A GEOMETRIC THEORY FOR PRECONDITIONED INVERSE ITERATION

II: CONVERGENCE ESTIMATES

KLAUS NEYMEYR

ABSTRACT. The topic of this paper is a convergence analysis of preconditioned inverse iteration (PINVIT). A sharp estimate for the eigenvalue approximations is derived; the eigenvector approximations are controlled by an upper bound for the residual vector. The analysis is mainly based on extremal properties of various quantities which define the geometry of PINVIT.

1. INTRODUCTION

Let A be a symmetric positive definite matrix whose eigenvalues of arbitrary multiplicity are given by $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n$. Preconditioned inverse iteration (PINVIT), as introduced in Part I, maps a given vector x with the Rayleigh quotient

$$\lambda := \lambda(x) = \frac{(x, Ax)}{(x, x)} \tag{1.1}$$

to the next iterate

$$x' = x - B^{-1}(Ax - \lambda x). \tag{1.2}$$

Therein B^{-1} is a symmetric and positive definite matrix which approximates the inverse of A so that with respect to the A–norm

$$\|I - B^{-1}A\|_A \leq \gamma \quad \text{for a given} \quad \gamma \in [0, 1[. \tag{1.3}$$

In order to derive a sharp estimate for the Rayleigh quotient of x', one has to analyze its dependence on the choice of the preconditioner as well as on all vectors x having a fixed Rayleigh quotient equal to λ.

In Part I the dependence on the preconditioner is analyzed: For $x \neq 0$ and for $\gamma \in [0, 1[$ the Rayleigh quotient $\lambda(x')$ takes its supremum with respect to all preconditioners satisfying (1.3) in a vector of the form $w = \beta(A + \alpha I)^{-1}x$. Therein β is a scaling constant and α is a positive shift parameter. Hence w can be represented by applying inverse iteration with a positive shift to the vector x.

Here we analyze the dependence of these suprema on all those x, whose Rayleigh quotient has a fixed value. To be more precise we determine for given $\lambda \in [\lambda_1, \lambda_n]$ and $\gamma \in [0, 1[$ the maximum

$$\sup\{\lambda(x');\ B \text{ satisfies (1.3)},\ x \neq 0 \text{ with } \lambda(x) = \lambda\}. \tag{1.4}$$

This maximum provides a practicable convergence estimate for PINVIT, since the Rayleigh quotient of x and the constant γ, which describes the quality of the preconditioner, are known quantities. The maximum (1.4) represents the case of poorest convergence, i.e. the

1991 *Mathematics Subject Classification.* Primary 65F15, 65N25; Secondary 65N30.

minimal decrease of the Rayleigh quotient if PINVIT is applied to an arbitrary iterate x with $\lambda = \lambda(x)$.

Our main result concerning the eigenvalue approximations is given by the following theorem. Section 3.2 contains a convergence estimate for the eigenvalue approximations.

Theorem 1.1. *Let $x^{(0)} \neq 0$ be an initial vector with the Rayleigh quotient $\lambda^{(0)} := \lambda(x^{(0)})$ and denote the sequence of iterates of preconditioned inverse iteration (1.2) by*

$$(x^{(j)}, \lambda^{(j)}), \qquad j = 0, 1, 2, \ldots,$$

where $\lambda^{(j)} = \lambda(x^{(j)})$. The preconditioner is assumed to satisfy (1.3) for some $\gamma \in [0, 1[$.

Then the sequence of Rayleigh quotients $\lambda^{(j)}$ decreases monotonically and $(x^{(j)}, \lambda^{(j)})$ converges to an eigenpair of A. Moreover, denote by x some iterate and let $\lambda = \lambda(x)$ be its Rayleigh quotient. Then for the new iterate x', given by (1.2), with $\lambda' = \lambda(x')$ it holds that:

(1) *If $\lambda = \lambda_1$ or $\lambda = \lambda_n$, then PINVIT is stationary in an eigenvector of A.*
If $\lambda = \lambda_i$, with $2 \leq i \leq n-1$, then λ' takes its maximal value $\lambda' = \lambda_i$ for PINVIT beeing applied to $x = x_i$, where x_i is an eigenvector of A corresponding to λ_i.

(2) *If $\lambda_i < \lambda < \lambda_{i+1}$, then the Rayleigh quotient takes its maximal value $\lambda' = \lambda_{i,i+1}(\lambda, \gamma)$ (under all x with $\lambda = \lambda(x)$ and all admissible preconditioners) for PINVIT being applied to the vector $x = x_{i,i+1}$ with*

$$x_{i,i+1} = \omega_1 x_i + \omega_2 x_{i+1}.$$

(Therein x_i is an eigenvector of A corresponding to the eigenvalue λ_i. The values ω_1^2 and ω_2^2 are uniquely determined by $\lambda(x_{i,i+1}) = \lambda$ and $|x_{i,i+1}| = 1$.) The supremum $\lambda' = \lambda_{i,i+1}(\lambda, \gamma)$ is given by

$$\begin{aligned} \lambda_{i,j}(\lambda,\gamma) \;=\; & \lambda\lambda_i\lambda_j(\lambda_i+\lambda_j-\lambda)^2 / \\ & \big(\gamma^2(\lambda_j-\lambda)(\lambda-\lambda_i)(\lambda\lambda_j+\lambda\lambda_i-\lambda_i^2-\lambda_j^2) \\ & -2\gamma\sqrt{\lambda_i\lambda_j}(\lambda-\lambda_i)(\lambda_j-\lambda) \\ & \quad \sqrt{\lambda_i\lambda_j+(1-\gamma^2)(\lambda-\lambda_i)(\lambda_j-\lambda)} \\ & -\lambda(\lambda_i+\lambda_j-\lambda)(\lambda\lambda_j+\lambda\lambda_i-\lambda_i^2-\lambda_i\lambda_j-\lambda_j^2)\big). \end{aligned} \tag{1.5}$$

For the relative decrease of $\lambda' = \lambda_{i,i+1}(\lambda, \gamma)$ towards the nearest eigenvalue λ_i smaller than λ it holds

$$\Phi_{i,i+1}(\lambda,\gamma) = \frac{\lambda_{i,i+1}(\lambda,\gamma)-\lambda_i}{\lambda-\lambda_i} < 1. \tag{1.6}$$

The proof of Theorem 1.1 is given in Section 3.

