Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Numerical computation of inner eigenvalues using the Dunford Cauchy integral

by

Wendy Kress

Preprint no.: 52 2007

Numerical computation of inner eigenvalues using the Dunford Cauchy integral

Wendy Kress

June 1, 2007

Abstract

When computing the eigenvalues of a matrix using iterative Krylov subspace methods, convergence is usually best for the extreme eigenvalues. We present a projection technique that enables us to efficiently compute the eigenvalues that lie in a specified interval in the interior of the spectrum. To obtain such a projection, the Dunford Cauchy integral is used. The technique requires fast inversion algorithms which are available for some classes of matrices like $\mathcal{H}$-matrices.

1 Introduction

Most methods for computing eigenvalues of a matrix concentrate on finding either the complete spectrum of the matrix or specific eigenvalues – in most cases the extreme eigenvalues. One can, e.g., use Krylov subspace methods. Given a matrix $\mathbf{A}$, they find approximations to $\mathbf{A}$'s eigenvectors from a subspace $\mathcal{K}_k(\mathbf{A}, u_0) := span\left\{\mathbf{A}u_0, \ldots, \mathbf{A}^k u_0\right\}$ for some vector u_0. In addition, approximate eigenvalues, so called Ritz values, are found. As k is increased, the Ritz values converge to the eigenvalues of $\mathbf{A}$. Convergence is fastest for those eigenvalues belonging to the upper and lower spectrum.

When only interested in eigenvalues in the inner region of the spectrum, the basic Krylov subspace method is not efficient. We can modify the method in the following way. To guarantee that the Krylov subspace method arrives only at the interesting eigenvalues, we change the search space from $\mathcal{K}_k(\mathbf{A}, u_0)$ to $\mathcal{K}_k(\mathbf{A}_p, u_0)$, where $\mathbf{A}_p$ is a projection of $\mathbf{A}$ onto the space spanned by the eigenvectors corresponding to the interesting eigenvalues. The Krylov method will then find only the nonzero eigenvalues and corresponding eigenvectors of $\mathbf{A}_p$. In Section 3, we will describe the Krylov method in a little more detail. For an overview on Krylov methods for the numerical computation of eigenvalues see [9]. For other methods for the computation of eigenvalues, e.g., the QR algorithm, see [4] or [8] which also covers Krylov methods.

In this work, we present a technique to efficiently compute the projection $\mathbf{A}_p$ using the Dunford Cauchy integral. For this formulation, we will require a cheap inversion algorithm for matrices $(z\mathbf{I} - \mathbf{A})$ with $z \in \mathbb{C}$, which is available for matrices permitting a representation by $\mathcal{H}$-matrices [6], [3], in almost linear complexity.

2 Prerequisites and notations

In the following, we consider $\mathbf{A} \in \mathbb{R}^{M \times M}$ with real eigenvalues $\lambda_1 \leq \ldots \leq \lambda_M$ and we assume that $\mathbf{A}$ possesses an orthonormal set of eigenvectors $\mathbf{V} = [v_1, \ldots, v_M]$. Consequently, $\mathbf{V}^T\mathbf{A}\mathbf{V}$ is diagonal and consists of $\mathbf{A}$'s eigenvalues. The above assumptions are fulfilled for real symmetric matrices. Although, in this work, we consider only matrices of the above kind, the procedure will also work for other diagonalizable matrices, the analysis being slightly more involved.

Given a specified interval, we denote the eigenvalues of $\mathbf{A}$ lying in that interval by $\lambda_{i+1}, \ldots, \lambda_{i+n}$ and the corresponding eigenvectors by $v_{i+1}, \ldots, v_{i+n}$. We assume that no eigenvalue is exactly equal to the endpoints of the interval.

Using scaling and translation, we assume the specified interval to be $[-1, 1]$.

3 Krylov subspace methods and projection $\mathbf{A}_p$

In a Krylov subspace method, we consider the subspace $\mathcal{K}_k(\mathbf{A}, u_0) := span\left\{\mathbf{A}u_0, \ldots, \mathbf{A}^k u_0\right\}$ for some initial vector u_0 and small k and seek approximate eigenvectors $x \in \mathcal{K}_k(\mathbf{A}, u_0)$ fulfilling

$$(v, \mathbf{A}x - \theta\, x) = 0\,, \quad \forall v \in \mathcal{K}_k(\mathbf{A}, u_0)\,, \tag{1}$$

for some value $\theta \in \mathbb{R}$. Here, $(\cdot,\cdot)$ denotes the usual scalar product in $\mathbb{R}^M$. Such vectors x and values θ are called Ritz vectors and Ritz values, respectively. If U_k is an orthonormal basis of $\mathcal{K}_k(\mathbf{A}, u_0)$, the Ritz values are the eigenvalues of the matrix $S_k = U_k^T \mathbf{A} U_k$, and the Ritz vectors are the eigenvectors of S_k multiplied by U_k. To see this, let θ be an eigenvalue of S_k with corresponding eigenvector y. Then taking $x = U_k y$, $U_k^T(\mathbf{A}x - \theta x) = 0$ which is just another formulation of (1). On the other hand, if x and θ fulfill (1), then for all $w \in \mathbf{R}^M$, $(U_k w, U_k S_k U_k^T x - \theta x) = 0$ and shifting U_k to the second argument of the scalar product leads to $S_k y = \theta y$ with $y = U_k^T x$.

