BIT Numerical Mathematics (2006)46:000-000
DOI:10.1007/s10543-000-0000-x

RATIONAL APPROXIMATION TO TRIGONOMETRIC OPERATORS *

V. GRIMM AND M. HOCHBRUCK

Mathematisches Institut, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, D–40225 Düsseldorf, Germany. email: {grimm, marlis}@am.uni-duesseldorf.de

Abstract.

We consider the approximation of trigonometric operator functions that arise in the numerical solution of wave equations by trigonometric integrators. It is well known that Krylov subspace methods for matrix functions without exponential decay show superlinear convergence behavior if the number of steps is larger than the norm of the operator. Thus, Krylov approximations may fail to converge for unbounded operators. In this paper, we propose and analyze a rational Krylov subspace method which converges not only for finite element or finite difference approximations to differential operators but even for abstract, unbounded operators. In contrast to standard Krylov methods, the convergence will be independent of the norm of the operator and thus of its spatial discretization. We will discuss efficient implementations for finite element discretizations and illustrate our analysis with numerical experiments.

AMS subject classification (2000): 65F10, 65L60, 65M60, 65N22

Key words: Rational Krylov subspace methods, trigonometric operator function, Hilbert space, wave equations, trigonometric integrators, highly oscillatory problems, finite element discretization

1 Introduction

Recently, the numerical solution of large time-dependent hyperbolic problems, especially wave equations, has been an active area of research. Several new schemes have been proposed and analyzed. Short time error bounds have been shown in [9, 11, 12, 13, 14, 19, 28] while long time properties have been studied in [2, 3, 4, 15, 16, 17]. All these integrators require the evaluation or approximation of the product of a trigonometric operator function with a vector. For simplicity, we will restrict ourselves to functions arising in the solution of linear wave equations, namely we consider the approximation of

$$y(\tau) = f(\tau^2 A)v, \tag{1.1}$$

for

$$f(x) = \cos\sqrt{x} \qquad \text{or} \qquad f(x) = \operatorname{sinc}\sqrt{x}, \tag{1.2}$$

where $\operatorname{sinc} x = \sin x/x$. Here, A is a positive self-adjoined operator with compact resolvent or a matrix that results from a finite difference or finite element

*Received **. Revised **. Communicated by **.

discretization of such an operator and v is a function in a suitable Hilbert space. However, our analysis applies to other matrix functions as well, in particular to the functions arising in trigonometric integrators for nonlinear wave equations.

Recently, rational Krylov approximations for the approximation of functions related to the matrix exponential operator have been studied in a number of papers [1, 6, 7, 8, 20, 21, 22, 25, 31]. All these papers deal with the case that the differential operator is symmetric or sectorial, so that the functions decay if the eigenvalues become larger and larger. For such problems, the idea of using rational approximations arises naturally since this leads to fast convergence of small eigenvalues. For hyperbolic problems, the situation is less obvious, since trigonometric functions do not decay (or do not decay fast in the case of the sinc function) if the eigenvalues tend to infinity. Thus uniform approximations of highly oscillatory functions on the whole positive real line will not be possible for polynomials or rational functions. Theoretical results on the approximation of the cosine function by rational functions can be found in [5].

A crucial idea for our method is to take the origin of the problem into account, i.e. we exploit that weak solutions of the wave equation can be defined only if the initial data is spatially smooth. Under this assumption, we will prove that our algorithm allows to approximate the matrix cosine operator with a constant number of steps, where the constant is independent of the norm of the matrix. The main computational cost per step is the solution of a linear system, which can be done either by a fast direct solver [21] or by a preconditioned iterative solver [31]. For finite element discretizations, it turns out that the standard method requires almost the same amount of work as the rational version.

The paper is organized as follows: We motivate our method in Section 2 and recall some results from functional analysis. The main theorem and its proof are given in Section 3. In Section 4, we discuss the implementation of our new method and present some numerical experiments illustrating the efficiency and the theoretical results.

2 Motivation

Let A be a positive self-adjoint operator with compact resolvent on a Hilbert space H, so that A has an orthonormal basis of eigenvectors $e_i \in H$ with eigenvalues $\lambda_i > 0$, $i = 1, 2, \ldots$. Fractional powers A^α are defined on appropriate domains

$$\mathcal{D}(A^\alpha) := \left\{ v = \sum_{i=1}^{\infty} \mu_i e_i \in H \mid \sum_{i=1}^{\infty} |\lambda_i|^{2\alpha} |\mu_i|^2 < \infty \right\} = \{ v \mid \|A^\alpha v\| < \infty \}.$$

EXAMPLE 2.1. *For the negative Laplacian, $A = -\Delta$, with homogeneous Dirichlet boundary conditions on the unit square $\Omega = [0,1]^2$, and $H = L^2(\Omega)$ we have $\mathcal{D}(A^{\frac{1}{2}}) = H_0^1(\Omega)$ and $\mathcal{D}(A) = H_0^1(\Omega) \cap H^2(\Omega)$.*

For any function F bounded on $[0, \infty)$, one can define $F(A)$ on H via

$$F(A)\, v := \sum_{i=1}^{\infty} F(\lambda_i)\, \mu_i e_i, \qquad v = \sum_{i=1}^{\infty} \mu_i e_i \in H. \tag{2.1}$$

