SIAM J. MATRIX ANAL. APPL.
Vol. 22, No. 1, pp. 56–65
© 2000 Society for Industrial and Applied Mathematics

CONDITIONING OF THE STABLE, DISCRETE-TIME LYAPUNOV OPERATOR*

MICHAEL K. TIPPETT†, STEPHEN E. COHN‡, RICARDO TODLING§, AND DAN MARCHESIN¶

Abstract. The Schatten p-norm condition of the discrete-time Lyapunov operator $\mathcal{L}_A$ defined on matrices $P \in \mathbb{R}^{n\times n}$ by $\mathcal{L}_A P \equiv P - APA^T$ is studied for stable matrices $A \in \mathbb{R}^{n\times n}$. Bounds are obtained for the norm of $\mathcal{L}_A$ and its inverse that depend on the spectrum, singular values, and radius of stability of A. Since the solution P of the discrete-time algebraic Lyapunov equation (DALE) $\mathcal{L}_A P = Q$ can be ill-conditioned only when either $\mathcal{L}_A$ or Q is ill-conditioned, these bounds are useful in determining whether P admits a low-rank approximation, which is important in the numerical solution of the DALE for large n.

Key words. Lyapunov matrix equation, condition estimates, large-scale systems, radius of stability

AMS subject classifications. 15A12, 93C55, 93A15, 47B65

PII. S0895479899354822

1. Introduction.

Properties of the solution P of the discrete algebraic Lyapunov equation (DALE), $P = APA^T + Q$, are closely related to the stability properties of A. For instance, the DALE has a unique solution $P = P^T > 0$ for any $Q = Q^T > 0$ if A is stable [11], a fact also true in infinite-dimensional Hilbert spaces [18]. In the setting treated here with A, Q, $P \in \mathbb{R}^{n\times n}$, A is stable if its eigenvalues $\lambda_i(A)$, $i = 1, \ldots, n$, lie inside the unit circle; the eigenvalues are ordered so that $|\lambda_1(A)| \geq |\lambda_2(A)| \geq \cdots \geq |\lambda_n(A)|$. Here A is always assumed to be stable.

In applications where the dimension n is very large, direct solution of the DALE or even storage of P is impractical or impossible. For instance, in numerical weather prediction applications A is the matrix that evolves atmospheric state perturbations. The DALE and its continuous-time analogues can be solved directly for simplified atmospheric models [6, 23], but in realistic models n is about $10^6 - 10^7$ and even the storage of P is impossible. Krylov subspace [5] and Monte Carlo [9] methods have been used to find low-rank approximations of the right-hand side of the DALE and of the solution of the DALE [10].

The solution P of the DALE can be well approximated by a rank-deficient matrix if P has some small singular values. Therefore, it is useful to identify properties of A

*Received by the editors April 23, 1999; accepted for publication (in revised form) by R. Bhatia November 1, 1999; published electronically May 31, 2000. This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico grants 91.0029/95-4, 381737/97-7, and 30.0204/83-3, Financiadora de Estudos e Projetos grant 77.97.0315.00, and the NASA EOS Interdisciplinary Project on Data Assimilation. This work was performed by an employee of the U.S. Government or under U.S. Government contract. The U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. Copyright is owned by SIAM to the extent not limited by these rights.

http://www.siam.org/journals/simax/22-1/35482.html

†IRI, Lamont–Doherty Earth Observatory of Columbia University, Palisades, NY 10964-8000 (tippett@iri.ldeo.columbia.edu). This work was done while the author was with the Centro de Previsão de Tempo e Estudos Climáticos, Cachoeira Paulista, SP, Brazil.

‡Data Assimilation Office, Code 910.3, NASA/GSFC, Greenbelt, MD 20771 (cohn@dao.gsfc.nasa.gov).

§General Sciences Corp./SAIC, Code 910.3, NASA/GSFC/DAO, Greenbelt, MD 20771 (todling@dao.gsfc.nasa.gov).

