$$

in terms of the inner product (2.1). Differentiating with respect to t gives

$$\frac{d}{dt} \left(\frac{1}{2} \langle w, w \rangle \right) = \langle w, w_t \rangle - \frac{1}{2} \int_0^1 w^* A^{-1} A' A^{-1} w \, dx. \tag{4.17}$$

Elementary manipulations show that in addition to the boundary conditions (4.15), the function w satisfies the system

$$w_t = A(t) w_{xx} + A' A^{-1} w. \tag{4.18}$$

Taking the inner product of both sides with w yields

$$\langle w, w_t \rangle = \int_0^1 w^* w_{xx} \, dx + \int_0^1 w^* A^{-1} A' A^{-1} w \, dx. \tag{4.19}$$

Now because w satisfies the boundary conditions (4.15) and because of Assumption \mathcal{A} (that all the eigenvalues of (4.6) are positive), we know that the Rayleigh quotient

$$\frac{\int_0^1 w^* w_{xx} \, dx}{\langle w, w \rangle}$$

is negative. In particular its numerator is negative, so from (4.19) we have

$$\langle w, w_t \rangle \leq \int_0^1 w^* A^{-1} A' A^{-1} w \, dx. \quad (4.20)$$

Substituting (4.20) into the right-hand side of (4.17) yields

$$\begin{aligned} \frac{d}{dt} \left(\frac{1}{2} \langle w, w \rangle \right) &\leq \frac{1}{2} \int_0^1 w^* A^{-1} A' A^{-1} w \, dx \\ &= \frac{1}{2} \int_0^1 w^* A^{-1/2} (A^{-1/2} A' A^{-1/2}) A^{-1/2} w \, dx \\ &\leq \mu(A^{-1/2} A' A^{-1/2}) \int_0^1 (A^{-1/2} w)^* A^{-1/2} w \, dx \\ &= \mu(A^{-1/2} A' A^{-1/2}) \langle w, w \rangle, \end{aligned} \quad (4.21)$$

where $\mu(\cdot)$ denotes the logarithmic norm introduced in (4.1), (4.2). For each $T > 0$ let

$$\mu_T := \sup_{t \in [0, T]} \mu(A^{-1/2} A' A^{-1/2}). \quad (4.22)$$

Combining (4.22) with (4.21) and (4.16) we obtain the bound

$$\int_0^1 u_{xx}(x, t)^* A(t) u_{xx}(x, t) \, dx \leq \exp(2t\mu_T) \int_0^1 u_{xx}(x, 0)^* A(0) u_{xx}(x, 0) \, dx, \quad t \in [0, T], \quad T > 0. \quad (4.23)$$

We now substitute this back into (4.13) to obtain

$$\begin{aligned} \sum_{j=0}^{\infty} \lambda_j |c_j(t)|^2 &\leq e^{-\lambda_0(t-t_k)} \sum_{j=0}^{\infty} \lambda_j |c_j(t)|^2 \\ &\quad + E(k, A)^2 \exp(2t_{k+1}\mu_{t_{k+1}}) \|f''\|_{A(0)^{-1}}^2 \sum_{j=0}^{\infty} \frac{(1 - e^{-\lambda_j(t-t_k)})}{\lambda_j}, \end{aligned} \quad (4.24)$$

where we have used the fact that $u_{xx}(x, 0) = f''(x)$ together with the notation $\|\cdot\|_{A(0)^{-1}}$ to denote the norm defined, by analogy with (2.2), by

$$\|f\|_{A(0)^{-1}} := \int_0^1 f^* A(0) f \, dx. \quad (4.25)$$

From the eigenfunction expansion of v in (4.8) it is easy to verify that

$$\sum_{j=0}^{\infty} \lambda_j |c_j(t)|^2 = \int_0^1 v^*(x, t) (-v_{xx}(x, t)) \, dx. \quad (4.26)$$

We denote the quantity on the right-hand side of (4.26) by $|v|_1^2$ (the notation indicating that this is a quantity closely related to the usual H^1 seminorm). Then (4.24) may be written as

$$|v(\cdot, t_{k+1})|_1^2 \leq e^{-\lambda_0(t_{k+1}-t_k)} |v(\cdot, t_k)|_1^2 + E(k, A)^2 \exp(2t_{k+1}\mu_{t_{k+1}}) \|f''\|_{A(0)^{-1}}^2 \sum_{j=0}^{\infty} \frac{(1 - e^{-\lambda_j(t_{k+1}-t_k)})}{\lambda_j}. \tag{4.27}$$