We explain the result by discussing the five–point finite difference discretization of the eigenproblem for the Laplacian on the square $[0, \pi]^2$ with homogeneous Dirichlet boundary conditions. The eigenvalues of the continuous problem $\lambda^{(k,l)}$ and of the finite difference discretization $\lambda_h^{(k,l)}$, for the mesh size h, are given by

$$\lambda^{(k,l)} = k^2 + l^2, \qquad \lambda_h^{(k,l)} = \frac{4}{h^2}\left(\sin^2(\frac{kh}{2}) + \sin^2(\frac{lh}{2})\right). \tag{1.7}$$

The 10 smallest eigenvalues (with multiplicity) read $(2, 5, 5, 8, 10, 10, 13, 13, 17, 17)$; for $h = \pi/50$ these eigenvalues coincide with $\lambda_h^{(k,l)}$ within the 1 percent range. Figure 1

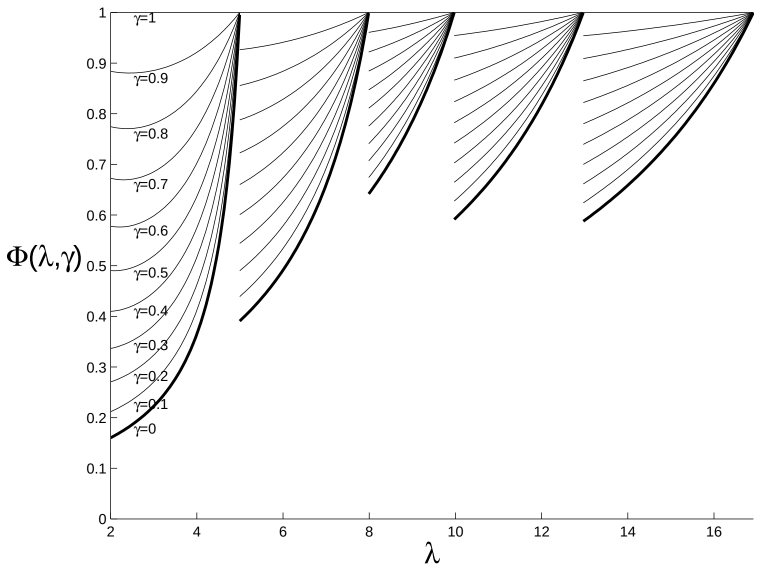

FIGURE 1. *Convergence estimates $\Phi_{i,i+1}(\lambda,\gamma)$ for the 10 smallest eigenvalues $\lambda_h^{(k,l)}$ given by Equation (1.7).*

shows the convergence estimates $\Phi_{i,i+1}(\lambda,\gamma)$ for the eigenvalues $\lambda_i := \lambda_h^{(k,l)}$. Note that the estimates are valid independently of the multiplicity of the eigenvalues.

The bold lines represent the case $\gamma = 0$, i.e. $B = A$, for which PINVIT is identical with the inverse iteration procedure (INVIT). We explicitly derive this estimate describing the poorest convergence of INVIT, by inserting $\gamma = 0$ and $j = i + 1$ in (1.5) and obtain

$$\lambda_{i,i+1}(\lambda, 0) = \left(\lambda_i^{-1} + \lambda_{i+1}^{-1} - (\lambda_i + \lambda_{i+1} - \lambda)^{-1}\right)^{-1},$$

and

$$\Phi_{i,i+1}(\lambda, 0) = \frac{\lambda_i^2}{\lambda_i^2 + (\lambda_{i+1} - \lambda)(\lambda_i + \lambda_{i+1})}.$$

In each interval $[\lambda_i, \lambda_{i+1}[$ INVIT attains its poorest convergence in those vectors which are spanned by the eigenvectors corresponding to λ_i and λ_{i+1}.

For $\lambda = \lambda_{i+1}$ we have $\Phi_{i,i+1}(\lambda_{i+1}, \gamma) = 1$, which expresses the fact that INVIT and PINVIT are stationary in the eigenvectors of A. The curves in Figure 1 for $\gamma > 0$ describe the case of poorest convergence of PINVIT: For $\gamma = 1$ PINVIT is stationary and for smaller γ PINVIT behaves more and more like inverse iteration. By Theorem 1.1 this poorest convergence is attained in the same vectors in which inverse iteration attains its poorest convergence, but additionally the preconditioner is to be chosen appropriately.

Consider the sequence of iterates

$$(x^{(j)}, \lambda^{(j)}), \qquad j = 0, 1, 2, \ldots$$

of PINVIT. If one starts with an initial eigenvalue approximation larger than λ_k, it cannot be said in principle when the Rayleigh quotients $\lambda^{(j)}$ move from one interval $[\lambda_k, \lambda_{k+1}]$ to the next interval of smaller eigenvalues. For the moment we assume the Rayleigh quotients to converge to λ_1; the general case is discussed in the following. Once having reached the interval $[\lambda_1, \lambda_2]$ then the "one–step" estimates Φ can be used to define a *convergence rate* estimate $\Theta_{1,2}(\lambda, \gamma)$ for PINVIT

$$\Theta_{1,2}(\lambda,\gamma) := \sup_{\lambda_1<\tilde{\lambda}\leq\lambda} \Phi_{1,2}(\tilde{\lambda},\gamma), \qquad \lambda \in]\lambda_1, \lambda_2]. \tag{1.8}$$

Confer Figure 1 to see that $\Theta_{1,2}(\lambda, \gamma)$ only slightly differs from $\Phi_{1,2}(\lambda, \gamma)$. (E.g. in the interval $[2, 5]$ the curve $\gamma = 0.9$ takes its minimum in $\lambda \approx 2.44$.) The convergence rate $\Theta_{1,2}(\lambda, \gamma)$ is the guaranteed relative decrease of the Rayleigh quotients in the sense of Equation (1.6) for all further iterates of PINVIT. Hence the Rayleigh quotients $\lambda^{(j)}$ converge *linearly* to λ_1 with the convergence rate $\Theta_{1,2}$. In any case $\Theta_{1,2}(\lambda, \gamma)$ can be bounded away from 1 by using the unsharp estimate (3.3) for $\Phi_{1,2}(\lambda, \gamma)$.

In principle, it cannot be guaranteed that PINVIT converges to the *smallest* eigenvalue λ_1 and corresponding eigenvector of A, since PINVIT for some choice of the preconditioner may reach stationarity in higher eigenvectors/values, even if the initial vector has some contribution from the eigenvector to the smallest eigenvalue. But note that all eigenvectors to eigenvalues larger than λ_1 form a null set. In practice, due to rounding errors such an early breakdown of PINVIT is as unlikely as that inverse iteration may get stuck in higher eigenvalues. Hence as a result of rounding errors INVIT as well as PINVIT converge from scratch to the smallest eigenvalue λ_1 and a corresponding eigenvector. In exact arithmetic convergence of PINVIT to λ_1 is guaranteed if the Rayleigh quotient of the initial vector is less than λ_2. Depending on the choice of the preconditioner and on the eigenvector expansion of the vector x, PINVIT may converge much more rapidly than suggested by the estimate (1.5).

It is an important result that the convergence of PINVIT in the case that A is the mesh analog of a differential operator does not depend on the mesh size h and hence on the number of the variables since Equation (1.5) is a function of λ, λ_i, λ_{i+1} and γ and does not depend on the largest eigenvalue. We assume that there is no implicit dependence on λ_n or the mesh size via γ: For the best multigrid preconditioners, (1.3) is satisfied for γ bounded away from 1 independently on the mesh size. Furthermore, in case of adaptive multigrid eigenvalue computation with a good coarse grid approximation, one expects that the eigenvalue approximations on all levels of refinement are located in $[\lambda_1, \lambda_2[$ if the discretization error is small in comparison to $\lambda_2 - \lambda_1$. In this case the bound Θ gives a reliable convergence rate estimate.