¶Instituto de Matemática Pura e Aplicada, Rio de Janeiro, RJ, Brazil (marchesi@impa.br).

or Q that lead to P being ill-conditioned. If A is normal, then

$$\frac{\lambda_1(P)}{\lambda_n(P)} \leq \frac{\lambda_1(Q)}{\lambda_n(Q)} \frac{1-|\lambda_n(A)|^2}{1-|\lambda_1(A)|^2}\,; \tag{1.1}$$

the conditioning of P is controlled by that of Q and by the spectrum of A. In the general case, the conditioning of Q and of the discrete-time Lyapunov operator $\mathcal{L}_A$ defined by $\mathcal{L}_A P \equiv P - APA^T$ determine when P may be ill-conditioned.

THEOREM 1.1. *Let A be a stable matrix and suppose that $\mathcal{L}_A P = Q$ for $Q = Q^T > 0$. Then*

$$\|P\|_p \, \|P^{-1}\|_p \leq \|\mathcal{L}_A\|_p \, \|\mathcal{L}_A^{-1}\|_p \, \|Q\|_p \, \|Q^{-1}\|_p\,, \qquad p = \infty\,, \tag{1.2}$$

where $\|\cdot\|_p$ is the Schatten p-norm (see (2.2)).

Theorem 1.1 (see proof in the appendix) follows from $\mathcal{L}_A^{-1}$ and its adjoint being positive operators. Therefore, the same connection between rank-deficient approximate solutions and operator conditioning exists for matrix equations such as the continuous algebraic Lyapunov equation. We note that Theorem 1.1 also holds for $1 \leq p < \infty$ if either A is singular or $\sigma_1^2(A) \geq 2$; $\sigma_1(A)$ is the largest singular value of A.

Here we characterize the Schatten p-norm condition of $\mathcal{L}_A$. The main results are the following. Theorem 3.1 bounds $\|\mathcal{L}_A\|_p$ in terms of the singular values of A. A lower bound for $\|\mathcal{L}_A^{-1}\|_p$ depending on $\lambda_1(A)$ is presented in Theorem 4.1, generalizing results of [7]. Theorem 4.2 gives lower bounds for $\|\mathcal{L}_A^{-1}\|_1$ and $\|\mathcal{L}_A^{-1}\|_\infty$ in terms of the singular values of A. Theorem 4.6 gives an upper bound for $\|\mathcal{L}_A^{-1}\|_p$ depending on the radius of stability of A and generalizes results in [20]. Three examples illustrating the results are included. The issue of whether $\mathcal{L}_A$ and $\mathcal{L}_A^{-1}$ achieve their norms on symmetric, positive definite matrices is addressed in the concluding remarks.

2. Preliminaries. We investigate the condition number $\kappa(\mathcal{L}_A) = \|\mathcal{L}_A\| \, \|\mathcal{L}_A^{-1}\|$, where $\|\cdot\|$ is a norm on $\mathbb{R}^{n^2 \times n^2}$ induced by a matrix norm on $\mathbb{R}^{n \times n}$. Specifically, for $\mathcal{M} \in \mathbb{R}^{n^2 \times n^2}$ we consider norms defined by

$$\|\mathcal{M}\|_p = \max_{S \neq 0 \in \mathbb{R}^{n \times n}} \frac{\|\mathcal{M}S\|_p}{\|S\|_p}\,, \quad 1 \leq p \leq \infty\,, \tag{2.1}$$

where the Schatten matrix p-norm for $S \in \mathbb{R}^{n \times n}$ is defined by

$$\|S\|_p = \left(\sum_{i=1}^{n} (\sigma_i(S))^p \right)^{1/p}\,; \tag{2.2}$$

$\sigma_i(S)$ are the singular values of S with ordering $\sigma_1(S) \geq \sigma_2(S) \geq \cdots \geq \sigma_n(S) \geq 0$. On $\mathbb{R}^{n \times n}$, $\|\cdot\|_2$ is the Frobenius norm and $\|\cdot\|_\infty = \sigma_1(\cdot)$. If $S = S^T \geq 0$, then $\|S\|_1 = \operatorname{tr} S$. The following lemma about the Schatten p-norms follows from their being unitarily invariant [1, p. 94].

