CONVERGENCE ANALYSIS OF PROJECTION METHODS FOR THE NUMERICAL SOLUTION OF LARGE LYAPUNOV EQUATIONS

V. SIMONCINI† AND V. DRUSKIN ‡

Abstract. The numerical solution of large-scale continuous-time Lyapunov matrix equations is of great importance in many application areas. Assuming that the coefficient matrix is positive definite, but not necessarily symmetric, in this paper we analyze the convergence of projection type methods for approximating the solution matrix. Under suitable hypotheses on the coefficient matrix, we provide new asymptotic estimates for the error matrix when a Galerkin method is used in a Krylov subspace, Numerical experiments confirm the good behavior of our upper bounds when linear convergence of the solver is observed.

1. The problem. We are interested in the approximate solution of the following Lyapunov matrix equation:

\[ AX + XA^\top = BB^\top, \]

with \( A \) a real matrix of large dimension and \( B \) a real tall matrix. Here \( A^\top \) indicates the transpose of \( A \). We assume that the \( n \times n \) matrix \( A \) is either symmetric and positive definite, or nonsymmetric with positive definite symmetric part, that is, \((A + A^\top)/2\) is positive definite. In the following we mostly deal with the case of \( B \) having a single column, that is \( B = b \) and we assume that \( b \) has unit Euclidean norm, that is \( \|b\| = 1 \). Nonetheless, our results can be extended to the multiple vector case.

This problem arises in a large variety of applications, such as signal processing and system and control theory. The symmetric solution \( X \) carries important information on the stability and energy of an associated dynamical linear system and on the feasibility of order reduction techniques [1], [3], [5]. The analytic solution of (1.1) can be written as

\[ X = \int_{0}^{\infty} e^{-tA}BB^\top e^{-tA^\top} dt = \int_{0}^{\infty} xx^\top dt, \]

where we have set \( x = e^{-tA}B \). Let \( \alpha_{\text{min}} \) be the smallest eigenvalue of the symmetric part of \( A \), \( \alpha_{\text{min}} = \lambda_{\text{min}}((A + A^\top)/2) > 0 \). Then it can be shown that \( \|x\| \leq \exp(-t\alpha_{\text{min}})\|B\| \); see, e.g., [5].

Projection type methods seek an approximate solution \( X_m \) in a subspace of \( \mathbb{R}^n \), by requiring that the residual \( BB^\top - (AX_m + X_mA^\top) \) be orthogonal to this subspace. A particularly effective choice as approximation space is given by (here for \( B = b \)) the Krylov subspace \( K_m(A, b) = \text{span}\{b, Ab, \ldots, A^{m-1}b\} \), of dimension \( m \leq n \) [16], [17], [26]. Abundant experimental evidence over the years has shown that the use of this space allows one to often obtain a satisfactorily accurate approximation \( X_m \) in a space of much lower dimension than \( n \). A particularly attractive feature is that \( X_m \) may be written as a low rank matrix, \( X_m = U_mU_m^\top \) with \( U_m \) of low column rank, so that only the matrix \( U_m \) needs to be stored.

†Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato 5, I-40127 Bologna, Italy and CIRSA, Ravenna, Italy (valeria@dm.unibo.it).
‡Scientific Advisor, Ph.D., 1 Hampshire st. Cambridge, MA 02139 USA (druskin1@boston.oilfield.slb.com)
To the best of our knowledge, no asymptotic convergence analysis of this Galerkin method is available in literature. The aim of this paper is to fill this gap. We also refer to [25] for a-priori estimates on the residual norm when solving the Sylvester equation with projection type methods. To provide our error estimates, we shall use the integral representation (1.2) for both $X$ and $X_m$, and explicitly bound the norm of the error matrix $X - X_m$; we refer to [26] for early considerations in this direction. Our approach is highly inspired by, and fully relies on, the papers [8], [18], where general estimates for the error in approximating matrix operators by polynomial methods are derived.

We provide explicit estimates when $A$ is symmetric, and when $A$ is nonsymmetric with its field of values (or spectrum) contained in certain not necessarily convex sets of $\mathbb{C}^+$. We are interested in finding a-priori bounds for the 2-norm of the error matrix, that is $\|X - X_m\|$, where the 2-norm is the matrix norm induced by the vector Euclidean norm. We start by observing that

\[
K_m(A, b) \subset \mathbb{R}^m, \quad \text{where the 2-norm is the matrix norm induced by the vector Euclidean norm.}
\]

Our estimates are asymptotic, and thus linear, that is, they do not capture the possibly superlinear convergence behavior of the method that is sometimes observed [24]. In the linear system setting, the superlinear behavior is due to the fact that Krylov based methods tend to adapt to the (discrete) spectrum of $A$; accelerating convergence as spectral information is gained while enlarging the space; see, e.g., [29] for a discussion and more references.