We consider trigonometric functions that appear in the solution of the linear wave equation

$$u'' = Au, \qquad u(0) = u_0, \quad u'(0) = u_0'. \tag{2.2}$$

For initial data $u_0 \in \mathcal{D}(A)$ and $u_0' \in V := \mathcal{D}(A^{\frac{1}{2}})$, (2.2) has a strong solution $u \in C^2(\mathbb{R}, H) \cap C^1(\mathbb{R}, V)$. For $u_0 \in V$ and $u_0' \in H$, (2.2) has a weak solution $u \in C^1(\mathbb{R}, H) \cap C(\mathbb{R}, V)$, that satisfies

$$\frac{d^2}{dt^2}(u(t), v) = a(u(t), v), \qquad \text{for all} \quad v \in V, \tag{2.3}$$

with the inner product $(\cdot, \cdot)$ in H and $a(u, v) = (A^{\frac{1}{2}}u, A^{\frac{1}{2}}v)$ for $u, v \in \mathcal{D}(A^{\frac{1}{2}})$. A strong solution is also a weak solution and any weak solution can be written as

$$u(t) = \cos(tA^{\frac{1}{2}})u_0 + t\,\mathrm{sinc}(tA^{\frac{1}{2}})u_0',$$

cf. [29]. The key observation here is that, despite of u being defined for all $u_0 \in H$ and $u_0' \in \mathcal{D}(A^{-\frac{1}{2}})$, u is a (weak) solution of (2.2) only for $u_0 \in \mathcal{D}(A^{\frac{1}{2}})$ and $u_0' \in H$. That is, all valid initial data u_0 occurring in applications is at least in $\mathcal{D}(A^{\frac{1}{2}})$, while for u_0' only $u_0' \in H$ is required. Note that this statement does not change when we are dealing with nonlinear wave equations. It is at the very heart of a weak solution to be in V.

For the approximation of $y(\tau) = f(\tau^2 A)v$ by Krylov methods, it is known from [18], that the approximation in the Krylov space

$$\mathcal{K}_m(A, v) = \mathrm{span}\,\{v, Av, A^2v, \cdots, A^{m-1}v\} \tag{2.4}$$

starts to converge superlinearly after $m \geq \|\tau A^{\frac{1}{2}}\|$ steps. Moreover, it requires the initial data to be in $\mathcal{D}(A^{m-1})$ which is much smoother than the natural requirements for the initial data. The key idea now is to use a variant of a rational Krylov method [26]. Following [21, 31] we approximate $y(\tau)$ in the rational Krylov space

$$\mathcal{K}_m((I + \gamma\tau^2 A)^{-1}, v), \qquad \gamma > 0.$$

The resolvent $(I + \gamma\tau^2 A)^{-1}$ maps a vector $v \in H$ into $\mathcal{D}(A)$, which implies that this space is well defined for all $v \in H$ and all m without additional smoothness assumptions. However, this approximation will not converge with a convergence rate independent of the norm of τA if the function to be approximated does not decay. Therefore, it is necessary to exploit the additional smoothness of the data appearing in wave equations. This can be done as follows. For $v \in \mathcal{D}(A^\alpha)$, we write

$$f(A)v = \psi_\alpha(A)A^\alpha v + f(0)v, \quad \psi_\alpha(x) = \frac{f(x) - f(0)}{x^\alpha} \tag{2.5}$$

and approximate ψ_α in the Krylov space $\mathcal{K}_m((I + \gamma\tau^2 A)^{-1}, A^\alpha v)$. In short, our new method is to proceed analogously to [31] for the function ψ_α and the starting vector $A^\alpha v$. For this method, a convergence rate will be shown.

3 Convergence of rational Lanczos approximations

The Lanczos algorithm based on the Krylov subspace

$$\mathcal{K}_m((I+\gamma\tau^2 A)^{-1}, A^\alpha v), \qquad 0 \le \alpha \le 1$$

yields an orthonormal basis V_m and the Lanczos recursion

$$(I+\gamma\tau^2 A)^{-1} V_m = V_m T_m + \beta_m v_{m+1} e_m^T, \tag{3.1}$$

where $\beta V_m e_1 = A^\alpha v$ and $V_m^* V_m = I$. Here, A could be either a matrix or an abstract operator. To improve the readability of the paper, we do not clearly distinguish between matrices and operators in our notation. Following [31], we use the relationship (2.5) between a given function f and ψ_α to define

$$g_\gamma^\alpha(x) := \psi_\alpha\left(\Big(\frac{1}{x}-1\Big)\frac{1}{\gamma}\right), \qquad x \in (0,1].$$

Note that

$$g_\gamma^\alpha((I+\gamma\tau^2 A)^{-1}) = \psi_\alpha(\tau^2 A). \tag{3.2}$$

This yields the approximation

$$y_m^\alpha(\tau) \approx y(\tau) = f(\tau^2 A)v$$

defined by

$$y_m^\alpha(\tau) = f(0)v + \tau^{2\alpha}\beta V_m \psi_\alpha\big(\widetilde{T}_m\big)e_1, \qquad \beta = \|A^\alpha v\| \tag{3.3}$$

where

$$\widetilde{T}_m = \frac{1}{\gamma}(T_m^{-1} - I). \tag{3.4}$$

We will now prove that the artificial decay introduced by exploiting the smoothness of the initial data enables us to derive a convergence rate. For simplicity, we state the theorem for the most common cases only. From the following discussion it becomes clear that similar theorems can be derived for initial data in different Hilbert spaces and for other functions f.