Since S_k is a small matrix, its eigenvalues and eigenvectors can be computed at low cost using, e.g., a QR algorithm.

As mentioned above, a projection $\mathbf{A}_p$ of $\mathbf{A}$ onto $span\{v_{i+1}, \ldots, v_{i+n}\}$ can be used in a modification of a Krylov subspace method. We consider a Krylov subspace method using the subspace $\mathcal{K}_k(\mathbf{A}_p, u_0)$ instead of $\mathcal{K}_k(\mathbf{A}, u_0)$: Find Ritz vectors $x \in \mathcal{K}_k(\mathbf{A}_p, u_0)$ and Ritz values θ such that

$$(v, \mathbf{A}x - \theta\, x) = 0\,, \quad \forall v \in \mathcal{K}_k(\mathbf{A}_p, u_0)\,. \tag{2}$$

The approximate eigenvalues obtained in the Krylov method can only correspond to eigenvalues $\lambda_{i+1}, \ldots, \lambda_{i+n}$, since the Krylov subspace is a subspace of the space spanned by the eigenvectors $v_{i+1}, \ldots, v_{i+n}$.

4 Dunford Cauchy integral

One way to compute the projection of a matrix is to use the *Dunford Cauchy integral*. We recall the Cauchy integral formula: Given a simply closed curve Γ in the complex plane, we have for $x \notin \Gamma$

$$x_p := \frac{1}{2\pi i} \oint_\Gamma \frac{z}{z-x} dz = \begin{cases} x & \text{if } x \text{ is enclosed by } \Gamma\,, \\ 0 & \text{if } x \text{ is not enclosed by } \Gamma\,. \end{cases}$$

The Cauchy integral can be extended to matrices in the following way,

$$\mathbf{A}_p := \frac{1}{2\pi i} \oint_\Gamma z\,(z\mathbf{I} - \mathbf{A})^{-1} dz\,. \tag{3}$$

The right hand side is called the *Dunford Cauchy integral* [1]. If we assume that Γ encloses only the eigenvalues $\lambda_{i+1}, \ldots, \lambda_{i+n}$, $\mathbf{A}_p$ is a projection of $\mathbf{A}$ onto the subspace $v_{i+1}, \ldots, v_{i+n}$. To see this, we note that

$$\mathbf{A} = \mathbf{V} diag(\lambda_1, \ldots, \lambda_M) \mathbf{V}^T\,.$$

Then

$$\begin{aligned} \mathbf{A}_p &= \mathbf{V} \left(\frac{1}{2\pi i} \oint_\Gamma z (diag(z-\lambda_1, \ldots, z-\lambda_M))^{-1} dz \right) \mathbf{V}^T \\ &= \mathbf{V} diag \left(\frac{1}{2\pi i} \oint_\Gamma \frac{z}{z-\lambda_1} dz, \ldots, \frac{1}{2\pi i} \oint_\Gamma \frac{z}{z-\lambda_M} dz \right) \mathbf{V}^T \\ &= \mathbf{V}\, diag\,(0, \ldots, 0, \lambda_i, \lambda_{i+1}, \ldots, \lambda_{i+n}, 0, \ldots, 0)\, \mathbf{V}^T\,, \end{aligned}$$

which is the projection onto the desired subspace.

To be more precise about the conditions under which the above integral is well defined, let $\gamma : [0, 2\pi] \to \mathbb{C}$ be a smooth parametrisation of the curve Γ.

Remark 1 *The Dunford Cauchy integral (3) exists as a proper integral under the assumption that $\gamma(\omega)$ does not coincide with an eigenvalue of $\mathbf{A}$.*

To numerically compute $\mathbf{A}_p$, we employ the trapezoidal rule to obtain an approximation $\tilde{\mathbf{A}}_p$,

$$\tilde{\mathbf{A}}_p = \frac{1}{2\pi i}\frac{2\pi}{N}\sum_{k=0}^{N-1}\gamma\left(\frac{2\pi k}{N}\right)\gamma'\left(\frac{2\pi k}{N}\right)\left(\gamma\left(\frac{2\pi k}{N}\right)\mathbf{I}-\mathbf{A}\right)^{-1}, \tag{4}$$

which for periodic functions has the following exponential convergence properties,

Theorem 2 *If $(\gamma(z)\mathbf{I}-\mathbf{A})$ is nonsingular in a strip $D_d := \{z\in\mathbb{C} : |\mathrm{Im}(z)|\le d\}$, we have*

$$\left\|\mathbf{A}_p - \tilde{\mathbf{A}}_p\right\|_2 \le \frac{2M(d)}{e^{Nd}-1} \tag{5}$$

with

$$M(d) \le \max_{\lambda\in\sigma(\mathbf{A})}\int_0^{2\pi}\frac{|\gamma(\omega+id)\gamma'(\omega+id)|}{|\gamma(\omega+id)-\lambda|}d\omega\,, \tag{6}$$

where $\sigma(\mathbf{A})$ denotes the spectrum of $\mathbf{A}$.