Throughout the paper we assume exact arithmetic.

2. Numerical solution and preliminary considerations. Given the Krylov subspace $K_m(A, b)$ and a matrix $V_m$ whose orthonormal columns span $K_m(A, b)$, with $b = V e_1$, we seek an approximation in the form $X_m = V_m Y_m V_m^\top$. Here and in the following, $e_i$ denotes the $i$th column of the identity matrix of given dimension. Imposing that the residual $R_m = b b^\top - (A X_m + X_m A^\top)$ be orthogonal to the given space, the Galerkin condition, yields the equation

\[
V_m^\top R_m V_m = 0 \quad \Leftrightarrow \quad T_m Y + Y T_m^\top = e_1 e_1^\top,
\]

where $T_m = V_m^\top A V_m$; see, e.g., [1], [26]. The $m \times m$ matrix $Y_m$ can thus be computed by solving the resulting small size Lyapunov equation.

The matrix $X_m$ can be equivalently written in integral form. Indeed, let $x_m = x_m(t) = V_m e^{-t T_m} e_1$ be the so-called Krylov approximation to $x = x(t)$ in $K_m(A, b)$. Then $X_m$ can be written as

\[
X_m = V_m \left( \int_0^\infty e^{-t T_m} e_1 e_1^\top e^{-t T_m^\top} dt \right) V_m^\top
\]

\[
= \int_0^\infty V_m e^{-t T_m} e_1 e_1^\top e^{-t T_m^\top} V_m^\top dt = \int_0^\infty x_m(x_m)^\top dt.
\]

We are interested in finding a-priori bounds for the 2-norm of the error matrix, that is for $\|X - X_m\|$, where the 2-norm is the matrix norm induced by the vector Euclidean norm. We start by observing that $\|X - X_m\| = \| \int_0^\infty (x x^\top - x_m x_m^\top) dt \|$, and that

\[
\|x x^\top - x_m x_m^\top\| = \|x(x - x_m)^\top - (x - x_m) x_m^\top\| \leq (\|x\| + \|x_m\|) \|x - x_m\|.
\]

It holds that $\lambda_{\min}((T_m + T_m^\top)/2) \geq \alpha_{\min}$. Using $\|x_m\| \leq \exp(-t \alpha_{\min}((T_m + T_m^\top)/2)) \leq \exp(-t \alpha_{\min})$, we have

\[
\|X - X_m\| \leq \int_0^\infty \|x x^\top - x_m x_m^\top\| dt \leq \int_0^\infty (\|x\| + \|x_m\|) \|x - x_m\| dt
\]

\[
\leq 2 \int_0^\infty e^{-t \alpha_{\min}} \|x - x_m\| dt.
\]

(2.1)
We notice that
\[
e^{-t \alpha_{\min}} \|x - x_m\| = \| \exp(-t(A + \alpha_{\min}I))b - V_m \exp(-t(T_m + \alpha_{\min}I))e_1 \|
\]
which is the error in the approximation of the exponential of the *shifted* matrix \(A + \alpha_{\min}I\) with the Krylov subspace solution. Therefore,
\[
\|X - X_m\| \leq 2 \int_0^\infty \| \hat{x} - \hat{x}_m \| dt. \quad (2.2)
\]
In the following we will bound \(\|X - X_m\|\) by judiciously integrating an upper bound of the integrand function.

The matrix \(V_m = [v_1, \ldots, v_m]\) can be generated one vector at the time, by means of the following Arnoldi recursion:
\[
AV_m = V_m T_m + v_{m+1} t_{m+1,m} e_m^\top, \quad v_1 = b/\|b\|, \quad (2.3)
\]
where \(V_{m+1} = [V_m, v_{m+1}]\) has orthonormal columns and spans \(K_{m+1}(A, b)\). In general, \(T_m\) is upper Hessenberg, and it is symmetric, and thus tridiagonal, when \(A\) is itself symmetric.

We conclude this section with a technical lemma, whose proof is included for completeness; see, e.g., [18] for a similar result in finite precision arithmetic. More details on the admissibility of series expansions in terms of Faber polynomials are given in section 4.