Hence, depending on the quality of the preconditioner, eigenvector/eigenvalue computation can be done with a grid independent rate while the convergence rates are of comparable magnitude with that of multigrid methods for boundary value problems. Therefore PINVIT can be viewed as the eigenproblem counterpart of multigrid algorithms for the solution of boundary value problems, see also [5].

The outline of the remainder of this paper is as follows: In Section 2 we analyze extremal properties of some geometric quantities, which define the geometry of the set of the iterates of PINVIT, with respect to all vectors having a fixed Rayleigh quotient. These quantities are for instance the Euclidean norm of the gradient vector of the Rayleigh quotient and various opening angles of cones to be defined later. In Section 3 these results and the mini–dimensional analysis of PINVIT, as given in Part I, are combined to derive sharp convergence estimates for PINVIT.

2. Extremal quantities on level sets of the Rayleigh quotient

For the following analysis we adopt the notation introduced in Part I. We make use of the c–basis introduced in Section 2 and assume (AC), see Section 4 of Part I, summarizing some nonrestrictive assumptions on the vector c. Furthermore, we consider an eigenvalue problem with only simple eigenvalues, cf. Section 3 of Part I .

2.1. **Extremal behavior of** $|\nabla\lambda(c)|$**.**

In this section we analyze the extremal behavior of the Euclidean norm of the Rayleigh quotient $|\nabla\lambda(c)|$ on the level set $L(\lambda)$, which is defined to consist of all nonnegative vectors on the unit sphere whose Rayleigh quotient is equal to λ

$$L(\lambda) = \{c \in \mathbf{R}^n;\ |c| = 1,\ c \geq 0,\ \lambda(c) = \lambda\}. \tag{2.1}$$

Theorem 2.1. *Let* $\lambda \in]\lambda_1, \lambda_n[$*. The gradient of the Rayleigh quotient with respect to the* c*–basis reads*

$$\nabla\lambda(c) = \frac{2}{(c, \Lambda^{-1}c)}(I - \lambda\Lambda^{-1})c. \tag{2.2}$$

For its Euclidean norm $|\nabla\lambda(c)|$ *on* $L(\lambda)$ *holds:*

(a) *If* $\lambda = \lambda_i$*, with* $i \in \{2, \dots, n-1\}$*, then* $|\nabla\lambda(c)|$ *takes its minimum* $|\nabla\lambda(e_i)| = 0$ *in the* i*–th unit vector* e_i*.*

(b) *If* $\lambda_i < \lambda < \lambda_{i+1}$*, then* $|\nabla\lambda(c)|$ *takes its minimum in the vector*

$$c_{i,i+1} := (0, \dots, 0, c_i, c_{i+1}, 0, \dots, 0)^T \in L(\lambda),$$

which has only two non–zero components c_i *and* c_{i+1}*.*

Proof. Property (a) is a consequence of Lemma 4.1 in Part I. We employ the method of Lagrange multipliers to determine necessary conditions for constrained relative extrema of $|\nabla\lambda(c)|$ with respect to the constraints $|c| = 1$ and $(c, \Lambda^{-1}c) = 1/\lambda$. Inserting the constraint $\lambda(c) = \lambda$ in (2.2) we obtain $|\nabla\lambda(c)| = 2\lambda|(I-\lambda\Lambda^{-1})c|$. Since $2\lambda|(I-\lambda\Lambda^{-1})c|$ takes its extremal value on $L(\lambda)$ in the same arguments as $|(I - \lambda\Lambda^{-1})c|^2$, we consider a Lagrange function with multipliers μ and ν of the form

$$L(c) = |(I - \lambda\Lambda^{-1})c|^2 + \mu(|c|^2 - 1) + \nu((c, \Lambda^{-1}c) - \lambda^{-1}). \tag{2.3}$$

Hence, $\nabla L = 0$ reads

$$(I - \lambda\Lambda^{-1})^2 c + \mu c + \nu\Lambda^{-1}c = 0. \tag{2.4}$$

If $\lambda \in]\lambda_i, \lambda_{i+1}[$, there are at least two indexes $k < l$ so that the components c_k and c_l are nonzero, because c is not equal to any of the unit vectors. Hence, the Lagrange multipliers μ and ν can be determined from (2.4) by solving the linear system

$$\begin{pmatrix} 1 & \lambda_k^{-1} \\ 1 & \lambda_l^{-1} \end{pmatrix} \begin{pmatrix} \mu \\ \nu \end{pmatrix} = \begin{pmatrix} -(1-\lambda\lambda_k^{-1})^2 \\ -(1-\lambda\lambda_l^{-1})^2 \end{pmatrix}.$$

Since $\lambda_l^{-1} - \lambda_k^{-1} \neq 0$, the unique solution reads

$$\mu = \frac{\lambda^2}{\lambda_k\lambda_l} - 1 \quad \text{and} \quad \nu = \frac{\lambda(2\lambda_k\lambda_l - \lambda(\lambda_k + \lambda_l))}{\lambda_k\lambda_l}.$$

Inserting μ and ν in the j-th component of (2.4) we obtain

$$\frac{\lambda^2(\lambda_j - \lambda_k)(\lambda_j - \lambda_l)}{\lambda_j^2\lambda_k\lambda_l} c_j = 0,$$

so that $c_j = 0$ for $j \neq k, l$. Hence, c has necessarily the form

$$c = c_{k,l} := (0, \dots, 0, c_k, 0, \dots, 0, c_l, 0, \dots, 0)^T \in L(\lambda). \tag{2.5}$$

We conclude that $\lambda = \lambda(c) \in]\lambda_k, \lambda_l[$. From $|c| = 1$ and $\lambda(c) = \lambda$ we get

$$c_k^2 = \frac{\lambda_k(\lambda_l - \lambda)}{\lambda(\lambda_l - \lambda_k)} \quad \text{and} \quad c_l^2 = \frac{\lambda_l(\lambda - \lambda_k)}{\lambda(\lambda_l - \lambda_k)}.$$

By direct computation follows

$$|\nabla\lambda(c)|^2 = \frac{4}{(c, \Lambda^{-1}c)^2}|(I - \lambda\Lambda^{-1})c|^2 = \frac{4\lambda^2(\lambda - \lambda_k)(\lambda_l - \lambda)}{\lambda_k\lambda_l}. \tag{2.6}$$

Since $\lambda_k < \lambda < \lambda_l$ one obtains

$$\frac{d}{d\lambda_k}|\nabla\lambda(c)|^2 = -\frac{4\lambda^3(\lambda_l - \lambda)}{\lambda_l\lambda_k^2} < 0 \quad \text{and} \quad \frac{d}{d\lambda_l}|\nabla\lambda(c)|^2 = \frac{4\lambda^3(\lambda - \lambda_k)}{\lambda_k\lambda_l^2} > 0.$$

Thus $|\nabla\lambda(c)|$ takes its minimal value in $c_{i,i+1}$. □

2.2. Extremal properties of the cone $C_\gamma(c)$.

The opening angle $\varphi_\gamma(c)$ of the circular cone $C_\gamma(c)$,

$$C_\gamma(c) = \{\zeta d;\ d \in E_\gamma(c),\ \zeta > 0\},$$

enclosing $E_\gamma(c)$ is defined by

$$\varphi_\gamma(c) := \sup_{z \in C_\gamma(c)} \arccos\left(\frac{\lambda\Lambda^{-1}c}{|\lambda\Lambda^{-1}c|}, \frac{z}{|z|}\right). \tag{2.7}$$

We show in the next lemma that the extremal properties of $|\nabla\lambda(c)|$ gained in Theorem 2.1 transfer to the opening angle $\varphi_\gamma(c)$.