This bound still requires some simplification; in particular we need to bound the term

$$F_k := \sum_{j=0}^{\infty} \lambda_j^{-1} (1 - e^{-\lambda_j(t_{k+1}-t_k)}). \tag{4.28}$$

We split the sum into two parts: from $j = 0$ to $4n - 1$ and from $j = 4n$ to infinity. Over the first part we simply use the inequality

$$\lambda_j^{-1} (1 - e^{-\lambda_j(t_{k+1}-t_k)}) \leq (t_{k+1} - t_k),$$

while in the second part we shall bound λ_j using

$$\lambda_j \geq \left(\frac{j}{n} - 2\right)^2 \pi^2 \alpha_k^2, \tag{4.29}$$

in which $\alpha_k^2 > 0$ is the least eigenvalue of A_k . This bound is an immediate corollary of Lemma 2.1. Together these two steps give

$$F_k \leq 4n(t_{k+1} - t_k) + \sum_{j=0}^{\infty} \left(\frac{j}{n} + 2\right)^{-2} \pi^{-2} \alpha_k^{-2} (1 - e^{-((j/n)+2)^2 \pi^2 \alpha_k^2 (t_{k+1}-t_k)}). \tag{4.30}$$

We now deal with the remaining terms using the *integral test* inequality, which states that for any monotone decreasing continuous function ψ ,

$$\int_0^{\infty} \psi(x) dx + \psi(0) \geq \sum_{j=0}^{\infty} \psi(j).$$

Applying this to the sum on the right-hand side of (4.30) results in the bound

$$F_k \leq 4n(t_{k+1} - t_k) + \left(\frac{1 - e^{-4\pi^2 \alpha_k^2 (t_{k+1}-t_k)}}{4\pi^2 \alpha_k^2}\right) + \int_0^{\infty} \left(\frac{x}{n} + 2\right)^{-2} \pi^{-2} \alpha_k^{-2} (1 - e^{-((x/n)+2)^2 \pi^2 \alpha_k^2 (t_{k+1}-t_k)}) dx. \tag{4.31}$$

Making the substitution $w = ((x/n) + 2)^2 \pi^2 \alpha_k^2 (t_{k+1} - t_k)$ allows this to be evaluated in terms of Γ functions, yielding

$$F_k \leq 4n(t_{k+1} - t_k) + \left(\frac{1 - e^{-4\pi^2 \alpha_k^2 (t_{k+1}-t_k)}}{4\pi^2 \alpha_k^2}\right) + \frac{\sqrt{t_{k+1} - t_k}}{\alpha_k \sqrt{\pi}}.$$

The inequality $1 - \exp(-x) \leq x$, valid for any $x \geq 0$, deals with the middle term, so we finally obtain the simple bound

$$F_k \leq (4n + 1)(t_{k+1} - t_k) + \frac{\sqrt{t_{k+1} - t_k}}{\alpha_k \sqrt{\pi}}. \quad (4.32)$$

Now suppose that $T > 0$ is fixed with $T = t_K$ for some $K > 0$. From (4.27), given an integer $K > 0$ and using the fact that $v(\cdot, 0) = 0$ we have

$$|v(\cdot, T)|_1^2 = \sum_{k=0}^{K-1} \{|v(\cdot, t_{k+1})|_1^2 - |v(\cdot, t_k)|_1^2\} \leq \left\{ \sum_{k=0}^{K-1} F_k E(k, A)^2 \right\} \exp(2T\mu_T) \|f''\|_{A(0)^{-1}}^2, \quad (4.33)$$

where we have used the assumption $\lambda_0 \geq 0$ and also the fact that $\exp(t_k \mu_{t_k}) \leq \exp(T\mu_T)$ for all $0 \leq k \leq K$. We can now prove the following result.