**Lemma 2.1.** Let \(\Phi_k\) be a Faber polynomial of degree at most \(k\). Let \(f(z) = \sum_k f_k \Phi_k(z)\) be a series expansion of the analytic function \(f\) and assume that the expansion of \(f(A)\) and \(f(T_m)\) is well defined. Then
\[
\|f(A)b - V_m f(T_m)e_1\| \leq \sum_{k=m}^\infty |f_k| (\|\Phi_k(A)\| + \|\Phi_k(T_m)\|).
\]

**Proof.** We have
\[
f(A)b - V_m f(T_m)e_1 = \sum_{k=0}^{m-1} f_k(\Phi_k(A))b - V_m \Phi_k(T_m)e_1 + \sum_{k=m}^\infty f_k(\Phi_k(A))b - V_m \Phi_k(T_m)e_1.
\]
Using the Arnoldi relation and the fact that \(T_m\) is upper Hessenberg, \(A^k V_m e_1 = V_m T_m^k e_1\) for \(k = 1, \ldots, m - 1\), and thus \(\Phi_k(A) = \Phi_k(T_m)\) and \(V_m \Phi_k(T_m) e_1 = V_m \Phi_k(T_m) e_1\), \(k = 1, \ldots, m - 1\), so that
\[
f(A)b - V_m f(T_m)e_1 = \sum_{k=m}^\infty f_k(\Phi_k(A))b - V_m \Phi_k(T_m)e_1.
\]
Taking norms, the result follows. \(\Box\)
3. The symmetric case. In the symmetric case, we show that the asymptotic convergence rate of the Krylov subspace solver is the same as that of the Conjugate Gradient method applied to the shifted system \((A + \alpha_{\text{min}} I)x = b\), where \(\alpha_{\text{min}} = \lambda_{\text{min}}\), the smallest eigenvalue of the positive definite matrix \(A\) [12]; see also section 5.

**Proposition 3.1.** Let \(A\) be symmetric and positive definite, and let \(\lambda_{\text{min}}\) be the smallest eigenvalue of \(A\). Let \(\lambda_{\text{min}}, \lambda_{\text{max}}\), be the extreme eigenvalues of \(A + \lambda_{\text{min}} I\) and \(\hat{\kappa} = \lambda_{\text{max}}/\lambda_{\text{min}}\). Then

\[
\|X - X_m\| \leq \frac{\sqrt{\hat{\kappa} + 1}}{\lambda_{\text{min}} \sqrt{\hat{\kappa}}} \left(\frac{\sqrt{\hat{\kappa} - 1}}{\sqrt{\hat{\kappa} + 1}}\right)^m. \tag{3.1}
\]

**Proof.** Using (2.1) we are left to estimate \(\int_0^\infty e^{-\alpha_{\text{min}}} \|x - x_m\| dt\). Let \(\lambda_{\text{max}}\) be the largest eigenvalue of \(A\). From [7] it follows that

\[
\|x - x_m\| \leq 4 \exp(-t \frac{\lambda_{\text{max}} + \lambda_{\text{min}}}{2}) \sum_{k=m}^\infty I_k(t \frac{\lambda_{\text{max}} - \lambda_{\text{min}}}{2}),
\]

where \(I_k\) is the Bessel function of an imaginary argument. Therefore, setting \(p = (3\lambda_{\text{min}} + \lambda_{\text{max}})/(\lambda_{\text{max}} - \lambda_{\text{min}}) = (\hat{\kappa} + 1)/(\hat{\kappa} - 1)\) and \(\rho = p + \sqrt{p^2 - 1}\), we have

\[
\|X - X_m\| \leq 4 \int_0^\infty \|\hat{x} - \hat{x}_m\| dt
\]

\[
\leq 8 \sum_{k=m}^\infty \int_0^\infty \exp(-t(\frac{3}{2} \lambda_{\text{min}} + \frac{1}{2} \lambda_{\text{max}})) I_k(t \frac{\lambda_{\text{max}} - \lambda_{\text{min}}}{2}) dt
\]

\[
= \frac{8}{\sqrt{(\frac{3}{2} \lambda_{\text{min}} + \frac{1}{2} \lambda_{\text{max}})^2 - (\lambda_{\text{max}} - \lambda_{\text{min}})^2}} \sum_{k=m}^\infty \frac{1}{(p + \sqrt{p^2 - 1})^k} \tag{3.2}
\]

\[
= \frac{8(\hat{\kappa} + 1)}{\sqrt{\hat{\kappa}(3\lambda_{\text{min}} + \lambda_{\text{max}})}} \sum_{k=m}^\infty \frac{1}{p + \sqrt{p^2 - 1})^k}
\]

\[
= \frac{4(\hat{\kappa} + 1)}{\sqrt{\hat{\kappa}(3\lambda_{\text{min}} + \lambda_{\text{max}})}} \frac{2\rho - 1}{\rho - 1} \rho^m.
\]