Proof. This result follows directly from the exponential convergence of the trapezoidal rule for periodic functions [2], extended to matrices by diagonalization. ■

Equation (4) gives a method to compute an approximate projection of the matrix $\mathbf{A}$ which requires N matrix inversions. Normally, matrix inversion is considered very expensive, but in a number of cases, cheap inversion procedures are available, e.g., in case of $\mathcal{H}$-matrices [5], [7], inversion can be accomplished in $\mathcal{O}(M\log^2 M)$ operations.

5 Error analysis

We now turn to an error analysis, to determine the number of quadrature points required for a given accuracy. In addition to the number of quadrature points, the error will highly depend on the chosen curve $\gamma(\omega)$. We first determine the optimal choice of $\gamma(\omega)$, restricting ourselves to ellipses passing through -1 and 1.

5.1 Optimal choice of the integration curve Γ

Let Γ be parametrised by

$$\gamma(\omega) = \cos(\omega) + i\eta\sin(\omega). \tag{7}$$

Using scaling and translation, we can assume that the desired eigenvalues of $\mathbf{A}$ lie in an interval $[-b,b]$ with $b<1$ and that the other eigenvalues of $\mathbf{A}$ lie in $(-\infty,-B]\cup[B,\infty)$ for $B>1$ (Figure 1). To determine the quadrature error, we consider the requirements in Theorem 2. The width of the strip D_d in Theorem 2 depends on the parameter η. In our case, the larger we can choose the width d of D_d, the better the convergence behavior of the trapezoidal rule. In Theorem 3, we show that the largest possible width d^* is obtained for

$$\eta^* = \sqrt{1-\beta^2} \text{ with } \beta = \max\left\{b, -\frac{B}{2}+\frac{\sqrt{B^2+8}}{2}\right\}.$$

Geometrically, this is precisely the minimal value of η, for which the distance of $\gamma(\omega)$ to points on the interval $(-b,b)$ is minimal for $\omega=0$ or $\omega=\pi$.

Theorem 3 *For $\eta^* = \sqrt{1-\beta^2}$, with*

$$\beta = \max\left\{b, -\frac{B}{2}+\frac{\sqrt{B^2+8}}{2}\right\},$$

$(\gamma(z)\mathbf{I}-\mathbf{A})$ is nonsingular in a strip $D_{d^*} := \{z\in\mathbb{C}\,|\,|\mathrm{Im}(z)|<d^*\}$ for maximal possible $d^* := \cosh^{-1}(1/\beta)$.

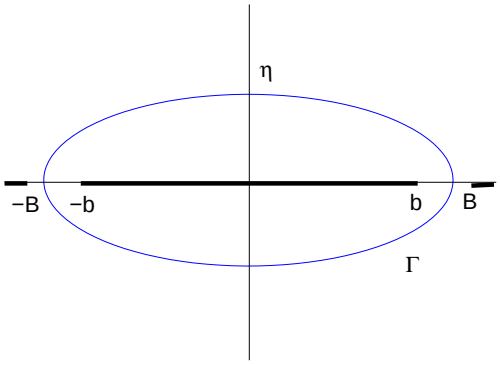

Figure 1: Ellipse $\gamma(\omega)$

Proof. To guarantee nonsingular $(\gamma(z)\mathbf{I}-\mathbf{A})$, we require that $\gamma(z)$ does not coincide with an eigenvalue of $\mathbf{A}$ for any $z = x+iy \in D_d$. We have assumed that the spectrum of $\mathbf{A}$ lies in the set $[-b,b]\cup(-\infty,-B]\cup[B,\infty)$. Thus, regularity in a strip D_{d^*} is guaranteed if $\gamma(z) \notin [-b,b]\cup(-\infty,-B]\cup[B,\infty)$ for all $z \in D_{d^*}$. We have

$$\begin{aligned}\mathrm{Re}(\gamma(x+iy)) &= \cos(x)\left(\frac{1+\eta}{2}e^{-y}+\frac{1-\eta}{2}e^{y}\right),\\ \mathrm{Im}(\gamma(x+iy)) &= \sin(x)\left(\frac{1+\eta}{2}e^{-y}-\frac{1-\eta}{2}e^{y}\right).\end{aligned}$$