Theorem 4.2. *Let u be the solution of our original problem (4.3) with time-dependent coefficient $A(\cdot)$ and let \hat{u} be the solution of the approximate problem (4.4) with piecewise-constant coefficient \hat{A} . Assume that all the eigenvalues of (4.6) are positive for some k (and hence for every k). Let $T = t_K$ for some integer K . Then the error $v = u - \hat{u}$ satisfies the error bound*

$$|v(\cdot, T)|_1 \leq \left\{ \sum_{k=0}^{K-1} \left((4n + 1)(t_{k+1} - t_k) + \frac{\sqrt{t_{k+1} - t_k}}{\alpha_k \sqrt{\pi}} \right) E(k, A)^2 \right\}^{1/2} \exp(T\mu_T) \|f''\|_{A(0)^{-1}}, \quad (4.34)$$

where $E(k, A)$ is defined by (4.11), μ_T is defined by (4.22), $\alpha_k^2 > 0$ is the least eigenvalue of A_k , and $|\cdot|_1$ is defined by

$$|v|_1^2 = - \int_0^1 v^* v_{xx} dx > 0, \quad (4.35)$$

the positivity of the right-hand side of (4.35) being an immediate consequence of the assumption that all the eigenvalues of (4.6) are positive.

If a constant step size $t_{k+1} - t_k = \delta$ is to be used, then the bound given in Theorem 4.2 can be simplified somewhat, and expressed in a form which allows the value of δ to be determined *a priori* in terms of a predetermined accuracy ϵ .

Theorem 4.3. *Let $T > 0$ be given. Suppose that a constant step size δ is used, where $T = K\delta$ for some integer K : in other words, $t_{k+1} - t_k = \delta$ for $k = 0, 1, \dots, K$ in (4.4). Suppose also that the function $\hat{A}(t)$ is given by*

$$\hat{A}(t) = A\left(\frac{1}{2}(t_k + t_{k+1})\right), \quad t \in [t_k, t_{k+1}), \quad k = 0, \dots, K. \quad (4.36)$$

Let $\alpha_T > 0$ be given by

$$\alpha_T^2 = \inf_{t \in [0, T]} \{\text{least eigenvalue of } A(t)\}, \quad (4.37)$$

and let

$$\|A'\|_{L^\infty[0,T]} := \sup_{t \in [0,T]} \|A'(t)\|, \tag{4.38}$$

where A' denotes the derivative of A and the norm on the right-hand side of (4.38) is the usual matrix 2-norm. Let

$$\gamma_T := \frac{\|A'\|_{L^\infty[0,T]}}{\alpha_T}. \tag{4.39}$$

Suppose also that all the hypotheses of Theorem 4.2 hold. Then

$$|v(\cdot, T)|_1 \leq \frac{\sqrt{T}\delta^{3/4}}{2} \gamma_T \exp(T\mu_T) \|f''\|_{A(0)^{-1}} \left((4n+1)\sqrt{\delta} + \frac{1}{\alpha_T\sqrt{\pi}} \right)^{1/2}. \tag{4.40}$$

Proof. In the bound given by (4.34) we need to have an estimate for $E(k, A)$. Bounding the right-hand side of (4.11) using standard inequalities (including the identity $\|A^{-1/2}\|^2 = \|A^{-1}\|$, valid because A is Hermitian positive definite and $\|\cdot\|$ is the 2-norm) we have

$$\begin{aligned} E(k, A)^2 &\leq \sup_{t \in [t_k, t_{k+1}]} \|A(t)^{-1/2}\|^2 \|A_k^{-1}\| \|A(t) - A_k\|^2 \\ &\leq \sup_{t \in [t_k, t_{k+1}]} \|A(t)^{-1/2}\|^2 \sup_{t \in [t_k, t_{k+1}]} \|A(t)^{-1}\| \sup_{t \in [t_k, t_{k+1}]} \|A(t) - A_k\|^2 \\ &= \sup_{t \in [t_k, t_{k+1}]} \|A(t)^{-1}\|^2 \sup_{t \in [t_k, t_{k+1}]} \|A(t) - A_k\|^2 \\ &\leq \frac{1}{\alpha_T^2} \sup_{t \in [t_k, t_{k+1}]} \|A(t) - A_k\|^2 \\ &\leq \frac{(\delta/2)^2 \|A'\|_{L^\infty[0,T]}^2}{\alpha_T^2}. \end{aligned} \tag{4.41}$$

Substituting (4.41) into the right-hand side of (4.34), together with $t_{k+1} - t_k = \delta$ for all k , all the terms in the summation become independent of the summation index k . Using the fact that $K\delta = T$, this yields

$$|v(\cdot, T)|_1^2 \leq \frac{T\delta}{4\alpha_T^2} \|A'\|_{L^\infty[0,T]}^2 \left((4n+1)\delta + \frac{\sqrt{\delta}}{\alpha_T\sqrt{\pi}} \right) \exp(2T\mu_T) \|f''\|_{A(0)^{-1}}^2. \tag{4.42}$$