To get (3.2) we used the following integral formula for Bessel functions in [13, Formula (6.611.4)]

\[
\int_0^\infty e^{-\alpha t} I_{\nu}(\beta t) dt = \frac{\beta^\nu}{\sqrt{\alpha^2 - \beta^2}(\alpha + \sqrt{\alpha^2 - \beta^2})^\nu}, \quad \text{for } \Re \nu > -1 \text{ and } \Re \alpha > |\Re \beta|.
\]

Standard algebraic manipulations give

\[
\rho = \frac{\sqrt{\hat{\kappa} + 1}}{\sqrt{\hat{\kappa} - 1}}, \quad \frac{2\rho}{\rho - 1} = \sqrt{\hat{\kappa} + 1}.
\]

In Figure 3.1 we report the behavior of the bound of Proposition 3.1 for a 400 x 400 diagonal matrix \(A\) having uniformly distributed eigenvalues between 1 and 10. Here \(\alpha_{\text{min}} = \lambda_{\text{min}} = 1\). The vector \(b\) is the normalized vector of all ones.

We explicitly observe that the linearity of the convergence rate is exactly reproduced by the upper bound of Proposition 3.1.
4. The nonsymmetric case. For $A$ nonsymmetric, the result of the previous section can be generalized whenever the field of values of $A$ is contained in a “well-behaved” set of the complex plane.

The following results make use of the theory of Faber polynomials, and of recently obtained results that have been used in the context of linear systems. To this end, we need some definitions on conformal mappings. Let $\mathbb{C} = \mathbb{C} \cup \{\infty\}$, and let $C(0, \varrho) = \{ |\tau| \leq \varrho \}$ be the closed disk centered at zero and radius $\varrho$; define

$$\psi : \mathbb{C} \setminus C(0,\varrho) \to \overline{\mathbb{C}}$$

a conformal mapping that maps $\mathbb{C} \setminus C(0,\varrho)$ onto a simply connected domain, and such that $\psi(\infty) = \infty$ and $\psi'(\infty) > 0$; see, e.g., [30]. Let $\phi$ denote the inverse of $\psi$. The principal (polynomial) part of the Laurent series of $\phi^k$ is the Faber polynomial $\Phi_k$, of exact degree $k$. Let $\Omega_\rho$ be a bounded set of the complex plane that is the image through $\psi$ of the domain $\mathbb{C} \setminus C(0,\varrho)$, such that its complement is simply connected. Let

$$f(\lambda) \equiv \exp(-\lambda t) = \sum_{k=0}^{\infty} f_k \Phi_k(\lambda),$$

be the expansion of $\exp(-\lambda t)$ in Faber series in $\Omega_\rho$ with $\varrho < r_2 < \infty$. For $\varrho < \rho < r_2$, the expansion coefficients are given as

$$f_k = \frac{1}{2\pi i} \int_{|\tau| = \rho} \frac{\exp(-t\psi(\tau))}{\tau^{k+1}} d\tau, \quad |f_k| \leq \frac{1}{\rho^k} \sup_{|\tau| = \rho} |\exp(-t\psi(\tau))|,$$

see, e.g., [30, sec. 2.1.3],[31]. Note that $f_k = f_k(t)$.

We start by recalling the following result in [30, Theorem 2(4), sec. 2.1.3, p.142].

**Theorem 4.1.** Let $\Omega$ be a bounded set such that its complement is simply connected. Let $f$ be regular in $\psi(C(0,\rho))$, $\varrho < r_2 < \infty$, and

$$\frac{1}{2\pi} \int_{0}^{2\pi} |f(\psi(re^{i\theta}))| d\theta \leq M.$$
for all \( r, \rho < r < r_2 \), where \( \mathcal{M} \) is a constant. If the closure \( \Omega \) is such that \( |\Phi_k(z)| \leq ag^k \), \( z \in \Omega \), \( k = 0, 1, \ldots \), where \( a \) is a constant, then for \( r \) such that \( \rho \leq r < r_2 \),

\[
|f(z) - \sum_{k=0}^{m} f_k \Phi_k(z)| \leq \sum_{k=m}^{\infty} |f_k| |\Phi_k(z)| \leq \mathcal{M} a - \frac{r_2}{r_2 - r} \left( \frac{r}{r_2} \right)^m, \quad z \in \psi(C(0, r)).
\]