For $x=\pi/2$ and $x=3\pi/2$, $\mathrm{Re}(\gamma(x+iy))=0$. To ensure that $\gamma(x+iy)\neq 0$, we require that $\left(\frac{1+\eta}{2}e^{-y}-\frac{1-\eta}{2}e^{y}\right) > 0$, for $|y|<d^*$, leading to

$$\eta \geq \frac{e^{d^*}-e^{-d^*}}{e^{d^*}+e^{-d^*}}. \tag{8}$$

For $\mathrm{Re}(z)=k\pi$, $\gamma(z)$ is purely real. To fulfill $\gamma(z)\notin[-b,b]\cup(-\infty,-B]\cup[B,\infty)$, we require that

$$\left(\frac{1+\eta}{2}e^{d^*}+\frac{1-\eta}{2}e^{-d^*}\right) > b, \tag{9}$$

$$\left(\frac{1+\eta}{2}e^{-d^*}+\frac{1-\eta}{2}e^{d^*}\right) < B. \tag{10}$$

Conditions (9) and (10) become less restrictive, the smaller η. Inserting the lower bound $\eta^* = \frac{e^{d^*}-e^{-d^*}}{e^{d^*}+e^{-d^*}}$, (9) becomes

$$\frac{2}{e^{d^*}+e^{-d^*}} = \frac{1}{\cosh d^*} \geq b\,.$$

Inequality (10) becomes

$$\frac{e^{2d^*}+e^{-2d^*}}{e^{d^*}+e^{-d^*}} \leq B\,,$$

which leads to

$$\cosh d^* \leq \frac{B}{4}+\frac{\sqrt{B^2+8}}{4}\,.$$

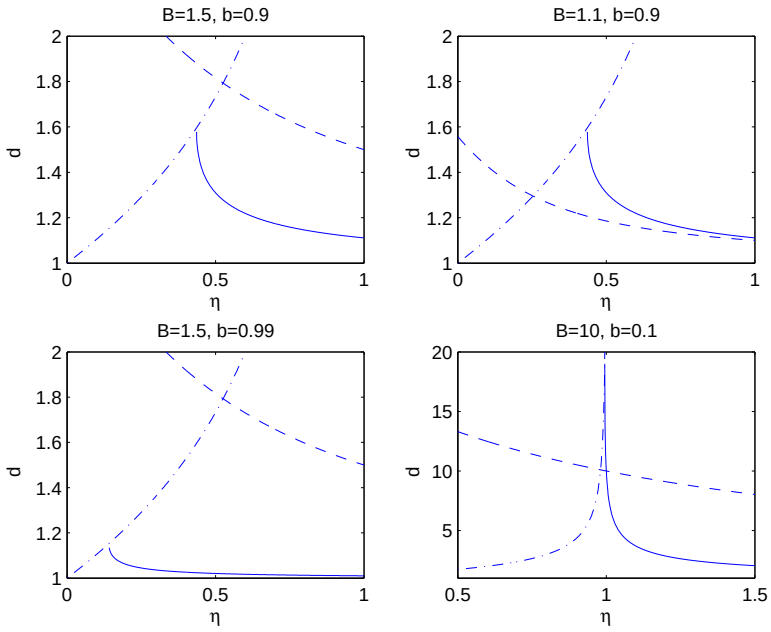

Figure 2: Bounds on d depending on η. (12)(dotted), (13) (solid), (14) (dashed)

Thus, the maximal choice of d^* fulfills

$$\cosh d^* = \min\left\{\frac{1}{b}, \frac{B}{4} + \frac{\sqrt{B^2+8}}{4}\right\} =: \frac{1}{\beta},$$

and is obtained for

$$\eta^* = \frac{\sinh d^*}{\cosh d^*} = \sqrt{1-\beta^2}.$$

■

5.2 General ellipse

For other η, the strip D_d becomes smaller, and thus the convergence rate will be slower. To illustrate the dependence of d on η, in Figure 2, we plot the lines representing upper bounds on d which need to be fulfilled for different values of b and B. The derivation of these upper bounds is given in Theorem 4.

Theorem 4 *Using the notation*

$$\beta_1 = \max\{b + \sqrt{b^2-(1-\eta^2)}, \frac{(1+\eta)^2}{B+\sqrt{B^2-(1-\eta^2)}}, \sqrt{1-\eta^2}\},$$

where the minimum is only taken over those terms that are real (e.g., if $b^2 < (1-\eta^2)$, the first term is not considered), $(\gamma(z)\mathbf{I} - \mathbf{A})$ is nonsingular in the strip $D_d := \{z \in \mathbb{C} : |\mathrm{Im}(z)| < d\}$ for d satisfying

$$e^d \le \frac{1+\eta}{\beta_1}. \tag{11}$$

Proof. As in Theorem 3, we require conditions (8), (9), (10) to be fulfilled, leading to

$$\eta \ge 1 \quad \text{or} \quad e^d \le \sqrt{\frac{1+\eta}{1-\eta}}, \tag{12}$$

$$1-\eta^2 \ge b^2 \quad \text{or} \quad e^{-d} \ge \frac{b+\sqrt{b^2-(1-\eta^2)}}{1+\eta}, \tag{13}$$

$$e^d \le \frac{B+\sqrt{B^2-(1-\eta^2)}}{1+\eta}. \tag{14}$$

In Figure 2, we plot the conditions on d. We take d to be defined by the strongest condition. ■