The factor

$$\frac{T\delta}{4\alpha_T^2} \|A'\|_{L^\infty[0,T]}^2$$

is precisely $\frac{1}{4}T\delta\gamma_T^2$. Substituting this into the right-hand side of (4.42) and taking square roots gives (4.40), as required. \square

Remark 4.4. It is also possible to bound μ_T in terms of γ_T . From (4.22) and (4.39),

$$\mu_T \leq \frac{\|A'\|_{L^\infty[0,T]}}{\alpha_T} = \gamma_T. \tag{4.43}$$

However it is important to note that μ_T may be negative, in which case (4.43) would result in the replacement of an exponentially small term in (4.40) by an exponentially large one, for large T .

Remark 4.5. We make some remarks on the practicality of Theorem 4.3. The term $\|f''\|_{A(0)^{-1}}$ may be assumed to be known, because the initial condition function f is known. If A is known symbolically, then it would be easy to compute A' symbolically and obtain a crude upper bound on $\|A'\|_{L^\infty[0,T]}$, perhaps even by using interval arithmetic and eigenvalue bounds based on Gershgorin's theorems [5, p. 341]. The most troublesome term is the quantity α_T defined by (4.37). If A is known symbolically and is small enough to be inverted symbolically, so that a symbolic expression for A^{-1} is available, then the expression

$$\alpha_T = \left(\sup_{t \in [0,T]} \|A^{-1}(t)\| \right)^{-1} =: \|A^{-1}\|_{L^\infty[0,T]}^{-1}$$

would allow α_T to be estimated in the same manner as $\|A'\|_{L^\infty[0,T]}$. Otherwise, the computation of α_T could be very difficult without further information.

Remark 4.6. The bound given by (4.40) is not tight. Given the formula (4.36), we would expect the difference between u and \hat{u} to be $O(\delta^2)$. This can indeed be proved, provided u is sufficiently smooth, although the resulting bound involves constants which cannot easily be determined *a priori*, unlike the bound in Theorem 4.2.

4.3. Propagation of error across time-steps

In each time-step $[t_k, t_{k+1})$ the solution \hat{u} must be found. Since \hat{A} is constant on $[t_k, t_{k+1})$, we would like to tackle this problem using the method of §3. However, our analysis of that method assumed exact initial conditions, and we do not have exact initial conditions: we have an approximation to $\hat{u}(\cdot, t_k)$ arising from the truncation of the eigenfunction expansions in the previous intervals $[0, t_1), \dots, [t_{k-1}, t_k)$. We need to account for the error due to inexact initial conditions and examine its propagation.

Lemma 4.7. *Let $|\cdot|_1$ be as above and let Assumption \mathcal{A} hold. Let w be any function satisfying the PDE $w_t = \hat{A}w$ together with the boundary conditions*

$$B_1 w(0, t) + B_2 w_x(0, t) = 0, \quad C_1 w(1, t) + C_2 w_x(1, t) = 0.$$

Then on each time-interval $[t_k, t_{k+1})$, the norm $|w(\cdot, t)|_1$ is monotone decreasing as a function of t .

Proof. This is a standard result and may be proved by differentiating the expression

$$|w|_1^2 = - \int_0^1 w^* w_{xx} \, dx$$

with respect to t . □

Consider the first time-interval $[0, t_1)$. Let \hat{u}_0 be the computed approximation to \hat{u} on this interval obtained by series truncation after $N_0 + 1$ terms, as in § 3. We can write this as

$$\hat{u}_0 = P_{N_0} \hat{u},$$

where P_{N_0} is the projection operator onto the space spanned by the first $N_0 + 1$ eigenfunctions of the eigenvalue problem (4.6) with $k = 0$. From the triangle inequality,

$$|u(\cdot, t_1) - \hat{u}_0(\cdot, t_1)|_1 \leq |u(\cdot, t_1) - \hat{u}(\cdot, t_1)|_1 + |\hat{u}(\cdot, t_1) - \hat{u}_0(\cdot, t_1)|_1.$$