We next discuss a few known examples of computable mappings \( \psi \) for which Theorem 4.1 can be effectively applied. We first consider the case when \( \Omega \) coincides with an ellipse. In such setting, Faber polynomials correspond to scaled Chebyshev polynomials. Moreover, it has been recently shown that their norm is bounded by 2 when \( \Omega \) is convex [2], allowing us to derive a simple error estimate. We recall that the field of values of a real matrix \( A \) in the Euclidean inner product is defined as \( F(A) = \{ x^T Ax, x \in \mathbb{C}^n, \|x\| = 1 \} \). The location of the field of values plays a crucial role in the behavior and analysis of polynomial type methods for the solution of linear systems; see, e.g., [10], [22].

**Corollary 4.2.** Assume the field of values of the real matrix \( A \) is contained in an ellipse \( E \subset \mathbb{C}^+ \) of center \( (c, 0) \), foci \( (c \pm d, 0) \) and major semi-axis \( a \). Let \( \alpha_{\min} = \lambda_{\min}(A + A^T)/2 \). Then

\[
\|X - X_m\| \leq \frac{4}{\alpha_{\min}} r_2 - r \left( \frac{r}{r_2} \right)^m
\]

where

\[
r = \frac{a}{d} + \sqrt{\left( \frac{a}{d} \right)^2} - 1, \quad r_2 = \frac{c + \alpha_{\min}}{d} + \sqrt{\left( \frac{c + \alpha_{\min}}{d} \right)^2} - 1.
\]

**Proof.** Let \( \alpha = \alpha_{\min} \) and let \( \hat{E} \) be the ellipse containing the field of values of \( A + \alpha I \). We consider the usual mapping whose boundary image is an elliptic curve in \( \mathbb{C}^+ \) containing the spectrum of \( A + \alpha I \), \( \psi(\tau) = c + \alpha - \frac{\alpha}{2} (\tau + \frac{1}{\tau}) \), with \( \tau = \sigma e^{i\theta} \in C(0, \sigma) \). For \( \sigma = r > 1 = \rho \), \( \psi(\tau) = r = \partial \hat{E} \) and it holds

\[
\exp(-2t\alpha) = \max_{\tau \in C(0, r)} \exp(-t\psi(\tau)),
\]

so that

\[
\frac{1}{2\pi} \int_0^{2\pi} |f(\psi(re^{i\theta}))|d\theta \leq \exp(-2t\alpha) =: \mathcal{M}(t).
\]

Moreover, the value \( r_2 \) corresponds to the reciprocal of the asymptotic convergence factor of the Faber expansion, namely \( r_2 = \psi^{-1}(0) \). Using the fact that \( \hat{E} \) is convex, it also follows that \( |\Phi_k(A + \alpha I)| \leq 2 \) for \( k = 0, 1, \ldots \); see [2]. The same holds for \( |\Phi_k(T_m + \alpha I)| \), since the field of values of \( T_m + \alpha I \) is included in that of \( A + \alpha I \). Therefore, Lemma 2.1 and Theorem 4.1 ensure that

\[
\|\hat{x} - \hat{x}_m\| \leq \sum_{k=m}^{\infty} |f_k| (|\Phi_k(A + \alpha I)| + |\Phi_k(T_m + \alpha I)|)
\]

\[
\leq 8 \exp(-2t\alpha) r_2 - r \left( \frac{r}{r_2} \right)^m, \quad z \in \psi(C(0, r))
\].
for $r$ satisfying $\varrho \leq r < r_2$. Finally,
\[ \|X - X_m\| \leq \int_0^\infty \|\dot{x} - \dot{x}_m\| dt \leq \frac{4}{\alpha} \frac{r_2}{r^2} \left( \frac{r}{r_2} \right)^m, \]
which completes the proof. \(\square\)

In the case of an ellipse, Knizhnerman suggested that one could simplify the derivation and proceed as in Proposition 3.1, where an additional multiplier of the type $\text{const} \kappa$ would arise in (3.2) in the proof [19].