LEMMA 2.1. *For any three matrices X, Y, and $Z \in \mathbb{R}^{n \times n}$,*

$$\|XYZ\|_p \leq \|X\|_\infty \|Y\|_p \|Z\|_\infty\,, \qquad 1 \leq p \leq \infty\,. \tag{2.3}$$

The $p = 2$ Schatten norm on $\mathbb{R}^{n \times n}$ is equivalently defined as $\|S\|_2^2 = (S, S)$, where $(\cdot, \cdot)$ is the inner product on $\mathbb{R}^{n \times n}$ defined by $(S_1, S_2) = \operatorname{tr} S_1^T S_2$. This norm

corresponds to the usual Euclidean norm on $\mathbb{R}^{n^2}$ since $\|S\|_2^2$ is equal to the sum of the squares of the entries of S. As a consequence $\kappa_2(\mathcal{L}_A) = \sigma_1(\mathcal{L}_A)/\sigma_{n^2}(\mathcal{L}_A)$, where $\sigma_1(\mathcal{L}_A)$ and $\sigma_{n^2}(\mathcal{L}_A)$ are, respectively, the largest and smallest singular values of $\mathcal{L}_A$. The adjoint of $\mathcal{L}_A$ is given by $\mathcal{L}_A^* S = \mathcal{L}_{A^T} S = S - A^T S A$.

We now state some lemmas about mappings $\mathcal{M} \in \mathbb{R}^{n^2 \times n^2}$ and about the spectra of $\mathcal{L}_A$ and A.

LEMMA 2.2 (see [2, equation (15)]). $\|\mathcal{M}\|_p \le \|\mathcal{M}\|_1^{1/p} \|\mathcal{M}\|_\infty^{1-1/p}, \quad 1 \le p \le \infty.$

LEMMA 2.3. $\|\mathcal{M}\|_1 = \|\mathcal{M}^*\|_\infty$.

LEMMA 2.4 (see [2, proof of Theorem 1]). *If $\mathcal{M}S > 0$ for all $S \in \mathbb{R}^{n\times n}$ such that $S > 0$, then $\|\mathcal{M}\|_\infty = \|\mathcal{M}I\|_\infty$.*

LEMMA 2.5 (see [13, 14]). *The n^2 eigenvalues of $\mathcal{L}_A$ are $1 - \lambda_i(A)\overline{\lambda_j(A)}$, $1 \le i, j \le n$.*

3. The norm of the Lyapunov operator. If A is normal, then $\mathcal{L}_A$ is normal, and its conditioning in the $p = 2$ Schatten norm depends only on its eigenvalues. Therefore, when A is normal,

$$\|\mathcal{L}_A^{-1}\|_2 = \frac{1}{\sigma_{n^2}(\mathcal{L}_A)} = \frac{1}{|\lambda_{n^2}(\mathcal{L}_A)|} = \frac{1}{1 - |\lambda_1(A)|^2} \tag{3.1}$$

and

$$\|\mathcal{L}_A\|_2 = \sigma_1(\mathcal{L}_A) = |\lambda_1(\mathcal{L}_A)| = \max_{i,j} |1 - \lambda_i(A)\overline{\lambda_j(A)}|\,. \tag{3.2}$$

For general A, the following theorem bounds $\|\mathcal{L}_A\|_p$ in terms of the singular values of A.