THEOREM 3.1. *For $f(x) = \cos\sqrt{x}$, the approximation (3.3) satisfies*

$$\|y_m^1(\tau) - \cos(\tau A^{\frac12})v\| \le \frac{C(\gamma)}{m^{\frac12}}\,\tau^2\,\|Av\|, \qquad \textit{for} \quad v \in D(A),$$
$$\|y_m^{\frac12}(\tau) - \cos(\tau A^{\frac12})v\| \le \frac{C(\gamma)}{m^{\frac14}}\,\tau\,\|A^{\frac12}v\|, \qquad \textit{for} \quad v \in D(A^{\frac12}),$$

while for $f(x) = \operatorname{sinc}\sqrt{x}$, the approximation satisfies

$$\|y_m^{\frac12}(\tau) - \operatorname{sinc}(\tau A^{\frac12})v\| \le \frac{C(\gamma)}{m^{\frac12}}\,\tau\,\|A^{\frac12}v\|, \qquad \textit{for} \quad v \in D(A^{\frac12}),$$
$$\|y_m^0(\tau) - \operatorname{sinc}(\tau A^{\frac12})v\| \le \frac{C(\gamma)}{m^{\frac14}}\,\|v\|, \qquad \textit{for} \quad v \in H.$$

The constants $C(\gamma)$ are independent of τ and of A.

Note that the error bounds are valid even for unbounded operators A. The convergence depends on the smoothness of the initial data only. For bounded operators resulting from finite difference or finite element approximations of an (unbounded) operator, these bounds show that the approximation is independent of the spatial discretization. The convergence is sublinear (cf. Definition 4.1.1 in [23]). In applications, where we aim for using just a small number of Krylov steps, sublinear convergence may actually be more favorable than linear convergence.

Before we prove this theorem, we need some auxiliary results. First we rewrite the Lanczos approximations (3.3) as

$$\begin{aligned}\beta V_m g_\gamma^\alpha(T_m)e_1 &= q_m\big((I+\gamma\tau^2 A)^{-1}\big)A^\alpha v\\ &= p_m(A)(I+\gamma\tau^2 A)^{-(m-1)}A^\alpha v,\end{aligned}$$

with polynomials $p_m, q_m \in \Pi_{m-1}$, where Π_{m-1} denotes the space of all polynomials of degree $m-1$ or less. As in [21, 31], the method can be alternatively characterized as constructing iterates from the class of *restricted* rational approximations defined by

$$R_i^j(\gamma) = \{\frac{p(x)}{(1+\gamma x)^i}\,|\,p\in\Pi_j\}.$$

LEMMA 3.2. *Let A be a self-adjoint positive operator and $0\le\alpha\le 1$. Then*

$$\|y_m^\alpha(\tau) - f(\tau^2A)v\| \le 2E_{m-1,\alpha}^{m-1}(\gamma)\,\tau^{2\alpha}\,\|A^\alpha v\|,$$

where

$$\begin{aligned}E_{m-1,\alpha}^{m-1}(\gamma) &:= \inf_{r\in R_{m-1}^{m-1}(\gamma)}\ \sup_{x\ge 0}\left|\frac{f(x)-f(0)}{x^\alpha} - r(x)\right|\\ &= \inf_{q\in\Pi_{m-1}}\ \sup_{x\in(0,1]}\left|g_\gamma^\alpha(x) - q(x)\right|.\end{aligned}$$

PROOF. According to [27] we have

$$q\big((I+\gamma\tau^2A)^{-1}\big)A^\alpha v = \beta V_m q(T_m)e_1 \qquad \text{for all } q\in\Pi_{m-1}. \tag{3.5}$$

From (2.5) and (3.2) we obtain

$$\frac{1}{\tau^{2\alpha}}\,\big(f(\tau^2A)v - y_m^\alpha(\tau)\big) = \psi_\alpha(\tau^2A)A^\alpha v - \beta V_m q(T_m)e_1 \tag{3.6}$$

$$+\,\beta V_m q(T_m)e_1 - \beta V_m\psi_\alpha(\widetilde{T}_m)e_1. \tag{3.7}$$

We first bound the right-hand side of (3.6). Using (3.5) we have

$$\left\|\psi_\alpha(\tau^2A)A^\alpha v - q((I+\gamma\tau^2A)^{-1})A^\alpha v\right\| \le \sup_{x\ge 0}\left|\psi_\alpha(x) - q\left(\frac{1}{1+\gamma x}\right)\right|\,\|A^\alpha v\|.$$

	$\alpha = 1$		$\alpha = 1/2$	
j	$E_j^j(\gamma_{\text{opt}})$	γ_{opt}	$E_j^j(\gamma_{\text{opt}})$	γ_{opt}
1	$4.5 \cdot 10^{-2}$	$1.64 \cdot 10^{-1}$	$3.6 \cdot 10^{-1}$	$2.00 \cdot 10^{0}$
5	$8.7 \cdot 10^{-3}$	$2.70 \cdot 10^{-2}$	$8.5 \cdot 10^{-2}$	$3.10 \cdot 10^{-2}$
10	$3.2 \cdot 10^{-3}$	$8.52 \cdot 10^{-3}$	$5.6 \cdot 10^{-2}$	$1.74 \cdot 10^{-2}$
15	$1.6 \cdot 10^{-3}$	$2.73 \cdot 10^{-3}$	$4.5 \cdot 10^{-2}$	$1.18 \cdot 10^{-2}$
20	$1.0 \cdot 10^{-3}$	$1.79 \cdot 10^{-3}$	$3.7 \cdot 10^{-2}$	$1.40 \cdot 10^{-2}$

Table 3.1: Numerical approximations to the optimal value of γ, γ_{opt}, and the corresponding value $E_j^j(\gamma_{\text{opt}})$ for $f(x) = \cos\sqrt{x}$