Lemma 2.2. *Let $\lambda \in]\lambda_1, \lambda_n[$ and $\gamma \in [0, 1]$ be given. On the level set $L(\lambda)$ the opening angle $\varphi_\gamma(c)$ satisfies:*

(a) *If $\lambda = \lambda_i$, with $i \in \{2, \dots, n-1\}$, then φ_γ takes its minimum $\varphi_\gamma(e_i) = 0$ in the i–th unit vector.*

(b) *If $\lambda_i < \lambda < \lambda_{i+1}$, then φ_γ takes its minimum in the vector $c_{i,i+1} \in L(\lambda)$.*

Proof. Property (a) follows from the fact that $E_\gamma(e_i) = \{e_i\}$ for any $\gamma \in [0, 1]$ and hence $\varphi_\gamma(e_i) = 0$.

Using the orthogonal decomposition of Theorem 4.3 (Part I) one obtains for the opening angle of $C_\gamma(c)$

$$\varphi_\gamma(c) = \arcsin\frac{\gamma|(I - \lambda\Lambda^{-1})c|}{|\lambda\Lambda^{-1}c|} \quad \text{and} \quad \varphi_1(c) = \arctan\frac{|(I - \lambda\Lambda^{-1})c|}{|c|}. \tag{2.8}$$

First we show the proposition for $\gamma = 1$ and then conclude on $\gamma \in [0, 1[$. Since $\arctan(\cdot)$ is strictly monotone increasing, it is sufficient to show the extremal properties for the argument $\frac{|(I-\lambda\Lambda^{-1})c|}{|c|}$. Moreover, since $c \in L(\lambda)$ we obtain the same necessary conditions for relative extrema if we analyze $|(I - \lambda\Lambda^{-1})c|^2$. Hence, a Lagrange multiplier ansatz leads to the same Lagrange function as considered in Theorem 2.1. In the same way we obtain instead of (2.6) for the residual

$$|(I - \lambda\Lambda^{-1})c|^2 = \frac{(\lambda - \lambda_k)(\lambda_l - \lambda)}{\lambda_k\lambda_l}.$$

Differentiation with respect to λ_k and λ_l as in the proof of Theorem 2.1 establishes the extremal property of $\varphi_1(c)$ on $L(\lambda)$. Now let $\gamma \in [0, 1[$, then we have from (2.8)

$$\sin(\varphi_\gamma(c)) = \gamma \sin(\varphi_1(c)). \tag{2.9}$$

Since $\sin(\cdot)$ is a strictly monotone increasing function on $[0, \frac{\pi}{2}]$ and with $\varphi_1(c_{1,n}) = \min\{\varphi_1(c);\ c \in L(\lambda)\}$, one obtains for $\gamma \in [0, 1[$

$$\gamma \sin(\varphi_1(c_{1,n})) = \min\{\gamma \sin(\varphi_1(c))\,;\ c \in L(\lambda)\}.$$

Applying (2.9) and once more the monotonicity of $\sin(\cdot)$ leads to

$$\varphi_\gamma(c_{1,n}) = \min\{\varphi_\gamma(c);\ c \in L(\lambda)\},$$

which establishes the required result. □

The action of PINVIT can be understood as a shrinking of the initial cone $C_1(c)$: While the vector c lies on the boundary of $C_1(c)$, the new iterate is an element of the shrinked cone $C_\gamma(c)$. We define a (complementary) shrinking angle ψ_γ by

$$\psi_\gamma(c) = \varphi_1(c) - \varphi_\gamma(c). \tag{2.10}$$

The following lemma exhibits extremal properties of $\psi_\gamma(c)$ on the level set $L(\lambda)$.

Lemma 2.3. *Let $\lambda \in]\lambda_1, \lambda_n[$ and $\gamma \in [0, 1]$ be given. On the level set $L(\lambda)$ the shrinking angle $\psi_\gamma(c)$ fulfills:*

(a) *If $\lambda = \lambda_i$, with $i \in \{2, \ldots, n-1\}$, then ψ_γ takes its minimum $\psi_\gamma(e_i) = 0$ in the i–th unit vector.*

(b) *If $\lambda_i < \lambda < \lambda_{i+1}$, then ψ_γ takes its minimum in the vector $c_{i,i+1} \in L(\lambda)$.*

Proof. Let $a := \frac{|(I-\lambda\Lambda^{-1})c|}{|\lambda\Lambda^{-1}c|}$, then

$$\Psi_\gamma(a) := \psi_\gamma(c) = \arcsin(a) - \arcsin(\gamma a).$$

Property (a) is an immediate consequence of property (a) of Lemma 2.2. The extremal behavior of $\varphi_1(c)$ as described in Lemma 2.2 transfers to $a = \sin\varphi_1$, since $\sin(\cdot)$ is a strictly monotone increasing function on $[0, \frac{\pi}{2}]$. Differentiation of $\Psi_\gamma(a)$ shows that for $\gamma \in]0, 1[$

$$\frac{d}{da}\Psi_\gamma(a) = \frac{\sqrt{1-\gamma^2 a^2} - \sqrt{1-a^2}}{\sqrt{(1-a^2)(1-\gamma^2 a^2)}} > 0.$$

Thus $\Psi_\gamma(a)$ is strictly monotone increasing in a which completes the proof. □

2.3. **Orientation of the gradient $\nabla\lambda(w)$.**

In the sequel we determine the orientation of the gradient vector $\nabla\lambda(w)$ within points of suprema w of the Rayleigh quotient on $C_\gamma(c)$. The results gained here are used in the next section to analyze the angle dependence of the Rayleigh quotient within these points. Figure 2 illustrates the content of Lemma 2.4.