On the second time-interval $[t_1, t_2)$ we introduce a function \hat{u}_1 , which is the approximation to \hat{u}_0 on this interval obtained by the projection

$$\hat{u}_1 = P_{N_1} \hat{u}_0,$$

where P_{N_1} is the projection operator onto the space spanned by the first $N_1 + 1$ eigenfunctions of the eigenvalue problem (4.6) with $k = 1$. (Note that expanded in terms of the eigenfunctions of (4.6) with $k = 0$, \hat{u}_0 has a finite expansion, but expanded in terms of the eigenfunctions of (4.6) with $k = 1$, \hat{u}_0 will generally have an infinite expansion, it is this expansion which we truncate to get \hat{u}_1 .) Following the same methods as before,

$$\begin{aligned} &|u(\cdot, t_2) - \hat{u}_1(\cdot, t_2)|_1 \\ &\leq |u(\cdot, t_2) - \hat{u}(\cdot, t_2)|_1 + |\hat{u}(\cdot, t_2) - \hat{u}_0(\cdot, t_2)|_1 + |\hat{u}_0(\cdot, t_2) - \hat{u}_1(\cdot, t_2)|_1 \\ &\leq |u(\cdot, t_2) - \hat{u}(\cdot, t_2)|_1 + |\hat{u}(\cdot, t_1) - \hat{u}_0(\cdot, t_1)|_1 + |\hat{u}_0(\cdot, t_2) - \hat{u}_1(\cdot, t_2)|_1 \\ &= |u(\cdot, t_2) - \hat{u}(\cdot, t_2)|_1 + |\hat{u}(\cdot, t_1) - P_{N_0} \hat{u}(\cdot, t_1)|_1 + |\hat{u}_0(\cdot, t_2) - P_{N_1} \hat{u}_0(\cdot, t_2)|_1, \end{aligned}$$

where we have used Lemma 4.7 on the middle term.

After k steps, we obtain in an analogous way,

$$\begin{aligned} |u(\cdot, t_k) - \hat{u}_{k-1}(\cdot, t_k)|_1 &\leq |u(\cdot, t_k) - \hat{u}(\cdot, t_k)|_1 + |\hat{u}(\cdot, t_1) - P_{N_0} \hat{u}(\cdot, t_1)|_1 \\ &\quad + \sum_{j=0}^{k-2} |\hat{u}_j(\cdot, t_{j+2}) - P_{N_{j+1}} \hat{u}_j(\cdot, t_{j+2})|_1. \end{aligned} \tag{4.44}$$

The first term on the right-hand side of (4.44) is the error term for which we obtained a bound in Theorem 4.2. The remaining error terms are all due to the projection operators P_{N_j} , which project a function onto the space spanned by the first $N_j + 1$ eigenfunctions of the problem (4.6) with $k = j$.

Given $\epsilon > 0$ and $T > 0$ the following algorithm allows one to compute an approximation to the exact solution u at time T with error at most ϵ in the $|\cdot|_1$ norm.

Algorithm 4.8.

- (1) Use the bound of Theorem 4.2 to choose a step size $\delta = T/K$ (K integer) such that $|u(\cdot, T) - \hat{u}(\cdot, T)|_1 \leq \epsilon/2$.
- (2) Choose N_0 such that $|\hat{u}(\cdot, t_1) - P_{N_0} \hat{u}(\cdot, t_1)|_1 \leq \epsilon/(2K)$.

(3) Starting from the initial condition $\hat{u}_0 = P_{N_0}f$, for j from 0 to $K - 2$,

- (i) compute $\hat{u}_j(\cdot, t_{j+1})$ from the initial condition $\hat{u}_j(\cdot, t_j) = P_{N_j}\hat{u}_{j-1}(\cdot, t_j)$,
- (ii) choose N_{j+1} such that

$$|\hat{u}_j(\cdot, t_{j+2}) - \hat{u}_{j+1}(\cdot, t_{j+2})| = |\hat{u}_j(\cdot, t_{j+2}) - P_{N_{j+1}}\hat{u}_j(\cdot, t_{j+2})|_1 \leq \epsilon/(2K).$$

In order to realize this algorithm in practice, we need to be able to choose N_{j+1} at each step so that $|\hat{u}_j(\cdot, t_{j+2}) - P_{N_{j+1}}\hat{u}_j(\cdot, t_{j+2})|_1 \leq \epsilon/(2K)$.