By (3.2) and (3.4), the difference in (3.7) can be bounded by

$$\begin{aligned}\left\|q(T_m) - g_\gamma^\alpha(T_m)\right\| \|A^\alpha v\| &\leq \sup_{\widetilde{x}\in(0,1]} \left|q(\widetilde{x}) - g_\gamma^\alpha(\widetilde{x})\right| \|A^\alpha v\| \\ &\leq \sup_{x\geq 0} \left| q\left(\frac{1}{1+\gamma x}\right) - \psi_\alpha(x)\right| \|A^\alpha v\|,\end{aligned}$$

by setting $\widetilde{x} = 1/(1+\gamma x)$. Since $q \in \Pi_{m-1}$ was an arbitrary polynomial, the statement of the lemma is proved. □

Proof [of Theorem 3.1]. According to Lemma 3.2 it remains to bound

$$E_{m-1,\alpha}^{m-1}(\gamma) = \inf_{q\in\Pi_{m-1}} \sup_{x\in(0,1]} \left|q(x) - g_\gamma^\alpha(x)\right| \tag{3.8}$$

for $\alpha = 1$ and $\alpha = \frac{1}{2}$, respectively. The error of this polynomial approximation problem can be bounded by Jackson's theorem (cf. [30]). We have

$$E_{m-1,\alpha}^{m-1}(\gamma) \leq C\omega\left(g_\alpha^\gamma; \frac{1}{m}\right),$$

where $\omega\left(g_\alpha^\gamma;\delta\right)$ is the modulus of continuity of g_α^γ on the interval $[0,1]$. For $f(x) = \cos\sqrt{x}$, a straightforward but tedious calculation shows that $\omega\left(g_\alpha^\gamma;\delta\right) \leq C(\gamma)\delta^{\frac{\alpha}{2}}$, for $\alpha \in \{\frac{1}{2}, 1\}$. For the $f(x) = \operatorname{sinc}\sqrt{x}$ we have $\omega\left(g_\alpha^\gamma;\delta\right) \leq C(\gamma)\delta^{\frac{\alpha}{2}+\frac{1}{4}}$, for $\alpha \in \{0, \frac{1}{2}\}$. □

For the practical application of this method one would like to choose γ such that the required accuracy is attained after a minimal number of steps. Lemma 3.2 shows that a priori error bounds for the proposed method can be given by choosing γ minimizing $E_{m-1,\alpha}^{m-1}(\gamma)$. We have computed numerical approximations to this optimal γ with the help of the Remez algorithm [24, Section 1.3] to solve the polynomial approximation problem (3.8). For the cosine function, the results for $\alpha = 1$ and $\alpha = \frac{1}{2}$ are summarized in Table 3.1. For the sinc function, the results for $\alpha = \frac{1}{2}$ and $\alpha = 0$ are summarized in Table 3.2.

	$\alpha = 1/2$		$\alpha = 0$	
j	$E_j^j(\gamma_{\rm opt})$	$\gamma_{\rm opt}$	$E_j^j(\gamma_{\rm opt})$	$\gamma_{\rm opt}$
1	$1.6 \cdot 10^{-1}$	$1.65 \cdot 10^{0}$	$2.1 \cdot 10^{-1}$	$5.00 \cdot 10^{-1}$
5	$2.0 \cdot 10^{-2}$	$3.13 \cdot 10^{-1}$	$8.6 \cdot 10^{-2}$	$1.49 \cdot 10^{-2}$
10	$7.8 \cdot 10^{-3}$	$2.15 \cdot 10^{-1}$	$5.2 \cdot 10^{-2}$	$6.58 \cdot 10^{-3}$
15	$5.2 \cdot 10^{-3}$	$1.04 \cdot 10^{-1}$	$3.8 \cdot 10^{-2}$	$3.51 \cdot 10^{-3}$
20	$3.8 \cdot 10^{-3}$	$1.02 \cdot 10^{-1}$	$3.0 \cdot 10^{-2}$	$2.20 \cdot 10^{-3}$

Table 3.2: Numerical approximations to the optimal value of γ, $\gamma_{\rm opt}$, and the corresponding value $E_j^j(\gamma_{\rm opt})$ for $f(x) = \operatorname{sinc}\sqrt{x}$.

4 Numerical experiments

4.1 Pseudospectral discretization

We start to illustrate the new error bounds by considering the following simple problem:

$$u'' = -Au, \qquad u(0) = u_0 \in V = H_0^1(0,1),$$
$$u'(0) = u_0' \in H = L^2(0,1),$$

with $A = -u_{xx}$,

$$u_0' = \begin{cases} x & \text{for} \quad 0 \le x \le \frac{1}{2} \\ 1-x & \text{for} \quad \frac{1}{2} < x \le 1 \end{cases}$$

and $u_0 = 0$. The solution is given by $u(\tau) = \tau \operatorname{sinc}(\tau A^{\frac{1}{2}}) u_0'$. The eigenpairs (φ_k, λ_k) of A are

$$\varphi_k = \sqrt{2}\sin(k\pi x), \qquad \lambda_k = (k\pi)^2, \qquad k = 1, 2, \cdots.$$