5.3 Convergence rate of the trapezoidal rule

In Theorem 2, we have stated the exponential convergence of the quadrature error in (4). We now derive bounds on the term $M(d)$ in equation (5). To obtain a bound on $M(d)$, we need the following conjecture,

Conjecture 5 *We have*

$$\max_{\substack{0\le\alpha\le 1\\ |\lambda|\le\frac{1+\alpha}{2}}} \int_0^{2\pi} \sqrt{\frac{e^{4\delta}+(\alpha e^{\delta})^4-2\alpha^2\cos 4\omega}{(\cos\omega(e^{\delta}+\alpha e^{-\delta})-2\lambda)^2+\sin^2\omega(e^{\delta}-\alpha e^{-\delta})}}d\omega \le c\log 1/\delta\,, \tag{15}$$

with $c\approx 4.6$.

This conjecture can be validated by numerical evaluation of the integral. The integral on the left hand side is maximal for $\lambda=\frac{1+\alpha}{2}$ and becomes singular for $\delta=0$. In Figure 3 we investigate the growth rate of the integral for $\lambda=\frac{1+\alpha}{2}$ as $\delta\to 0$. The logarithmic growth is clearly validated and the constant c in Conjecture 5 can be estimated by $c\approx 4.6$. We have the following error estimate.

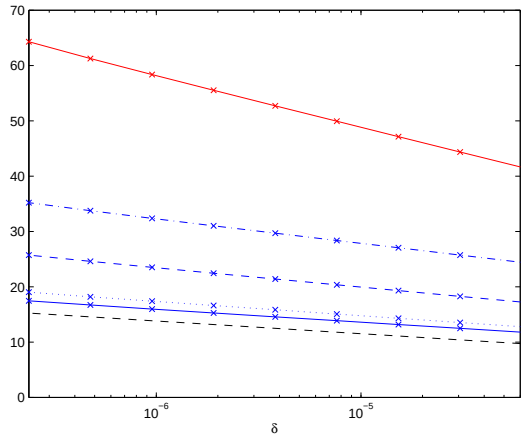

Figure 3: Lhs. of (15) for different values of α. From top to bottom: $\alpha=1$, $\alpha=0.9$, $\alpha=0.5$, $\alpha=0.1$, $\alpha=0$. Bottom dashed line: $\log(1/\delta)$

Theorem 6 *Assuming that (15) holds, the error using the trapezoidal rule (4) on $\gamma(\omega)$ as in (7) is bounded by*

$$\left\|\mathbf{A}_p-\tilde{\mathbf{A}}_p\right\|_2 \le \beta_1 c\log(N)\left(e\cdot\left(\frac{1+\eta}{\beta_1}\right)^N-1\right)^{-1}$$

with $c\approx 4.6$.

Proof. Given η, let d^* be the maximal value for d such that $\gamma(z)$ is analytic in D_d. According to Theorem 4,

$$e^{d^*}=\frac{1+\eta}{\beta_1}\,.$$

Defining in addition,

$$\beta_2=\min\{b-\sqrt{b^2-(1-\eta^2)},\frac{(1-\eta)^2}{B-\sqrt{B^2-(1-\eta^2)}},\sqrt{1-\eta^2}\}\,,$$

we rewrite

$$\begin{aligned}\gamma(\omega + i(d^* - \delta)) &= \frac{1}{2}(e^{i\omega}e^{-d^*}e^{\delta}(1+\eta) + e^{-i\omega}e^{d^*}e^{-\delta}(1-\eta^2)/(1+\eta)) \\ &= \frac{1}{2}\left(e^{i\omega}\beta_1 e^{\delta} + e^{-i\omega}\frac{(1-\eta^2)}{\beta_1}e^{-\delta}\right) \\ &= \frac{1}{2}\left(e^{i\omega}\beta_1 e^{\delta} + e^{-i\omega}\beta_2 e^{-\delta}\right)\,,\end{aligned}$$

and with some calculations, we arrive at

$$|\gamma(\omega + i(d^* - \delta))\gamma'(\omega + i(d^* - \delta))| = \frac{1}{4}\sqrt{(\beta_1^4 e^{4\delta} + \beta_2^4 e^{-4\delta}) - 2(\beta_1\beta_2)^2\cos 4\omega}\,.$$