We may suppose that, for $t \in [t_{j+1}, t_{j+2})$, the function \hat{u}_j has an expansion

$$\hat{u}_j(x, t) = \sum_{\nu=0}^{\infty} c_{\nu, j+1} \mathbf{y}_{\nu, j+1}(x) \exp(-\lambda_{\nu, j+1}(t - t_{j+1})), \tag{4.45}$$

in which the coefficients $c_{\nu, j+1}$ are given by

$$\overline{c_{\nu, j+1}} = \int_0^1 (\hat{u}_j(x, t_{j+1}))^* A_{j+1}^{-1} \mathbf{y}_{\nu, j+1}(x) \, dx. \tag{4.46}$$

From (4.45) we have

$$(\hat{u}_j - P_{N_{j+1}}\hat{u}_j)(x, t) = \sum_{\nu=N_{j+1}+1}^{\infty} c_{\nu, j+1} \mathbf{y}_{\nu, j+1}(x) \exp(-\lambda_{\nu, j+1}(t - t_{j+1})),$$

and hence

$$|(\hat{u}_j - P_{N_{j+1}}\hat{u}_j)(\cdot, t)|_1^2 = \sum_{\nu=N_{j+1}+1}^{\infty} |c_{\nu, j+1}|^2 \lambda_{\nu, j+1} \exp(-2\lambda_{\nu, j+1}(t - t_{j+1})). \tag{4.47}$$

Using the fact that $-\mathbf{y}_{\nu, j+1}'' = \lambda_{\nu, j+1} A_{j+1}^{-1} \mathbf{y}_{\nu, j+1}$ together with the fact that both $\mathbf{y}_{\nu, j+1}$ and \hat{u}_j satisfy the boundary conditions, we obtain from (4.46) the expression

$$\begin{aligned} \overline{c_{\nu, j+1}} &= -\frac{1}{\lambda_{\nu, j+1}} \int_0^1 (\hat{u}_j''(x, t_{j+1}))^* \mathbf{y}_{\nu, j+1} \, dx \\ &= -\frac{1}{\lambda_{\nu, j+1}} \int_0^1 (A_{j+1}^{1/2} \hat{u}_j''(x, t_{j+1}))^* A_{j+1}^{-1/2} \mathbf{y}_{\nu, j+1} \, dx, \end{aligned}$$

whence Cauchy–Schwarz yields

$$\lambda_{\nu, j+1} |c_{\nu, j+1}|^2 \leq \frac{1}{\lambda_{\nu, j+1}} \int_0^1 (\hat{u}_j''(x, t_{j+1}))^* A_{j+1} (\hat{u}_j''(x, t_{j+1})) \, dx. \tag{4.48}$$

Substituting back into (4.47) yields

$$|(\hat{u}_j - P_{N_{j+1}}\hat{u}_j)(\cdot, t)|_1^2 \leq I_1 I_2, \tag{4.49}$$

where

$$I_1 = \sum_{\nu=N_{j+1}+1}^{\infty} \frac{1}{\lambda_{\nu,j+1}} \exp(-2\lambda_{\nu,j+1}(t-t_{j+1})) \tag{4.50}$$

and

$$I_2 = \int_0^1 (\hat{u}_j''(x, t_{j+1}))^* A_{j+1} (\hat{u}_j''(x, t_{j+1})) dx. \tag{4.51}$$

We now proceed to estimate the terms I_1 and I_2 separately. For the term I_1 we require a lower bound on the eigenvalues $\lambda_{\nu,j+1}$. The lower bound we use comes from (4.29) and is

$$\lambda_{\nu,j+1} \geq \left(\frac{\nu}{n} - 2\right)^2 \pi^2 \alpha_{j+1}^2, \quad \nu \geq 2n, \tag{4.52}$$

where α_{j+1}^2 is the least eigenvalue of A_{j+1} . This gives

$$\begin{aligned} I_1 &\leq \sum_{\nu=N_{j+1}+1}^{\infty} \left(\frac{\nu}{n} - 2\right)^{-2} \pi^{-2} \alpha_{j+1}^{-2} \exp\left(-2\left(\frac{\nu}{n} - 2\right)^2 \pi^2 \alpha_{j+1}^2 (t-t_{j+1})\right) \\ &\leq \frac{1}{\pi^2 \alpha_{j+1}^2} \int_{N_{j+1}}^{\infty} \left(\frac{x}{n} - 2\right)^{-2} \exp\left(-2\left(\frac{x}{n} - 2\right)^2 \pi^2 \alpha_{j+1}^2 (t-t_{j+1})\right) dx. \end{aligned}$$