It is easy to check that $u_0' \in \mathcal{D}(A^{\frac{1}{2}}) = H_0^1(0,1)$, but $u_0' \notin \mathcal{D}(A)$. Galerkin discretization with respect to the subspace $V_N = \{\varphi_1, \cdots, \varphi_N\}$ and orthogonal projection onto the subspace lead to the following system of ordinary differential equations

$$u_N'' = -A_N u_N, \qquad u_N(0) = 0, \qquad u_N'(0) = u_{N,0}', \tag{4.1}$$

where

$$A_N = \operatorname{diag}\left(\pi^2, (2\pi)^2, \cdots, (N\pi)^2\right), \qquad u_{N,0}' = \left\{ \frac{2\sqrt{2}}{(k\pi)^2} \sin\left(\frac{k\pi}{2}\right) \right\}_{k=1,\cdots,N}.$$

With respect to the relation between A_N and A, one has

$$\|A_N^{\frac{1}{2}} u_{N,0}'\|_2 \le \|A^{\frac{1}{2}} u_0'\| = 1$$

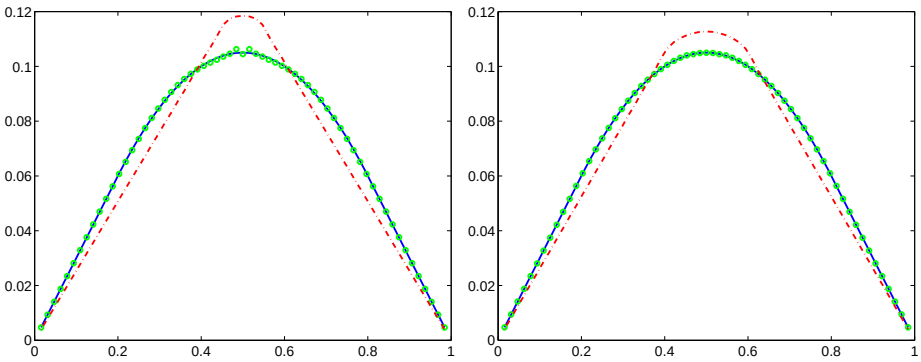

Figure 4.1: Plot of $u_N(\tau) = \tau\,\mathrm{sinc}(\tau A_N^{\frac{1}{2}})u'_{N,0}$ for $\tau = 0.3$ and $N = 63$ (solid) and of the approximations with the standard Krylov space (dash-dotted) and the rational Krylov space (circles) for 5 steps (left picture) and 10 steps (right picture).

independent of N. Here, $\|\cdot\|_2$ denotes the Euclidean norm in $\mathbb{C}^N$ and as before, $\|\cdot\|$ denotes the norm of the Hilbert space H (in this example $H = L^2(0,1)$). This is a particularly simple example of what has been discussed in Section 2. This property, possibly with a constant different from one, holds for higher dimensions, as well as for finite-difference (cf. [13]) and finite element discretizations when the projections are properly chosen.

In a first experiment, we compare the standard Krylov approximation to the solution $y_N(\tau) = u_N(\tau) = \tau\,\mathrm{sinc}(\tau A_N^{\frac{1}{2}})u'_{N,0}$ of (4.1) with the rational Krylov approximation for dimensions $N = 63$ and $N = 1023$. The parameter in the rational method has been chosen as $\gamma = 0.214$, which is the optimal value for 10 steps. However, the method is not sensitive with respect to choosing the optimal γ. The results for $\tau = 0.3$ are shown in Figures 4.1 and 4.2. In both figures, the solid line corresponds to the exact solution. The dash-dotted line is the approximation in the standard Krylov-space and the line marked with large circles corresponds to the rational Krylov approximation.

Figure 4.1 shows the approximations for $N = 63$ and after 5 and 10 Krylov steps, respectively. As expected, the rational method first finds the low modes of the solution, whereas the standard Krylov method takes care of the high frequencies first. This can be seen in the softening of the kink in the first few iterations. With a finer spatial grid, more and more high frequencies come in and slow down the convergence of the standard Krylov method, see Figure 4.2 for $N = 1023$.

These simple experiments illustrate the result of our main theorem, which shows that the error bounds are independent of the norm of the matrix. However, they are obviously effected by the smoothness of the data.

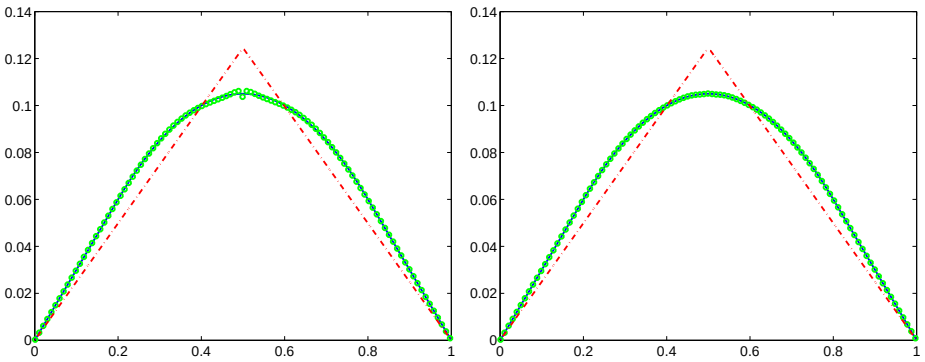

Figure 4.2: Plot of $u_N(\tau) = \tau\,\mathrm{sinc}(\tau A_N^{\frac{1}{2}})u'_{N,0}$ for $\tau = 0.3$ and $N = 1024$ (solid) and of the approximations with the standard Krylov space (dash-dotted) and the rational Krylov space (circles) for 5 steps (left picture) and 10 steps (right picture).