THEOREM 3.1.

$$|1 - \sigma_1^2(A)| \le \max_j |1 - \sigma_j^2(A)| \le \|\mathcal{L}_A\|_p \le 1 + \sigma_1^2(A)\,, \quad 1 \le p \le \infty\,. \tag{3.3}$$

Proof. Note that $\mathcal{L}_A v_j v_j^T = v_j v_j^T - \sigma_j^2 u_j u_j^T$, where u_j and v_j are, respectively, the jth left and right singular vectors of A such that $Av_j = \sigma_j u_j$. The lower bound follows from $\|u_j u_j^T\|_p = \|v_j v_j^T\|_p = 1$ and

$$\|\mathcal{L}_A\|_p \ge \|v_j v_j^T - \sigma_j^2 u_j u_j^T\|_p \ge \left|\|v_j v_j^T\|_p - \|\sigma_j^2 u_j u_j^T\|_p\right| = \left|1 - \sigma_j^2\right|\,. \tag{3.4}$$

The upper bound follows from

$$\|\mathcal{L}_A P\|_p \le \|P\|_p + \|APA^T\|_p \le \|P\|_p + \|A\|_\infty^2 \|P\|_p\,. \qquad \square \tag{3.5}$$

If A is normal, $\sigma_j(A)$ can be replaced by $|\lambda_j(A)|$ in Theorem 3.1, and $\|\mathcal{L}_A\|_p \le 1 + |\lambda_1(A)|^2$. If A is normal and $(-\overline{\lambda_1(A)})$ is an eigenvalue of A, then $1 + |\lambda_1(A)|^2$ is an eigenvalue of $\mathcal{L}_A$ and $\|\mathcal{L}_A\|_p = 1 + |\lambda_1(A)|^2$.

Theorem 3.1 shows that $\|\mathcal{L}_A\|_p$ is large and contributes to ill-conditioning if and only if $\sigma_1(A)$ is large, a situation that occurs in various applications [3, 22]. If $\sigma_1(A) \gg 1$ and $|\lambda_1(A)| < 1$, A is highly nonnormal [8, p. 314] and, as Corollary 4.8 will show, close to an unstable matrix.

4. The norm of the inverse Lyapunov operator. We first show that a sufficient condition for $\|\mathcal{L}_A^{-1}\|_p$ to be large is that $\lambda_1(A)$ be near the unit circle. The condition is necessary when A is normal.

THEOREM 4.1. *Let A be a stable matrix. Then*

$$\|\mathcal{L}_A^{-1}\|_p \geq \frac{1}{1-|\lambda_1(A)|^2}\,, \qquad 1 \leq p \leq \infty\,, \tag{4.1}$$

with equality holding if A is normal.

Proof. To obtain the lower bound, let z_1 be the leading eigenvector of A, $Az_1 = \lambda_1(A)z_1$, and note that $\mathcal{L}_A z_1 z_1^H = (1-|\lambda_1(A)|^2)z_1 z_1^H$, where $(\cdot)^H$ denotes conjugate transpose. Either $\operatorname{Re} z_1 z_1^H \neq 0$ or $\operatorname{Im} z_1 z_1^H \neq 0$ is an eigenvector of $\mathcal{L}_A$, and it follows that $\|\mathcal{L}_A^{-1}\|_p \geq (1-|\lambda_1(A)|^2)^{-1}$. Finally, if A is normal, then

$$\mathcal{L}_{A^T}^{-1} I = \mathcal{L}_A^{-1} I = \sum_{i=1}^{n} \frac{1}{1-|\lambda_i(A)|^2} z_i z_i^H\,, \tag{4.2}$$

and $\|\mathcal{L}_A^{-1}\|_\infty = \|\mathcal{L}_A^{-1}\|_1 = (1-|\lambda_1(A)|^2)^{-1}$. Using Lemma 2.2 gives $\|\mathcal{L}_A^{-1}\|_p \leq (1-|\lambda_1(A)|^2)^{-1}$ when A is normal, and therefore $\|\mathcal{L}_A^{-1}\|_p = (1-|\lambda_1(A)|^2)^{-1}$. ∎

When A is nonnormal, $\|\mathcal{L}_A^{-1}\|_p$ can be large without $\lambda_1(A)$ being near the unit circle. For instance, if $\sigma_1(A)$ is large or, more generally, if $\|A^k\|_\infty$ converges to zero slowly as a function of k, then $\|\mathcal{L}_A^{-1}\|_p$ is large. We show this fact first for $p = 1, \infty$.