Lemma 2.4. *Let w be a point of a supremum of the Rayleigh quotient on the cone $C_\gamma(c)$ for $\gamma > 0$. Then*

(a) *If $\mathrm{span}\{w, \lambda\Lambda^{-1}c\}$ denotes the linear space spanned by w and $\lambda\Lambda^{-1}c$, then*

$$\nabla\lambda(w) \in \mathrm{span}\{w, \lambda\Lambda^{-1}c\}. \tag{2.11}$$

(b) *There is a nonzero $\theta \in \mathbf{R}$ in a way that $w - \lambda\Lambda^{-1}c = \theta\nabla\lambda(w)$.*

Proof. Since $\nabla\lambda(w) = \frac{2}{(w,\Lambda^{-1}w)}(w - \lambda\Lambda^{-1}w)$, we only have to show that

$$\Lambda^{-1}w \in \mathrm{span}\{w, \Lambda^{-1}c\},$$

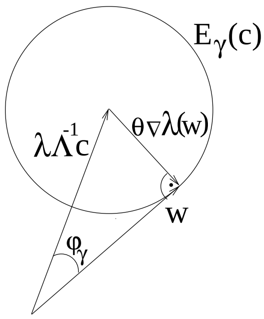

FIGURE 2. *Orientation of $\nabla\lambda(w)$ in a point w of an extremum.*

or equivalently $w \in \mathrm{span}\{\Lambda w, c\}$. By Theorem 4.8 (in Part I) any point of a supremum has the form

$$w = \beta(\alpha + \Lambda)^{-1}c. \tag{2.12}$$

Hence it remains to show that

$$(\alpha + \Lambda)^{-1}c \in \mathrm{span}\{\Lambda(\alpha + \Lambda)^{-1}c, c\},$$

or equivalently $c \in \mathrm{span}\{\Lambda c, (\alpha + \Lambda)c\} = \mathrm{span}\{\Lambda c, \alpha c\}$. Since $\alpha = 0$ only for $\gamma = 0$, see Theorem 4.10 in Part I, the last inclusion is true.

To establish the second assertion observe that w and $\nabla\lambda(w)$ are orthogonal since

$$(w, \nabla\lambda(w)) = \frac{2}{(w, \Lambda^{-1}w)}(w, w - \lambda(w)\Lambda^{-1}w) = 0.$$

We conclude that $w - \lambda\Lambda^{-1}c$ and ∇w are collinear vectors of the 2–dimensional space $\mathrm{span}\{w, \lambda\Lambda^{-1}c\}$ because $|w|^2 + |w - \lambda\Lambda^{-1}c|^2 = |\lambda\Lambda^{-1}c|^2$ defines an orthogonal decomposition (Theorem 4.3 in Part I). □

2.4. **Angle dependence of the Rayleigh quotient.**

We analyze the dependence of the Rayleigh quotient on the opening angle φ of the cone $C_\gamma(c)$ within the plane

$$P_{c,v} := \mathrm{span}\{\lambda\Lambda^{-1}c, v\}, \tag{2.13}$$

where $v = \frac{w}{|w|}$, w given by (2.12), denotes a point of a supremum of the Rayleigh quotient on the cone $C_\gamma(c)$. Now consider a parametrization of the unit circle in $P_{c,v}$ by $z(\varphi)$ in a way that $\varphi = \angle(z(\varphi), \lambda\Lambda^{-1}c)$. To make $z(\varphi)$ unique, we consider that parametrization (clockwise or anticlockwise) for which $z(\varphi^*) = v$ and $\varphi^* < \pi$ hold.

To express the angle dependence of the Rayleigh quotient in the plane $P_{c,v}$ we define

$$\lambda_{c,v}(\varphi) := \lambda(z(\varphi)). \tag{2.14}$$

Lemma 2.5. *On the assumptions of Lemma 2.4 let $v = \frac{w}{|w|}$. Then in $v = z(\varphi^*)$ holds*

$$\left|\frac{d\lambda_{c,v}}{d\varphi}(\varphi^*)\right| = |\nabla\lambda(v)|. \tag{2.15}$$

Proof. Applying the chain rule we obtain

$$\frac{d}{d\varphi}\lambda(z(\varphi)) = \nabla\lambda(z(\varphi))^T \frac{dz(\varphi)}{d\varphi}. \tag{2.16}$$

Since $|z(\varphi)| = 1$, its derivative with respect to φ is tangential to the unit circle in $P_{c,v}$, i.e. $(z(\varphi), \frac{d}{d\varphi}z(\varphi)) = 0$, while additionally $|\frac{d}{d\varphi}z(\varphi)| = 1$. Thus $\frac{dz}{d\varphi}(\varphi^*)$ and $\nabla\lambda(v)$ are collinear due to Lemma 2.4. We conclude

$$\frac{dz}{d\varphi}(\varphi^*) = \pm\frac{\nabla\lambda(v)}{|\nabla\lambda(v)|}; \tag{2.17}$$

inserting (2.17) in (2.16) for $\varphi = \varphi^*$ and recognizing (2.14) completes the proof. □

2.5. **The function** $\bar{\lambda}(c,\varphi)$.

While in the preceding section the derivative of the Rayleigh quotient with respect to the opening angle within the plane $P_{c,v}$ is determined, we consider now the derivative of the maximal Rayleigh quotient on the cone $C_\gamma(c)$ with respect to its opening angle φ.

Therefore we introduce the function $\bar{\lambda}(c,\varphi)$, which describes the maximal value of the Rayleigh quotient on the cone $C_\gamma(c)$ for $\gamma \in [0,1[$. To express the angle dependence of $C_\gamma(c)$ we also write $C_{\gamma(\varphi)}(c)$ for $\varphi \in [0, \varphi_{\max}(c)[$. Therein the maximal opening angle $\varphi_{\max}(c)$ is defined by $\varphi_{\max}(c) = \arccos(\frac{c}{|c|}, \frac{\Lambda^{-1}c}{|\Lambda^{-1}c|})$. Thus

$$\bar{\lambda}(c,\varphi) := \sup\lambda(C_{\gamma(\varphi)}(c)).$$

The next lemma shows that the derivatives $\frac{d\bar{\lambda}(c,\varphi)}{d\varphi}$ and $\frac{d\lambda_{c,v}(\varphi)}{d\varphi}$ coincide within the points of suprema.

Lemma 2.6. *Let w of the form (2.12) be a point of a supremum fulfilling the assumptions of Lemma 2.4 and $v(\alpha) = v = \frac{w}{|w|}$. Let $\varphi^* = \angle(\lambda\Lambda^{-1}c, v(\alpha))$.*

If $\alpha > 0$, i.e. in $v(\alpha)$ a supremum is attained, then

$$|\frac{d\bar{\lambda}}{d\varphi}(c,\varphi^*)| = |\frac{d\lambda_{c,v}}{d\varphi}(\varphi^*)| = |\nabla\lambda(v)|. \tag{2.18}$$

Proof. Since $\lambda(\cdot)$ is continuously differentiable in $\mathbf{R}^n \setminus \{0\}$ the function $\lambda_{c,v}(\varphi)$, see Lemma 2.5 for its definition, is continuously differentiable in φ. Furthermore, the function $\bar{\lambda}(c,\varphi)$ is continuously differentiable in φ as a consequence of the representation of the curve of suprema in the form (2.12), as derived in Section 4.3 of Part I.