**Example 4.3.** We consider a $400 \times 400$ (normal) diagonal matrix $A$ whose eigenvalues are $\lambda = c + a_1 \cos \theta + a_2 \sin \theta$, $\theta$ uniformly distributed in $[0, 2\pi]$ and $c = 20$, $a_1 = 10$ and $a_2 = 2$, so that the eigenvalues are on an elliptic curve with center $c$ and focal distance $d = \sqrt{a_1^2 - a_2^2} \approx 9.8$. Here $\alpha_{\min} \approx 10.001$, yielding $r \approx 1.2247$ and $r_2 \approx 5.9561$. The vector $b$ is the vector of all ones, normalized to have unit norm. In Figure 4.1 we report the error associated with the Krylov subspace approximation of the Lyapunov solution, and the estimate of Corollary 4.2. The agreement is impressive, as it should be expected since the spectrum exactly lies on the elliptic curve and the matrix is normal, so that the field of values coincides with the associated convex hull.

![Fig. 4.1. Example 4.3. True error and its estimate of Corollary 4.2 for the Krylov subspace solver of the Lyapunov equation.](image)

**Example 4.4.** We next consider the $400 \times 400$ matrix $A$ stemming from the centered finite difference discretization of the operator $\mathcal{L}(u) = -\Delta u + 40(x + y)u_x + 200u$ in the unit square, with Dirichlet boundary conditions. The spectrum of $A$, together with its field of values (computed with the Matlab function `fv.m` in [14]) and a surrounding ellipse, are shown in the left plot of Figure 4.2. Here $\alpha_{\min} = 0.4533$. The ellipse has parameters $c = 4.4535$, $a_1 = c - \alpha_{\min}$, $a_2 = 3.7$ and focal distance $d = \sqrt{a_1^2 - a_2^2} \approx 1.52$, yielding $r \approx 5.0649$ and $r_2 \approx 6.2963$. The right plot of Figure 4.2 shows the convergence history of the Krylov solver, together with the asymptotic factor $(r/r_2)^m$ in the estimate of Corollary 4.2. The initial asymptotic convergence rate is reasonably well captured by the estimate.

**Example 4.5.** We consider the $400 \times 400$ bidiagonal matrix $A$ with uniformly distributed diagonal elements in the interval $[10, 110]$ and unit upper diagonal. In this
Fig. 4.2. Example 4.4. Left plot: Spectrum of $A$, field of values (thin solid curve), and smallest computed elliptic curve including the field of value (thick solid curve). Right plot: True error and its asymptotic factor in the estimate of Corollary 4.2 for the Krylov subspace solver of the Lyapunov equation.

Fig. 4.3. Example 4.5. Left plot: real spectrum, field of values (thin solid curve), and smallest computed elliptic curve including the field of value (thick solid curve). Right plot: True error and its estimate of Corollary 4.2 for the Krylov subspace solver of the Lyapunov equation.

case $\alpha_{\min} = 9.4692$. The vector $b$ is the normalized vector of all ones. Our numerical computation reported in the left plot of Figure 4.3 showed that the field of values of $A$ (computed once again with $fv.m$ [14]) is contained in an ellipse with center $c = 60$, semi-axes $a_1 = 50.8$, $a_2 = 4.2$ and focal distance $d = \sqrt{a_1^2 - a_2^2} \approx 50.62$, yielding $r = 1.0864$ and $r_2 = 2.3118$. The right plot of Figure 4.3 shows the convergence history of the Krylov solver, together with the asymptotic factor in the estimate of Corollary 4.2. Once again, the asymptotic rate is a good estimate of the actual convergence rate. Even more accurate bounds for this example might be obtained by using more appropriate conformal mappings than the ellipse. It may be possible to include the field of values into a rectangle, for which the mapping $\psi$ could be
numerically estimated [9], [11].

The following mapping is a modified version of the external mapping used for instance in [15],

$$
\psi(\tau) = \gamma_1 - \gamma_2(1 - \frac{1}{\tau})^{2-\theta}, \quad \tau = \sigma e^{i\omega},
$$

(4.1)

for $0 < \theta < 1$ and $\gamma_1, \gamma_2 \in \mathbb{R}^+$. The function $\psi$ maps the exterior of the disc $C(0, \sigma)$ onto a wedge-shaped convex set $\Omega$ in $\mathbb{C}^+$. The following result holds.