Some transformations lead to

$$|\gamma(\omega + i(d^* - \delta)) - \lambda| = \frac{1}{2}\sqrt{(2\lambda - \cos\omega(e^{\delta}\beta_1 + e^{-\delta}\beta_2))^2 + \sin^2\omega(e^{\delta}\beta_1 - e^{-\delta}\beta_2)^2}\,,$$

and

$$M(d^* - \delta) \leq \max_{\lambda_i \in \sigma(\mathbf{A})} \frac{\beta_1}{2}\int_0^{2\pi} \sqrt{\frac{e^{4\delta} + (\frac{\beta_2}{\beta_1}e^{-\delta})^4 - 2(\frac{\beta_2}{\beta_1})^2\cos 4\omega}{(\cos\omega(e^{\delta} + e^{-\delta}\frac{\beta_2}{\beta_1}) - 2\frac{\lambda_i}{\beta_1})^2 + \sin^2\omega(e^{\delta} - e^{-\delta}\frac{\beta_2}{\beta_1})^2}}d\omega\,.$$

Since $\frac{|\lambda_i|}{\beta_1} \leq \frac{b}{\beta_1} \leq \frac{1+\frac{\beta_2}{\beta_1}}{2}$, we can apply Conjecture 5 with $\alpha = \frac{\beta_2}{\beta_1}$ to arrive at

$$M(d^* - \delta) \leq \beta_1\frac{c}{2}\log 1/\delta\,.$$

Choosing $\delta = \frac{1}{N}$, and combining this result with (5), we arrive at the theorem. ■

In the above estimates, we have assumed that for the computation of $\tilde{\mathbf{A}}_p$ in (4), the inverses of $(\gamma(2\pi k/N)\mathbf{I} - \mathbf{A})$ are computed exactly. In practice, e.g., with $\mathcal{H}$-matrix techniques, they will only be computed to some accuracy ε. In the following lemma, we show how this error affects the accuracy of the approximate projection $\tilde{\mathbf{A}}_p$.

Lemma 7 *For $\omega_k := \frac{2\pi k}{N}$, $k = 0, \ldots, N-1$, denote by $\mathbf{A}_{\omega_k}^{-1}$ an approximate inverse of $(\gamma(\omega_k)\mathbf{I} - \mathbf{A})$ with $\gamma(\omega)$ defined by (7). Assume that*

$$\|\mathbf{A}_{\omega_k}^{-1} - (\gamma(\omega_k)\mathbf{I} - \mathbf{A})^{-1}\|_2 \leq \varepsilon\,.$$

Then for

$$\hat{\mathbf{A}}_p := \frac{1}{2\pi i}\frac{2\pi}{N}\sum_{k=0}^{N-1}\gamma(\omega_k)\,\gamma'(\omega_k)\,\mathbf{A}_{\omega_k}^{-1}\,,$$

$$\|\tilde{\mathbf{A}}_p - \hat{\mathbf{A}}_p\|_2 \leq \varepsilon(1+\eta^2)\,.$$

Proof. We have

$$\begin{aligned}\|\tilde{\mathbf{A}}_p - \hat{\mathbf{A}}_p\|_2 &\leq \frac{\varepsilon}{N}\sum_{k=0}^{N-1}|\gamma(\omega_k)\,\gamma'(\omega_k)| \\ &= \frac{\varepsilon}{N}\sum_{k=0}^{N-1}\sqrt{(1+\eta^4)\cos^2(\omega_k)\sin^2(\omega_k) + \eta^2(\cos^4(\omega_k) + \sin^4(\omega_k))} \\ &\leq \varepsilon(1+\eta^2)\,.\end{aligned}$$

■

τ	η^{opt}	η^*	$err(50)$	$err(100)$	$err^*(50)$	$err^*(100)$
2.2	0.92	0.92	$1.8 \cdot 10^{-9}$	$< 10^{-16}$	$3.1 \cdot 10^{-8}$	$8.5 \cdot 10^{-19}$
2.1	0.64	0.64	$1.1 \cdot 10^{-6}$	$1.6 \cdot 10^{-13}$	$1.9 \cdot 10^{-6}$	$3.3 \cdot 10^{-13}$
2.05	0.44	0.45	$1.1 \cdot 10^{-4}$	$1.6 \cdot 10^{-9}$	$1.9 \cdot 10^{-4}$	$3.2 \cdot 10^{-8}$
2.025	0.33	0.32	$2.8 \cdot 10^{-3}$	$1.1 \cdot 10^{-6}$	$4.9 \cdot 10^{-2}$	$2.2 \cdot 10^{-6}$

Table 1: Approximation error $\|\mathbf{A} - \tilde{\mathbf{A}}\|_2$

6 Numerical experiments

The above results are validated by the following experiments. Above, we have assumed that $\mathbf{A}$'s eigenvalues lie in $[-b, b] \cup (-\infty, -B] \cup [B, \infty)$ for $b < 1$ and $B > 1$.