4.2 Finite element discretization

Applying the Galerkin method with linear finite elements to the wave equation of Example 2.1 in the weak form (2.3) gives

$$\frac{d^2}{dt^2}(u_h(t), v_h) = a(u_h(t), v_h), \qquad \text{for all} \quad v_h \in V_h, \tag{4.2}$$

where V_h is the finite element space. For the nodal basis $\{\varphi_1, \varphi_2, \ldots, \varphi_N\}$, the finite element approximation can be written as

$$u_h(t) = \sum_{i=1}^{N} \mu_i(t)\varphi_i, \qquad \mu_h = (\mu_i)_{i=1}^N. \tag{4.3}$$

The initial values are chosen such that

$$\begin{aligned} a(u_h(0), v_h) &= a(u(0), v_h), && \forall v_h \in V_h, \\ (u'_h(0), v_h) &= (u'(0), v_h), && \forall v_h \in V_h. \end{aligned}$$

Inserting (4.3) into (4.2) gives the following system of ordinary differential equations for the coefficients $\mu_h(t)$

$$M_h\mu''_h(t) + A_h\mu_h(t) = 0, \qquad \mu_h(0) = \mu_{h,0}, \qquad \mu'_h(0) = \mu'_{h,0},$$

with the mass matrix M_h and the stiffness matrix A_h, which are both symmetric positive definite. $\mu_{h,0}$ and $\mu'_{h,0}$ are the coefficients of $u_h(0)$ and $u'_h(0)$, respectively. The norm in the Hilbert space L^2 is given by the M_h-norm of the coefficients, that is $\|u_h\| = \|\mu_h\|_{M_h} := \|M_h^{\frac{1}{2}}\mu_h\|_2$.

In order to apply our convergence result, we consider the transformation

$$\widetilde{\mu}_h(t) = M_h^{\frac{1}{2}}\mu_h(t), \qquad \widetilde{A}_h = M_h^{-\frac{1}{2}}A_hM_h^{-\frac{1}{2}} \tag{4.4}$$

which leads to

$$\widetilde{\mu}_h''(t) + \widetilde{A}_h \widetilde{\mu}_h(t) = 0, \tag{4.5}$$

with initial values

$$\widetilde{\mu}_h(0) = M_h^{\frac{1}{2}} \mu_{h,0} =: \widetilde{\mu}_{h,0}, \qquad \widetilde{\mu}_h'(0) = M_h^{\frac{1}{2}} \mu_{h,0}' =: \widetilde{\mu}_{h,0}'.$$

One could interpret the transformation as choosing an orthonormal basis in the finite-dimensional space. We will now approximate the solution of (4.5) for $\widetilde{\mu}_{h,0}' = 0$ by directly approximating $y_h(\tau) = \mu_h(\tau) = \cos(\tau \widetilde{A}_h^{\frac{1}{2}}) \widetilde{\mu}_{h,0}$ with the new rational Krylov method. The approximation to the coefficient vector $\mu_h(\tau)$ after m Krylov steps is denoted by $\mu_{h,m}(\tau)$ which corresponds to the function $u_{h,m}(\tau)$.

The following two propositions are easy to prove.

PROPOSITION 4.1. *Let μ_h be the vector consisting of the coefficients of the Ritz projection* (4.3) *of u. Then the finite element discretization satisfies*

$$\|M_h^{-\frac{1}{2}} A_h \mu_h\|_2 = \|\widetilde{A}_h \widetilde{\mu}_h\|_2 \leq \|Au\|, \qquad \forall u \in \mathcal{D}(A),$$

where $\widetilde{\mu}_h$ and $\widetilde{A}_h$ are defined in (4.4).

With this proposition and Theorem 3.1, the more important following proposition can be shown.

PROPOSITION 4.2. *For $u_0 \in \mathcal{D}(A)$ let $u_h(\tau)$ be the finite element approximation of $u(\tau) = \cos(\tau A) u_0$ and let $u_{h,m}(\tau)$ be the rational Krylov approximation of $u_h(\tau)$. Then the error is bounded by*

$$\|u_h(\tau) - u_{h,m}(\tau)\| = \|\mu_h(\tau) - \mu_{h,m}(\tau)\|_{M_h} \leq \frac{\tau}{\sqrt{m}} C \|Au_0\|.$$

Note that the right-hand side is independent of the dimension of the finite element subspace.

In a similar way as for the preconditioned conjugate gradient method (cf. [10]), one can compute approximations to $\mu_h(t)$ without explicitly conducting the transformation. This yields the following algorithm to compute the approximation $\mu_{h,m}(\tau)$ to $\mu_h(\tau)$

$v = M_h^{-1} A_h \mu_{h,0},\ w = M_h v,\ \beta = \sqrt{v^H w} > 0;$
$v_1 = v/\beta,\ w_1 = w/\beta$
for $m = 1, 2, \ldots$ **do**
  $t_{m-1,m} = \overline{t_{m,m-1}}$
  $z_m = (M_h + \gamma \tau^2 A_h)^{-1} w_m$
  $t_{m,m} = w_m^H z_m$
  $\widetilde{v}_{m+1} = z_m - t_{m,m} v_m - t_{m-1,m} v_{m-1}$
  $\widetilde{w}_{m+1} = M_h \widetilde{v}_{m+1}$
  $t_{m+1,m} = \sqrt{\widetilde{v}_{m+1}^H \widetilde{w}_{m+1}}$
  $v_{m+1} = \widetilde{v}_{m+1}/t_{m+1,m},\ w_{m+1} = \widetilde{w}_{m+1}/t_{m+1,m}$
end for