**Corollary 4.6.** Let $\hat{\Omega} \subset \mathbb{C}^+$ be the wedge-shaped set which is the image through $\hat{\psi}$ of the disk $C(0, r)$, where $\hat{\psi}$ is as in (4.1). Let $r_2 = |\psi^{-1}(0)|$. Assume the field of values of the matrix $A + \alpha_{\min} I$, with $\alpha_{\min} = \lambda_{\min}((A + A^\top)/2)$, is contained in $\hat{\Omega}$. Then

$$
\|X - X_m\| \leq \frac{4}{\alpha_{\min}} \frac{r_2}{r_2 - r} \left( \frac{r}{r_2} \right)^m.
$$

**Proof.** The proof follows the same steps of that of Corollary 4.2.

**Example 4.7.** We consider the $400 \times 400$ (normal) diagonal matrix $A$ whose eigenvalues are on the curve $\psi(\tau) = 2 - 2(1 - 1/\tau)^{2-\omega}, \tau \in C(0, r)$ with $r = 1.1$ and $\omega = 0.3$ (see the left plot of Fig. 4.4). Here $\alpha_{\min} = 1.9627$. The image of the mapping $\psi(\tau) = \alpha_{\min} + \psi(\tau), \tau \in C(0, r)$ thus contains the spectrum of $A + \alpha_{\min} I$. Numerical computation yields $r_2 = |\psi^{-1}(0)| \approx 3.5063$. The vector $b$ is the normalized vector of all ones. The right plot of Figure 4.4 shows the convergence history of the Krylov solver, together with the asymptotic factor $(r/r_2)^m$ in the estimate of Corollary 4.6. The linear asymptotic convergence is fully captured by the estimate.

![Fig. 4.4. Example 4.7. Left plot: Spectrum of $A$. Right plot: True error and its asymptotic factor in the estimate of Corollary 4.6 for the Krylov subspace solver of the Lyapunov equation.](image)

The following mapping was analyzed in [21] and it is associated with a non-convex domain; the specialized case of an annular sector is discussed for instance in [4]. Given a set $\Omega$, assume that $\partial \Omega$ is an analytic Jordan curve. If $\Omega$ is of bounded (or finite)
boundary rotation, then

$$\max_{z \in \Omega} |\Phi_k(z)| \leq \frac{V(\Omega)}{\pi},$$

where $V(\Omega)$ is the boundary rotation of $\Omega$, defined as the total variation of the angle between the positive real axis and the tangent of $\partial \Omega$. In particular, this bound is scale-invariant, so that it also holds that $V(s\Omega) = V(\Omega)$ [21]. These important properties ensure that for a diagonalizable matrix $A$, $\|\Phi_k(A + \alpha_{\min} I)\|$ is bounded independently of $k$, on a non-convex set with bounded boundary rotation. Indeed, letting $A = Q\Lambda Q^{-1}$ be the spectral decomposition of $A$, then $\|\Phi_k(A + \alpha_{\min} I)\| \leq \kappa(Q)\|\Phi_k(A + \alpha_{\min} I)\|$, where $\kappa(Q) = \|Q\|\|Q^{-1}\|$, and the estimate above can be applied.

**Corollary 4.8.** Assume that $A$ is diagonalizable, and let $A = Q\Lambda Q^{-1}$ be its spectral decomposition. Let $0 \leq s \leq 1$. Assume the spectrum of $A + \alpha_{\min} I$ is contained in the set $s\Omega \subset \mathbb{C}^+$, with $\rho > 0$, whose boundary is the "bratwurst" image of $\psi$.

$$\psi(\tau) = (\tau - \lambda N)(\tau - \lambda M) \over (N - M)\tau + \lambda(NM - 1),$$

where $N, M$ and $\lambda$ are given, $\tau \in C(0, r)$, $0 \leq r < r_2 = \psi^{-1}(0)$, with $r$ such that $\psi(C(0, r)) \subset \mathbb{C}^+$. Then,

$$\|X - X_m\| = \sum_{k=m}^\infty |f_k| (\|\Phi_k(A + \alpha I)\| + \|\Phi_k(T_m + \alpha I)\|)
\leq 4M(t) \frac{V(\Omega)\kappa(Q)}{\pi} \frac{r_2}{r_2 - r} \left(\frac{r}{r_2}\right)^m, \quad z \in \psi(C(0, r)),$$

for $r$ satisfying $\rho \leq r < r_2$. Here $M(t) = \exp(-t\psi_{\min})$. Finally,

$$\|X - X_m\| \leq 2 \int_0^\infty \|\dot{x} - \dot{x}_m\| dt \leq \frac{8V(\Omega)\kappa(Q)}{\pi} \int_0^\infty M(t)dt \frac{r_2}{r_2 - r} \left(\frac{r}{r_2}\right)^m.$$