The parameters of the above section can be scaled to more general subsets $(a_1, b_1) \cup (-\infty, a_1 - \delta] \cup [b_1 + \delta, \infty)$ by

$$\gamma(\omega) = \frac{b_1 + a_1}{2} + \tau \cos\omega + i\eta \sin\omega \,,$$

for some $\tau \in \left(\frac{b_1 - a_1}{2}, \frac{b_1 - a_1}{2} + \delta\right)$. The optimal ellipse, yielding the smallest quadrature error is obtained for $\eta = \sqrt{\tau^2 - \beta^2}$, where

$$\beta = \max\left\{ \frac{b_1 - a_1}{2}, \frac{-(b_1 - a_1 + 2\delta) + \sqrt{(b_1 - a_1 + 2\delta)^2 + 8\tau^2}}{4} \right\} .$$

For optimal η, the error estimate of Theorem 6 becomes

$$\|\mathbf{A} - \tilde{\mathbf{A}}\|_2 \leq 4.6\beta \log(N) \left(e \cdot \left(\frac{\tau + \eta}{\beta} \right)^N - 1 \right)^{-1} =: err^*(N) \,.$$

We investigate the following matrix $\mathbf{A} \in \mathbb{R}^{M \times M}$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix} .$$

Its eigenvalues lie in the interval $[0, 4]$. We consider $M = 80$.

In the first example, we project the matrix onto itself, i.e., we investigate only the error in the quadrature $err(N) = \left\| \mathbf{A} - \tilde{\mathbf{A}} \right\|_2$. For the ellipse

$$\gamma(\omega) = 2 + \tau \cos\omega + i\eta \sin\omega \,,$$

we have experimentally determined the optimal choice η^{opt} depending on τ and compared this with the result from the previous section, which in this case yields $\eta^* = \sqrt{\tau^2 - 4}$. In Figure 4, we show the dependence of the error on η. The different curves denote different values of τ. We clearly see a sharp cusp in the error and a distinct value η^{opt} for which the error is minimal. In Table 1, we see that this value is almost exactly the value η^* predicted in the previous section. In Figure 5, we show the error for increasing number of quadrature points N for different values of η. We see exponential decay of the quadrature error $err(N)$ as N increases and the superior convergence properties when choosing optimal η. We also investigate the convergence behavior, as the ellipse approaches the extremal eigenvalues ($\tau \to 2$). We have used $N = 50$ and $N = 100$ quadrature points. In Table 1, we give the actual errors $err(\cdot)$ and the errors from the theoretical error bounds, $err^*(\cdot)$. Both the optimal choice of η and the convergence rate are as predicted by the theoretical results, the errors differing by a factor less than two.

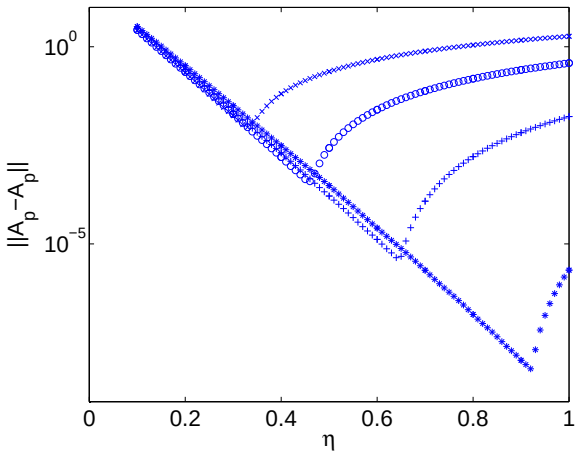

Figure 4: The l_2-error $err(50)$ for different values of τ. $\tau = 2.2$ $(*)$, $\tau = 2.1$ $(+)$, $\tau = 2.05$ $(\circ)$, $\tau = 2.025$ $(\times)$.

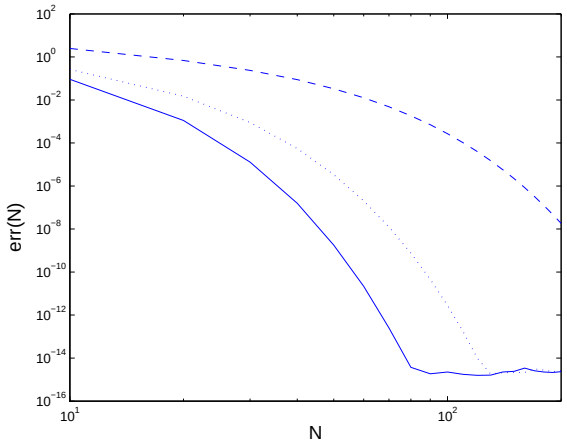

Figure 5: The l_2-error for $\tau = 2.2$ and different values of η, $\eta = \tau$ (dashed), $\eta = \eta^*$ (solid), $\eta = \eta^* + 0.1$ (dots).