For given $c \in \mathbf{R}^n$ and $\gamma \in]0,1[$ let $v(\alpha)$ be a point of a supremum of the Rayleigh quotient on $C_\gamma(c)$ with $\varphi^* = \angle(\lambda\Lambda^{-1}c, v(\alpha))$. Then by definition of $\bar{\lambda}(c,\varphi)$ for $\varphi \in [0, \varphi_{\max}(c)[$ is a dominating function of $\lambda_{c,v}(\varphi)$ so that

$$\lambda_{c,v}(\varphi) \le \bar{\lambda}(c,\varphi) \quad \text{and} \quad \lambda_{c,v}(\varphi^*) = \bar{\lambda}(c,\varphi^*). \tag{2.19}$$

The last equation results from the fact that $v(\alpha)$ lies in $P_{c,v}$. Since from (2.19) the positive function $\bar{\lambda}(c,\varphi) - \lambda_{c,v}(\varphi)$ takes its minimum in $\varphi = \varphi^*$, we conclude for the derivative

$$\frac{d\lambda_{c,v}}{d\varphi}(\varphi^*) = \frac{d\bar{\lambda}}{d\varphi}(c,\varphi^*). \tag{2.20}$$

From Equation (2.15) in Lemma 2.5 we obtain

$$|\frac{d\bar{\lambda}}{d\varphi}(c,\varphi^*)| = |\frac{d\lambda_{c,v}}{d\varphi}(\varphi^*)| = |\nabla\lambda(v)|.$$

□

Note that by Theorem 2.1 the derivative (2.18) on the level set $L(\lambda(v))$ is bounded from above by $v \in L(\lambda(v))$ of the form $v_{1,n} = (v_1, 0, \ldots, 0, v_n)^T$. This fact is of central importance for the convergence theorem in the next section.

3. Convergence estimates for PINVIT

3.1. **A convergence theorem on the Rayleigh quotient.** We are now in a position to formulate the main convergence estimates for the method of preconditioned inverse iteration by combining the various extremal properties proved in Section 2. Hence in this section our efforts for a convergence theory of PINVIT come to a close.

We present a sharp estimate from above for the Rayleigh quotient $\lambda(x')$ of the new iterate describing the case of poorest convergence. This estimate only depends on the Rayleigh quotient λ, on the spectral radius γ of the error propagation matrix $I - B^{-1}A$ and on the eigenvalues of A. Sharpness of the estimates means that for each λ and γ an initial vector as well as a preconditioner can be constructed so that the estimate is attained.

While Theorem 1.1 in Section 1 is stated in terms of the original basis, we here continue using the c–basis representation. It is important to note that the estimates for the Rayleigh quotient do not depend on the choice of the basis. We further note that the complex form of the upper bound (1.5) derives from Theorem 5.1 in Part I by inserting various quantities, each of which has a simple geometric interpretation, into the Rayleigh quotient.

Finally we note that the proof of Theorem 1.1 only has to treat the case of simple eigenvalues due to the analysis given in Section 3 of Part I. Later in Section 3.3 it is shown that the convergence estimates remain to be sharp for matrices with eigenvalues of arbitrary multiplicity.

Proof of Theorem 1.1. In terms of the c–basis the first assertion follows from the fact that PINVIT is stationary within each nonzero multiple of the unit vectors e_i, $i = 1, \ldots, n$.

To prove the second assertion we consider the case $\lambda_i < \lambda < \lambda_{i+1}$ and show that λ' takes its maximal value under all $c \in L(\lambda)$ and all admissible preconditioners in the vector

$$c_{i,i+1} = (0, \ldots, 0, c_i, c_{i+1}, 0, \ldots, 0)^T,$$

with $\lambda(c_{i,i+1}) = \lambda$. Moreover we have to show that the maximal value λ' is defined by Equation (1.5).

By Lemmata 2.2 and 2.3 the opening angle φ_γ and the shrinking angle ψ_γ take their minima on $L(\lambda)$ in $c_{i,i+1}$ so that

$$\varphi_\gamma(c_{i,i+1}) \leq \varphi_\gamma(c) \quad \text{and} \quad \psi_\gamma(c_{i,i+1}) \leq \psi_\gamma(c).$$

Now consider the normed curve of the points of suprema on $C_\gamma(c)$ for $\gamma \in [0,1]$, which is given by $\frac{(\alpha+\Lambda)^{-1}c}{|(\alpha+\Lambda)^{-1}c|}$ for $\alpha \in [0, \infty[$. The curve for $c_{i,i+1}$ is defined analogously. Let v ($v_{i,i+1}$) be two points on the curve defined by c ($c_{i,i+1}$) in a way that $\lambda(v) = \lambda(v_{i,i+1})$. The angles enclosed with the center of the cones are denoted by

$$\varphi^* = \angle(v, \lambda\Lambda^{-1}c) \quad \text{and} \quad \varphi^*_{i,i+1} = \angle(v_{i,i+1}, \lambda\Lambda^{-1}c_{i,i+1}).$$

For the derivatives of $\bar{\lambda}(c, \varphi)$ and $\bar{\lambda}(c_{i,i+1}, \varphi)$ with respect to φ from Lemma 2.6 and Theorem 2.1 follows

$$\frac{d\bar{\lambda}}{d\varphi}(c_{i,i+1}, \varphi^*_{i,i+1}) \leq \frac{d\bar{\lambda}}{d\varphi}(c, \varphi^*). \tag{3.1}$$

Note that $\bar{\lambda}(c, \cdot)$ and $\bar{\lambda}(c_{i,i+1}, \cdot)$ are monotone increasing positive functions in φ. So let us set $f(\varphi) = \bar{\lambda}(c_{i,i+1}, \varphi)$ and $g(\varphi) = \bar{\lambda}(c, \varphi)$. Furthermore, let us denote the opening angles of $C_1(c_{i,i+1})$ and $C_1(c)$ by $a = \varphi_1(c_{i,i+1})$ and $b = \varphi_1(c)$.

So we have the situation that $f, g : [0, b] \to \mathbf{R}$ for $b > 0$ are strictly monotone increasing functions and that $f(a) = g(b)$. Further by Equation (3.1) it holds that for $\alpha,\ \beta \in [0, b]$

with $f(\alpha) = g(\beta)$ the derivatives satisfy

$$f'(\alpha) \leq g'(\beta). \tag{3.2}$$

So it is easy to show by integrating the inverse functions $(g^{-1})'(y)$ and $(f^{-1})'(y)$ that for any $\xi \in [0, a]$

$$f(a - \xi) \geq g(b - \xi).$$

Hence, we have for ξ equal to the smaller shrinking angle $\psi_\gamma(c_{i,i+1}) \in [0, \varphi_1(c_{i,i+1})]$

$$\bar{\lambda}(c_{i,i+1}, \varphi_\gamma(c_{i,i+1})) \geq \bar{\lambda}(c, \varphi_1(c) - \psi_\gamma(c_{i,i+1})),$$

where $\varphi_\gamma(c_{i,i+1}) = \varphi_1(c_{i,i+1}) - \psi_\gamma(c_{i,i+1})$.