from which the result follows. $\Box$

**Example 4.9.** This example is take from [20]; see also [21] for more details. In this case, $A$ is the $225 \times 225$ matrix pde225 of the Matrix Market repository [23] and it is such that $\alpha_{\min} \approx 0.08249$. The spectrum of $A + \alpha_{\min} I$ is included in the set $2\Omega$ whose boundary is the bratwurst image of $\psi$ as in Corollary 4.8, with $\lambda = -1, N = 1.0457, M = 0.66039$ (exact to the first decimal digits). The left plot of Figure 4.5 shows the spectrum of $A + \alpha_{\min} I$ as ‘×’; the solid curve corresponds to the boundary of $\psi(C(0, r))$ with $r = 0.98$, enclosing the whole spectrum, while the dashed curve is the boundary of $\psi(C(0, r_2))$ with $r = N$. The right plot of Figure 4.5 shows the convergence curve of the Krylov subspace solver, together with the asymptotic quantity $\frac{r_2}{r_2 - r} \left(\frac{r}{r_2}\right)^m$, $m = 1, 2, \ldots$, from Corollary 4.8. We observe that the initial
convergence phase is well captured by the estimate. As expected, the estimate cannot reproduce the superlinear convergence of the solver at later stages.

We conclude this section by observing that other sets can be analyzed within our framework; indeed, more general mappings can be obtained and numerically approximated within the class of Schwarz-Christoffel conformal mappings [6].

5. Connections to linear system solvers and further considerations. The relation

\[ z^{-1} = \int_0^\infty e^{-tz} \, dt \]

can be used to show a close connection between our estimates and the solution of the linear system \((A + \alpha_{\min}I)d = b\) in the Krylov subspace. Let \(V_m(T_m + \alpha_{\min}I)^{-1}e_1\) be the Galerkin approximation to the linear system solution \(d\) in the Krylov subspace \(K_m(A, b) = K_m(A + \alpha_{\min}I, b)\). Then the system error can be written as

\[
(A + \alpha_{\min}I)^{-1}b - V_m(T_m + \alpha_{\min}I)^{-1}e_1 = \int_0^\infty (\exp(-t(A + \alpha_{\min}I))b - V_m \exp(-t(T_m + \alpha I))e_1) \, dt.
\]

Comparing the last integral with the error bound in (2.2), shows that the error norm \(\| (A + \alpha_{\min}I)^{-1}b - V_m(T_m + \alpha_{\min}I)^{-1}e_1 \|\) may be bounded by exactly the same tools we have used for the Lyapunov error and that the two initial integral bounds only differ by a factor of two. Indeed, the estimates of Proposition 3.1 (symmetric case) and of Corollary 4.2 (spectrum contained in an ellipse) employ the same asymptotic factors that characterize the convergence rate of methods such as the Conjugate Gradients in the symmetric case, and FOM or GMRES in the nonsymmetric case, when applied to the system \((A + \alpha_{\min}I)d = b\); see, e.g., [27]. Therefore, we have shown that the convergence of a Galerkin procedure in the Krylov subspace for solving (1.1) has the same convergence factor as a corresponding Krylov subspace method for the shifted (single vector) linear system.
As a natural consequence of the discussion above, the previous results can be generalized to the case when \( b \) is replaced by a matrix \( B \), with more than one column. A Galerkin approximation may be obtained by first generating the “block” Krylov subspace \( K_m(A, B) = \text{span}\{B, AB, \ldots, A^{m-1}B\} \) and then proceeding as described in section 2; see, e.g., [1]. Let \( B = [b_1, \ldots, b_s] \). Setting \( Z = \exp(-tA)B \) and letting \( Z_m \in K_m(A, B) \) be the associated Krylov approximation to the exponential, we can bound \( \|ZZ^T - Z_mZ_m^T\| \) for instance as

\[
\|ZZ^T - Z_mZ_m^T\| \leq \sum_{k=1}^s \|z_m^{(k)}(z_m^{(k)})^T - z_m^{(k)}(z_m^{(k)})^T\|
\]

where \( Z = [z_1, \ldots, z_s] \) and \( Z_m = [z_m^{(1)}, \ldots, z_m^{(s)}] \). The results of the previous sections can be thus applied to each term in the sum. Refined bounds may possibly be obtained by using the theory of matrix polynomials, but this is beyond the scope of this work; see, e.g., [27].

We also observe that our convergence results can be generalized to the case of accelerated methods, such as that described in [28], by using the theoretical matrix function framework described in [8].

Acknowledgments. We are deeply indebted to Leonid Knizhnerman for several insightful comments which helped improve a previous version of this manuscript.

REFERENCES