where we have used the integral test to bound the sum. Making the substitution

$$v = 2\left(\frac{x}{n} - 2\right)^2 \pi^2 \alpha_{j+1}^2 (t-t_{j+1}),$$

the integral simplifies to yield

$$I_1 \leq \frac{\sqrt{t-t_{j+1}}}{\sqrt{2}\pi\alpha_{j+1}} \int_{2((N_{j+1}/n)-2)^2\pi^2\alpha_{j+1}^2(t-t_{j+1})}^{\infty} v^{-3/2} \exp(-v) dv.$$

This integral is an incomplete Γ -function; we estimate it by using the fact that over the range of integration, $v^{-3/2}$ is monotone decreasing, and can therefore be replaced by its value at the lower limit without invalidating the upper bound on I_1 . This yields

$$I_1 \leq \frac{\exp(-2((N_{j+1}/n) - 2)^2 \pi^2 \alpha_{j+1}^2 (t-t_{j+1}))}{4\pi^4 \alpha_{j+1}^4 ((N_{j+1}/n) - 2)^3 (t-t_{j+1})}. \tag{4.53}$$

Next we estimate I_2 . We start with the inequality

$$I_2 \leq \|A_j^{-1/2} A_{j+1} A_j^{-1/2}\|_2 \int_0^1 (\hat{u}_j''(x, t_{j+1}))^* A_j (\hat{u}_j''(x, t_{j+1})) dx, \tag{4.54}$$

in which $\|\cdot\|_2$ denotes the usual matrix 2-norm. Next we expand \hat{u}_j in terms of the eigenfunctions $\mathbf{y}_{\nu,j}$ for $t \in (t_j, t_{j+1}]$:

$$\hat{u}_j(x, t) = \sum_{\nu=0}^{N_j} c_{\nu,j} \exp(-\lambda_{\nu,j}(t-t_j)) \mathbf{y}_{\nu,j}(x), \quad (x, t) \in (0, 1) \times (t_j, t_{j+1}].$$

Using the orthogonality

$$\int_0^1 \mathbf{y}_{\nu,j}^* A_j^{-1} \mathbf{y}_{\mu,j} \, dx = \delta_{\nu,\mu}$$

together with the fact that $\mathbf{y}_{\nu,j}'' = -\lambda_{\nu,j} A_j^{-1} \mathbf{y}_{\nu,j}$ gives

$$\begin{aligned} \int_0^1 (\hat{u}_j''(x, t_{j+1}))^* A_j (\hat{u}_j''(x, t_{j+1})) \, dx &= \sum_{\nu=0}^{N_j} |c_{\nu,j}|^2 \exp(-2\lambda_{\nu,j}(t - t_j)) \lambda_{\nu,j}^2 \\ &\leq \sum_{\nu=0}^{N_j} |c_{\nu,j}|^2 \lambda_{\nu,j} \{ \lambda_{\nu,j} \exp(-2\lambda_{\nu,j}(t - t_j)) \} \\ &\leq \frac{1}{2(t - t_j)} \sum_{\nu=0}^{N_j} |c_{\nu,j}|^2 \lambda_{\nu,j} \\ &= \frac{1}{2(t - t_j)} |\hat{u}_j(\cdot, t_j)|_1^2. \end{aligned}$$

Combining this with (4.54) yields

$$I_2 \leq \frac{\|A_j^{-1/2} A_{j+1} A_j^{-1/2}\|_2}{2(t - t_j)} |\hat{u}_j(\cdot, t_j)|_1^2. \tag{4.55}$$

Together with the bound (4.53) for I_1 and the inequality $|(\hat{u}_j - P_{N_{j+1}} \hat{u}_j)(\cdot, t)|_1^2 \leq I_1 I_2$, this yields

$$\begin{aligned} |(\hat{u}_j - P_{N_{j+1}} \hat{u}_j)(\cdot, t_{j+2})|_1^2 &\leq \frac{\|A_j^{-1/2} A_{j+1} A_j^{-1/2}\|_2 |\hat{u}_j(\cdot, t_j)|_1^2}{8\pi^4 \alpha_{j+1}^4 ((N_{j+1}/n) - 2)^3 \delta^2} \exp\left(-2\left(\frac{N_{j+1}}{n} - 2\right)^2 \pi^2 \alpha_{j+1}^2 \delta\right), \tag{4.56} \end{aligned}$$