If c is not of the form $c_{i,i+1}$ then Theorem 2.1 implies $|\nabla\lambda(c)| > |\nabla\lambda(c_{i,i+1})|$. Thus we have $\psi_\gamma(c) > \psi_\gamma(c_{i,i+1})$ from Lemma 2.3 so that

$$\bar{\lambda}(c, \varphi_1(c) - \psi_\gamma(c_{i,i+1})) > \bar{\lambda}(c, \varphi_1(c) - \psi_\gamma(c)) = \bar{\lambda}(c, \varphi_\gamma(c)),$$

which establishes $c_{i,i+1}$ as the vector of poorest convergence, i.e.

$$\bar{\lambda}(c_{i,i+1}, \varphi_\gamma(c_{i,i+1})) > \bar{\lambda}(c, \varphi_\gamma(c)).$$

Thus $c_{i,i+1}$ leads to the largest rate of convergence. Equation (1.5) derives from the fact that all points of suprema of the cones $C_\gamma(c_{i,i+1})$ remain in the plane $\mathrm{span}\{e_i, e_{i+1}\}$ due to the analysis of Part I. Thus by Lemma 4.11 in Part I all zero components of $c_{i,i+1}$ can be removed and the mini–dimensional analysis of Part I within the plane $\mathrm{span}\{e_i, e_{i+1}\}$ can be applied. Equation (1.5) results from Equation (5.4) in Theorem 5.1 (in Part I) where λ_1 and λ_2 are substituted by λ_i and λ_j.

To show that $\Phi_{i,i+1}(\lambda, \gamma) < 1$, we adapt the classical analysis of D'yakonov and Orekhov [2] and obtain the upper bound

$$\Phi_{i,i+1}(\lambda, \gamma) \leq \frac{1 - (1-\gamma)^2 \frac{\lambda_{i+1} - \lambda}{\lambda_{i+1}}}{1 + (1-\gamma)^2 \frac{(\lambda - \lambda_i)(\lambda_{i+1} - \lambda)}{\lambda_i \lambda_{i+1}}} < 1, \tag{3.3}$$

for $\lambda \in]\lambda_i, \lambda_{i+1}[$. Finally, for any sequence of iterates

$$(x^{(j)}, \lambda^{(j)}), \qquad j = 0, 1, 2, \ldots$$

by (1.6) the sequence of Rayleigh quotients $\lambda^{(j)}$ decreases monotonically and is bounded from below by λ_1. Its limes is an eigenvalue of A since PINVIT is stationary if and only if it is applied to some eigenvector of A, for which the residual $Ax - \lambda(x)x$ vanishes. Additionally, for all λ not equal to an eigenvalue of A, the estimate (3.3) gives a simple bound from above for the decrease of the Rayleigh quotient. □

3.2. **Convergence of eigenvector approximations.**

So far we have not given estimates on the convergence of the eigenvector approximations generated by PINVIT. The main reason for this reticence is that the sequence of the acute angles between the first eigenvector e_1 and the iterates of PINVIT is in general not monotone decreasing. To see this, let $\chi(d)$ be the acute angle between the vector d and e_1. Furthermore, for given $c \in \mathbf{R}^n$, with $|c| = 1$ and $c_1 > 0$, define the cone M by

$$M = \{d \in \mathbf{R}^n;\ \chi(d) \leq \chi(c)\}.$$

Then $c - \lambda\Lambda^{-1}c$ is the normal vector on the $n-1$ dimensional tangent plane of $E_1(c)$ in c. Furthermore, $c - \frac{|c|^2}{c_1}e_1$ is normal to the tangent plane of M in c. A necessary condition for $E_1(c) \subset M$ is that the normal vectors are collinear. In this case the acute angle between any new iterate and e_1 is less or equal to $\chi(c)$. Otherwise we have $E_1(c) \not\subset M$, so that a preconditioner with γ near to 1 can be constructed so that the considered angle is increased.

It is easy to see that in the $\mathbf{R}^2$ the normal vectors are collinear. But in the $\mathbf{R}^n$ for any c with $c_i \neq 0$ for pairwise different indexes $i = 1,\ k,\ l$ we obtain from

$$c - \frac{|c|^2}{c_1} e_1 = \omega(c - \lambda\Lambda^{-1}c)$$

both $\omega = \frac{\lambda_k}{\lambda_k - \lambda}$ and $\omega = \frac{\lambda_l}{\lambda_l - \lambda}$. Since $\lambda_k \neq \lambda_l$ we conclude the normal vectors are not collinear.

Nevertheless, Theorem (2.1) by taking the maximum of Equation (2.6) provides an upper bound for the Euclidean norm of the gradient by

$$|\nabla\lambda(c)|^2 \leq \frac{4\lambda^2(\lambda - \lambda_1)(\lambda_n - \lambda)}{\lambda_1\lambda_n}, \tag{3.4}$$

which is a variant of Temple's inequality; cf. Chapter III in [6].

From (3.4) we derive a simple bound depending on $\lambda - \lambda_1$ for the residual of the actual iterate.

Corollary 3.1. *Let $c \in \mathbf{R}^n$, $|c| = 1$ and $\lambda = \lambda(c)$. Then*

$$|(I - \lambda\Lambda^{-1})c| \leq \left(\frac{\lambda}{\lambda_1} - 1\right)^{1/2}$$

Proof. Combining (2.2) and (3.4) we have

$$|(I - \lambda\Lambda^{-1})c|^2 \leq \frac{(\lambda - \lambda_1)(\lambda_n - \lambda)}{\lambda_1\lambda_n}$$

From $1 - \frac{\lambda}{\lambda_n} \leq 1$ the proposition follows. □

3.3. **Convergence estimates for matrices with multiple eigenvalues.**

In Section 3 of Part I the convergence analysis of PINVIT for matrices with multiple eigenvalues is traced back to a reduced problem with only simple eigenvalues. By using the notation of that section it is shown that

$$\sup \lambda(E_\gamma(c)) \leq \sup \bar{\lambda}(E_\gamma(\bar{c})), \tag{3.5}$$

where the bar denotes those quantities which are associated with simple eigenvalues, see Part I. Therefore, the question arises whether or not PINVIT converges more rapdidly in the case of multiple eigenvalues. The next lemma shows that the worst–case–estimate of Theorem 1.1 is also sharp for matrices with eigenvalues of arbitrary multiplicity.