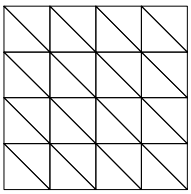

Figure 4.3: Coarsest mesh used in first experiment

number of grid points	9	961	3969	16129
error	$1.9 \cdot 10^{-9}$	$1.5 \cdot 10^{-8}$	$1.3 \cdot 10^{-8}$	$1.3 \cdot 10^{-8}$

Table 4.1: Exact error after 10 rational Krylov steps for $\tau = 0.3$ for the second experiment

The algorithm requires the solution of one linear system with the coefficient matrix $M_h + \gamma\tau^2 A_h$ per step and one solve with M_h in the initial step. The algorithm corresponds to the Lanczos iteration

$$(M_h + \gamma\tau^2 A_h)^{-1} M_h V_m = V_m T_m + \beta_m v_{m+1} e_m^T,$$

where $V_m = [v_1 \ \dots \ v_m]$ and $T_m = (t_{i,j})_{i,j=1}^m$ is a tridiagonal matrix. In finite element approximations, the solution of such linear systems can be computed efficiently by the preconditioned conjugate gradient method. The approximation to $y_h(\tau) = \mu_h(\tau) = M_h^{-\frac{1}{2}} \cos(\tau \widetilde{A}_h^{\frac{1}{2}}) \widetilde{\mu}_{h,0}$ for $\alpha = 1$ is then given by

$$y_{h,m}^1(\tau) = \mu_{h,m}(\tau) = \mu_{h,0} + \tau^2 \, V_m g_\gamma^1(T_m) e_1.$$

To illustrate our theoretical convergence results, we consider (4.2) in the two-dimensional unit square for the Laplace operator with homogeneous Dirichlet boundary conditions and initial values chosen as

$$u_0 \, : \, (x, y) \mapsto x(1-x)y(1-y), \qquad u_0' = 0.$$

The initial value u_0 is in $\mathcal{D}(A)$ but not in $\mathcal{D}(A^2)$. We use a sequence of refined regular triangulations. The coarsest grid is shown in Figure 4.3. In Table 4.2, we present the error of $\|u_h(0.3) - u_{h,10}(0.3)\|$ obtained by the rational Krylov method with γ being optimal for 10 steps. As predicted by Proposition 4.2, the error is independent of the spatial discretization.

In a second experiment, we use the domain and the mesh with 546 nodes, shown in Figure 4.4, as well as a coarser mesh with 43 nodes and a finer mesh with 5723 nodes. We use the same initial data as before on the inner square of the domain and zero elsewhere. Then we compare the rational Krylov method with the standard Krylov method, which can be applied to (4.5) in an analogous way (i.e. without using $M_h^{\frac{1}{2}}$ explicitly). The standard Krylov method requires the solution of a linear system with M_h in each step whereas the rational Krylov

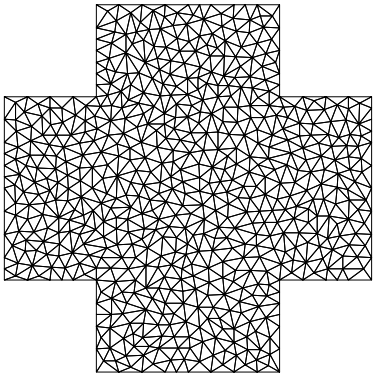

Figure 4.4: Mesh and domain for the second experiment.

method requires the solution of a linear system with $M_h + \gamma\tau^2 A_h$. We used a direct solver for both methods, so that both methods required roughly the same computational work per step. The error $\|u_h(0.3) - u_{h,m}(0.3)\|$ for both methods and for the three different meshes is presented in Figure 4.5.

Acknowledgement

This work has been supported by the Deutsche Forschungsgemeinschaft via the Transregio SFB 18.

REFERENCES

1. L. Bergamaschi and M. Vianello. Efficient computation of the exponential operator for large, sparse, symmetric matrices. *Numer. Linear Algebra Appl.*, 7:27–45, 2000.
2. D. Cohen, E. Hairer, and Ch. Lubich. Long-time analysis of nonlinearly perturbed wave equations via modulated fourier expansions. To appear in Arch. Ration. Mech. Anal., 2006.
3. D. Cohen, E. Hairer, and Ch. Lubich. Conservation of energy, momentum and actions in numerical discretizations of nonlinear wave equations. Preprint, April, 2007.
4. D. Cohen, E. Hairer, and Ch. Lubich. Energy conservation over long times of numerical discretizations for nonlinear wave equations. In *Proceedings of Equadiff07*, 2008. To appear.
5. V.A. Dougalis and S.M. Serbin. Remarks on a class of rational approximations to the cosine. *BIT*, 20:204–211, 1980.
6. V. L. Druskin and L. A. Knizhnerman. Extended Krylov subspaces: approximation of the matrix square root and related functions. *SIAM J. Matrix Analysis and Applications*, 19:755–771, 1998.