In the next example, we investigate the case of a projection. For this, we project the above operator onto the eigenspace of the eigenvalues in the interval $[0.95, 3.05]$. We here use $M = 40$. The closest eigenvalues are $\lambda_j = 2.9554$, $\lambda_{j+1} = 3.0871$. Again, in Figure 6 and Table 2, the error for different values of η is shown. We use three different values of τ for which the optimal ellipses are depicted in Figure 7. Again, convergence is as predicted.

7 Acknowledgement

The author would like to thank W. Hackbusch for giving the initial idea for this work.

References

[1] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology, Vol. 5, Evolutions problems I.* Springer, 1992.

[2] P. J. Davis. On the numerical integration of periodic analytic functions. In R. Langer, editor, *Symposium on Numerical Approximations*, pages 45–59. The University of Wisconsin Press, 1959.

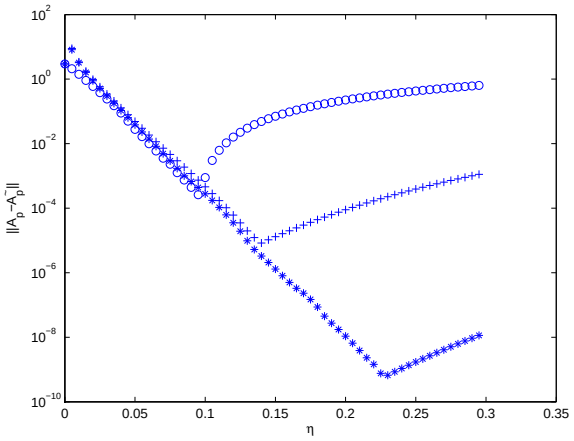

Figure 6: The error in the l_2-norm for different values of η. $\tau = 0.96$ ($\circ$), $\tau = 1.01$ ($*$), $\tau = 1.06$ ($+$).

τ	η^{opt}	η^*	$err(100)$	$err(200)$	$err^*(100)$	$err^*(200)$
0.9604	0.095	0.098	$2.6 \cdot 10^{-4}$	$1.3 \cdot 10^{-8}$	$2.7 \cdot 10^{-4}$	$1.1 \cdot 10^{-8}$
1.0104	0.23	0.22	$6.7 \cdot 10^{-10}$	$< 10^{-16}$	$1.2 \cdot 10^{-9}$	$2.3 \cdot 10^{-19}$
1.0604	0.14	0.14	$8.3 \cdot 10^{-6}$	$2.2 \cdot 10^{-11}$	$1.8 \cdot 10^{-5}$	$4.8 \cdot 10^{-11}$

Table 2: Approximation error $\|\mathbf{A}_p - \tilde{\mathbf{A}}_p\|_2$

[3] M. Espig and W. Hackbusch. On the robustness of elliptic resolvents computed by means of the technique of hierarchical matrices. Technical Report 11/2007, Max Planck Institute for Mathematics in the Sciences, 2007. to appear in: Applied Numerical Mathematics.

[4] G. H. Golub and H. A. van der Vorst. Eigenvalue computation in the 20th century. *J. Comp. Appl. Math.*, 123:35–65, 2000.

[5] W. Hackbusch. A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices. *Computing*, 62:89–108, 1999.

[6] W. Hackbusch and M. Bebendorf. Existence of $\mathcal{H}$-matrix approximants to the inverse FE-matrix of elliptic operators with l^∞-coefficients. *Numerische Mathematik*, 95:1–28, 2003.

[7] W. Hackbusch and B. N. Khoromskij. A sparse $\mathcal{H}$-matrix arithmetic. Part II: Application to multi-dimensional problems. *Computing*, 64:21–47, 2000.

[8] B. N. Parlett. *The Symmetric Eigenvalue Problem*. SIAM, 1998.

[9] D. C. Sorensen. Numerical methods for large eigenvalue problems. *Acta Numerica*, pages 519–584, 2002.

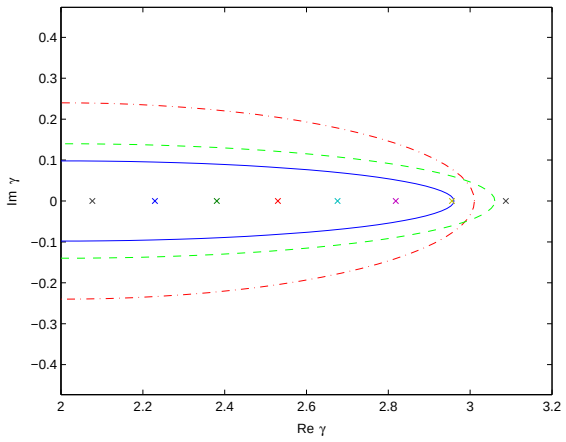

Figure 7: The optimal ellipse for $\tau = 0.9604$ (solid), $\tau = 1.0104$ (dash-dot), $\tau = 1.0604$ (dashed). The crosses denote the eigenvalues of $\mathbf{A}$.