Lemma 3.2. *Adopting the notation of Lemma 3.1 in Part I for any $c \in \mathbf{R}^m$ with $\bar{c} = Pc$ it holds that*

$$\sup \lambda(E_\gamma(c)) = \sup \bar{\lambda}(E_\gamma(\bar{c}))$$

Proof. Due to Theorem 4.10 in Part I for any $\bar{w} \in \arg\sup \bar{\lambda}(E_\gamma(\bar{c}))$ there are unique constants $\alpha,\ \beta \in \mathbf{R}$, so that

$$\bar{w} = \beta(\alpha + \bar{\Lambda})^{-1}\bar{c}.$$

Since $\bar{w} \in \partial E_\gamma(\bar{c})$, for its distance to the ball's center $\lambda\bar{\Lambda}^{-1}\bar{c}$ holds

$$\begin{aligned} |\bar{w} - \lambda\bar{\Lambda}^{-1}\bar{c}|^2 &= \sum_{i=1}^{n} \left(\beta(\alpha+\lambda_i)^{-1}\bar{c}_i - \lambda\lambda_i^{-1}\bar{c}_i\right)^2 \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m(i)} \left(\beta(\alpha+\lambda_i)^{-1}c_{i;j} - \lambda\lambda_i^{-1}c_{i;j}\right)^2 = |w - \lambda\Lambda^{-1}c|^2, \end{aligned}$$

where $w := \beta(\alpha+\Lambda)^{-1}c \in \mathbf{R}^m$. Hence, $w \in E_\gamma(c)$, since by Lemma 3.1 in Part I we have $|c - \lambda\Lambda^{-1}c| = |\bar{c} - \lambda\bar{\Lambda}^{-1}c|$ and $\bar{w} = Pw$. Additionally, it holds $\lambda(w) = \bar{\lambda}(Pw) = \bar{\lambda}(\bar{w})$. We conclude by using (3.5) that

$$\lambda(w) \le \sup \lambda(E_\gamma(c)) \le \sup \bar{\lambda}(E_\gamma(\bar{c})) = \bar{\lambda}(\bar{w}),$$

which establishes the proposition. □

4. Conclusion

A new theoretical framework for preconditioned gradient methods for the eigenvalue problem has been developed in a way that these methods are traced back to preconditioned inverse iteration. PINVIT, which derives from the well–known inverse iteration method where the associated system of linear equations is approximately solved through preconditioning, turns out to be an *efficient* and convergent algorithm for the iterative solution of *mesh* eigenproblems defined by self–adjoint and coercive elliptic differential operators, see [5] for more practical questions.

A sharp convergence estimate for the eigenvalue approximations has been derived which does not depend on the largest eigenvalue or on the mesh size. The convergence theory of PINVIT is the basis for the analysis of a preconditioned subspace iteration [4], which is the direct generalization of the subspace implementation of inverse iteration [1].

Since the given analysis provides a clear understanding of preconditioned inverse iteration and of its underlying geometry, the question, "Preconditioned eigensolvers –an oxymoron?", which is the title of a recent review article [3], definitely has a negative answer.

Appendix A. Exclusion of points of suprema in the unit vectors

By using the results of Section 2, we show that suprema of the Rayleigh quotient on $C_\gamma(c)$ are not attained in vectors of the form $z = \theta e_k$ (with $2 \le k \le n$, $\theta \ne 0$). Those points of suprema result from the analysis of Part I (see Lemma 4.5) from a Lagrange multiplier ansatz assuming one Lagrange multiplier to be equal to zero. To make the analysis of Part I complete we now exclude these points.

Lemma A.1. *Consider a nonnegative vector $c \in \mathbf{R}^n$, $|c| = 1$ and to avoid stationarity of PINVIT let $c \ne e_i$ for $i = 1, \dots, n$. Then $w \in \arg\sup \lambda(E_\gamma(c))$ is of the form $w = \beta(\alpha+\Lambda)^{-1}c$ for unique constants $\alpha,\ \beta \in \mathbf{R}$, so that points of suprema of the form θe_k are impossible.*

Proof. Assume $z \in \arg\sup \lambda(E_\gamma(c))$ to be of the form $z = \theta e_k$ with $2 \le k \le n$. Then $\lambda = \lambda(z) = \lambda_k$. Furthermore, due to Lemma 4.9 in Part I there is a $w = \beta(\alpha+\Lambda)^{-1}c \in E_\gamma(c)$ with $\lambda(w) = \lambda(z)$. Hence, $v = \frac{w}{|w|}$ has at least two nonzero components and

$$|\frac{d\bar{\lambda}}{d\varphi}[v]| = |\nabla\lambda(v)| \ne 0.$$

For the angle derivative (in the sense of Section 2.4) within the plane spanned by $z = \theta e_k$ and $\lambda\Lambda^{-1}c$ we have

$$|\frac{d\lambda}{d\varphi}[z]| = \nabla\lambda(z)^T \frac{dz}{d\varphi} = 0$$

since $\nabla\lambda(z) = 0$. Since $w \neq z$, there are disjoint ϵ–neighborhoods $U_\epsilon(z)$ and $U_\epsilon(w)$. Now decrease the opening angle φ of $C_\gamma(c)$ to $\varphi - \delta\varphi$ defining a new cone $C_{\gamma'}(c)$. Since by $w = \beta(\alpha + \Lambda)^{-1}c$ a continuous curve of points of extrema is defined, take the increment $\delta\varphi$ so small that $\tilde{w} \in \arg\sup E_{\gamma'}(c)$ is located in $U_\epsilon(w)$. If $\delta\varphi$ is sufficiently small, then because of

$$0 = |\frac{d\lambda}{d\varphi}[z]| < |\frac{d\bar{\lambda}}{d\varphi}[v]|$$

there is a further point of a supremum $\tilde{z} \in U_\epsilon(z)$ with $\lambda(\tilde{z}) = \lambda(\tilde{w})$. For $\delta\varphi < \pi/2$ the Rayleigh quotient $\lambda(\tilde{z})$ is not equal to any of the eigenvalues of A and hence $\tilde{z}$ is not collinear to any of the unit vectors. Furthermore, $\tilde{z}$ is not of the form $\tilde{\beta}(\tilde{\alpha} + \Lambda)^{-1}c$, since for a given Rayleigh quotient these points of suprema are unique. Hence, such $\tilde{z}$ does not satisfy the necessary conditions of the Lagrange multiplier ansatz of Part I so that $z = \theta e_k$ is not a point of a supremum. $\square$

REFERENCES

[1] J.H. Bramble, J.E. Pasciak, and A.V. Knyazev. A subspace preconditioning algorithm for eigenvector/eigenvalue computation. *Adv. Comput. Math.*, 6:159–189, 1996.

[2] E.G. D'yakonov and M.Y. Orekhov. Minimization of the computational labor in determining the first eigenvalues of differential operators. *Math. Notes*, 27:382–391, 1980.

[3] A.V. Knyazev. Preconditioned eigensolvers -an oxymoron? . *Electron. Trans. Numer. Anal.*, 7:104–123, 1998.

[4] K. Neymeyr. A geometric theory for preconditioned inverse iteration applied to a subspace. Submitted to Math. Comput., 1999.

[5] K. Neymeyr. A posteriori error estimation for elliptic eigenproblems. Submitted to Math. Comput., 1999.

[6] Y. Saad. *Numerical methods for large eigenvalue problems*. Manchester University Press, Manchester, 1992.

MATHEMATISCHES INSTITUT DER UNIVERSITÄT TÜBINGEN, AUF DER MORGENSTELLE 10, 72076 TÜBINGEN, GERMANY.

E-mail address: neymeyr@na.uni-tuebingen.de