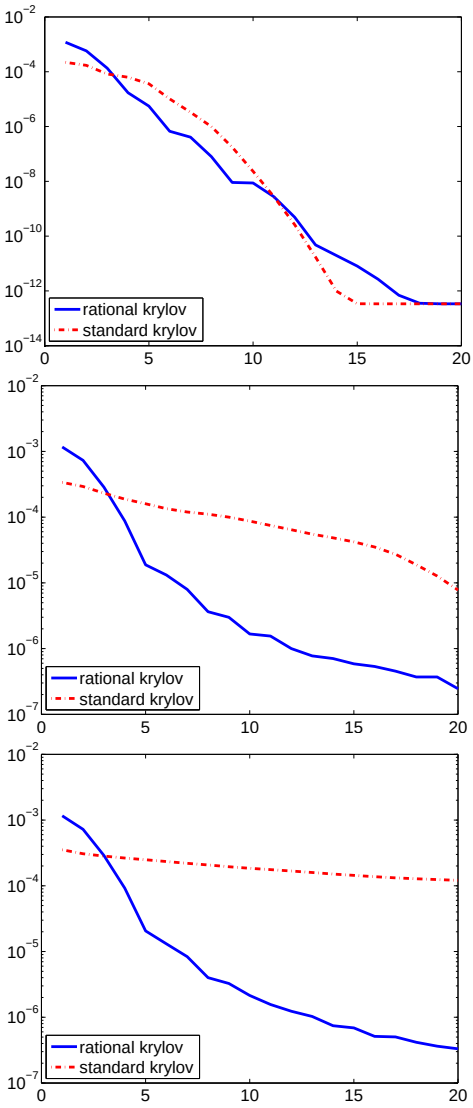

Figure 4.5: Comparison of the rational Krylov method and the standard Krylov method by the error $\|u_h(0.3) - u_{h,m}(0.3)\|$ for the grids with 43, 546 and 5723 nodes (top to bottom) and steps $m = 1, \ldots, 20$.

7. M. Eiermann and O. Ernst. A restarted Krylov subspace method for the evaluation of matrix functions. To appear in SIAM J. Numerical Analysis, 2005.
8. A. Frommer and V. Simoncini. Stopping criteria for rational matrix functions of Hermitian and symmetric matrices. To appear in SIAM J. Scientific Computing, 2007.
9. B. García-Archilla, J. Sanz-Serna, and R. Skeel. Long-time-step methods for oscillatory differential equations. *SIAM J. Sci. Comput.*, 30(3):930–963, 1998.
10. G. H. Golub and C. F. van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, second edition, 1989.
11. V. Grimm. A note on the Gautschi-type method for oscillatory second-order differential equations. *Numer. Math.*, 102:61–66, 2005.
12. V. Grimm. On error bounds for the Gautschi-type exponential integrator applied to oscillatory second-order differential equations. *Numer. Math.*, 100:71–89, 2005.
13. V. Grimm. On the Use of the Gautschi-Type Exponential Integrator for Wave Equations. In A. Bermúdez de Castro, D. Gómez, P. Quintela, and P. Salgado, editors, *Numerical Mathematics and Advanced Applications, ENUMATH 2005*, pages 557–563. Springer-Verlag, Berlin Heidelberg, 2006.
14. V. Grimm and M. Hochbruck. Error analysis of exponential integrators for oscillatory second-order differential equations. *J. Phys. A; Math. Gen.*, 39:5495–5507, 2006.
15. E. Hairer and C. Lubich. Long-time energy conservation of numerical methods for oscillatory differential equations. *SIAM J. Numer. Anal.*, 38:414–441, 2000.
16. E. Hairer and Ch. Lubich. Spectral semi-discretisations of weakly nonlinear wave equations over long times. To appear in Foundations of Comput. Math, 2007.
17. E. Hairer and Ch. Lubich. Oscillations over long times in numerical Hamiltonian systems. In *Proceedings of the HOP programme, Isaac Newton Institute*, 2008. To appear.
18. M. Hochbruck and Ch. Lubich. On Krylov subspace approximations to the matrix exponential operator. *SIAM J. Numer. Anal.*, 34:1911–1925, 1997.
19. M. Hochbruck and Ch. Lubich. A Gautschi-type method for oscillatory second-order differential equations. *Numer. Math.*, 83:403–426, 1999.
20. L. Lopez and V. Simoncini. Analysis of projection methods for rational function approximation to the matrix exponential. *SIAM J. Numerical Analysis*, 44(2):613–635, 2006.
21. I. Moret and P. Novati. RD-rational approximations of the matrix exponential. *BIT*, 44:595–615, 2004.

22. I. Moret and P. Novati. Interpolating functions of matrices on zeros of quasi-kernel polynomials. *Numer. Linear Algebra Appl.*, 11:337–353, 2005.

23. O. Nevanlinna. *Convergence of Iterations for Linear Equations.* Birkhäuser, Basel, 1993.

24. P. P. Petrushev and V. A. Popov. *Rational approximation of real functions.* Cambridge University Press, Cambridge, 1987.

25. M. Popolizio and V. Simoncini. Accerleration Techniques for Approximating the Matrix Exponential Operator. Preprint, October, 2006.

26. A. Ruhe. Rational Krylov sequence methods for eigenvalue computation. *Linear Alg. Appl.*, 58:391–405, 1984.

27. Y. Saad. Analysis of some Krylov subspace approximations to the matrix exponential operator. *SIAM J. Numer. Anal.*, 29:209–228, 1992.

28. J. M. Sanz-Serna. Mollified impulse methods for highly-oscillatory differential equations. To appear in SIAM J. Numer. Anal., 2007.

29. G.R. Sell and Y. You. *Dynamics of Evolutionary Equations.* Springer-Verlag, New Youk Berlin Heidelberg, 2002.

30. A. F. Timan. *Theory of Approximation of Functions of a Real Variable.* Pergamon Press, Oxford, 1963.

31. J. van den Eshof and M. Hochbruck. Preconditioning Lanczos approximations to the matrix exponential. *SIAM J. Sci. Comp.*, 24(4):1438–1457, 2006.