where we have used $t_{j+2} - t_{j+1} = t_{j+1} - t_j = \delta$. We can now choose the integer N_{j+1} to ensure the inequality

$$|(\hat{u}_j - P_{N_{j+1}} \hat{u}_j)(\cdot, t_{j+2})|_1 \leq \frac{\epsilon}{2K} = \frac{\epsilon \delta}{2T},$$

as required by Algorithm 4.8 above. Defining

$$M_{j+1}^2 := 2\left(\frac{N_{j+1}}{n} - 2\right)^2 \pi^2 \alpha_{j+1}^2 \delta, \tag{4.57}$$

from which N_{j+1} is easily found, some simple algebra shows that we require

$$M_{j+1}^{-3} \exp(-M_{j+1}^2) \leq \left\{ \frac{\pi \alpha_{j+1}}{2^{1/2} T^2 \|A_j^{-1/2} A_{j+1} A_j^{-1/2}\|_2 |\hat{u}_j(\cdot, t_j)|_1^2} \right\} \epsilon^2 \delta^{5/2}. \tag{4.58}$$

At each step the quantity $|\hat{u}_j(\cdot, t_j)|_1^2$ is known, since $\hat{u}_j(\cdot, t_j)$ is a computed quantity (computed in terms of a known set of eigenfunctions). The inequality (4.58) can be solved (e.g. numerically) to find an appropriate value for M_{j+1} and hence for N_{j+1} . We make three remarks on the size of M_{j+1} (and hence N_{j+1}) as follows.

- (1) As j increases, t_{j+2} increases and the solution of the PDE decays exponentially. Thus $|\hat{u}_j(\cdot, t_j)|_1^2$ may be expected to be exponentially small for large j , which is good news when choosing M_{j+1} .
- (2) At first sight it may appear that N_{j+1} increases only logarithmically with ϵ and δ . Unfortunately, while this is true for M_{j+1} it is not true for N_{j+1} . Since (4.57) involves δ and since δ may be expected to be $O(\epsilon^{4/3})$ from the error bound in Theorem 4.3, the dominant effect in the lower bound on N_{j+1} is $O(\epsilon^{-2/3})$. For large time, however, the previous remark means that that should not be a problem.
- (3) A further lower bound on N_{j+1} comes from the requirement $N_{j+1} > 2n$, which is explicit in our lower eigenvalue bound (4.52).

Acknowledgements. We thank the anonymous referee who read the first version of this paper so thoroughly and carefully.

References

1. M. H. ALEXANDER AND D. E. MANOPOULOS, A stable linear reference potential algorithm for solution of the quantum close-coupled equations in molecular scattering theory, *J. Chem. Phys.* **86** (1987), 2044–2050.
2. F. V. ATKINSON, *Discrete and continuous boundary value problems* (Academic, 1964).
3. J. CRANK, *The mathematics of diffusion* (Oxford University Press, 1995).
4. N. DUNFORD AND J. T. SCHWARTZ, *Linear operators, Part II, Spectral theory* (Wiley Interscience, 1963).
5. G. GOLUB AND C. VAN LOAN, *Matrix computations*, 2nd edn (Johns Hopkins University Press, 1989).
6. L. JÓDAR, E. NAVARRO AND J. CAMACHO, Analytic–numerical solutions with *a priori* error bounds for a class of strongly coupled mixed partial differential systems, *J. Comput. Appl. Math.* **104** (1999), 123–143.
7. J. A. MARTIN, E. NAVARRO AND L. JÓDAR, Exact and analytic numerical solutions of strongly coupled mixed diffusion problems, *Proc. Edinb. Math. Soc.* **43** (2000), 1–25.
8. V. S. MELEZHNIK, I. V. PUZYRNIN, T. P. PUZYRNINA AND L. N. SOMOV, Numerical solution of a system of integrodifferential equations arising the quantum-mechanical three-body problem with Coulomb interaction, *J. Comput. Phys.* **54** (1984), 221–236.
9. M. D. MIKHAILOV AND M. N. ÖSIZIK, *Unified analysis and solutions of heat and mass diffusion* (Wiley, 1984).
10. I. STAKGOLD, *Green's functions and boundary value problems* (Wiley, 1979).