THEOREM 4.2. *Let A be a stable matrix. For all $m \geq 1$,*

$$\|\mathcal{L}_A^{-1}\|_1 = \left\| \sum_{k=0}^{\infty} \left(A^k\right)^{\mathrm{T}} A^k \right\|_\infty \geq \left\| \sum_{k=0}^{m} \left(A^k\right)^{\mathrm{T}} A^k \right\|_\infty + \frac{\sigma_n^{2(m+1)}(A)}{1-\sigma_n^2(A)}\,, \tag{4.3}$$

$$\|\mathcal{L}_A^{-1}\|_\infty = \left\| \sum_{k=0}^{\infty} A^k \left(A^k\right)^{\mathrm{T}} \right\|_\infty \geq \left\| \sum_{k=0}^{m} A^k \left(A^k\right)^{\mathrm{T}} \right\|_\infty + \frac{\sigma_n^{2(m+1)}(A)}{1-\sigma_n^2(A)}\,. \tag{4.4}$$

In particular,

$$\|\mathcal{L}_A^{-1}\|_p \geq 1 + \sigma_1^2(A) + \frac{\sigma_n^4(A)}{1-\sigma_n^2(A)}\,, \qquad p = 1, \infty\,. \tag{4.5}$$

Proof. The operator $\mathcal{L}_A^{-1}$ applied to $S \in \mathbb{R}^{n\times n}$ can be expressed as [18]

$$\mathcal{L}_A^{-1} S = \sum_{k=0}^{\infty} A^k S \left(A^k\right)^{\mathrm{T}}. \tag{4.6}$$

Applying Lemma 2.4 gives $\|\mathcal{L}_A^{-1}\|_\infty = \|\mathcal{L}_A^{-1} I\|_\infty$, with the inequality in (4.4) being a consequence of

$$\left\| \sum_{k=0}^{\infty} A^k \left(A^k\right)^{\mathrm{T}} \right\|_\infty \geq \left\| \sum_{k=0}^{m} A^k \left(A^k\right)^{\mathrm{T}} \right\|_\infty + \lambda_n \left(\sum_{k=m+1}^{\infty} A^k \left(A^T\right)^k \right), \tag{4.7}$$

and

$$\lambda_n \left(\sum_{k=m+1}^{\infty} A^k \left(A^T\right)^k \right) \geq \sum_{k=m+1}^{\infty} \lambda_n \left(A^k \left(A^T\right)^k \right) \geq \sum_{k=m+1}^{\infty} \sigma_n^{2k}(A) = \frac{\sigma_n^{2(m+1)}(A)}{1-\sigma_n^2(A)}\,, \tag{4.8}$$

where we have used the facts that for matrices $W, X, Y \in \mathbb{R}^{n\times n}$ with X, Y being symmetric positive semidefinite, $\lambda_i(X+Y) \geq \lambda_i(X) + \lambda_n(Y)$, and $\lambda_i(WXW^T) \geq \sigma_n^2(W)\lambda_i(X)$ [17]. Likewise the $p = 1$ results follow from $\|\mathcal{L}_A^{-1}\|_1 = \|\mathcal{L}_{A^T}^{-1} I\|_\infty$. ∎

Lower bounds for $1 < p < \infty$ follow trivially, e.g.,

$$\|\mathcal{L}_A^{-1}\|_p \geq \frac{\|\mathcal{L}_A^{-1} I\|_p}{\|I\|_p} = \frac{\|\mathcal{L}_A^{-1} I\|_p}{n^{1/p}} \geq n^{-1/p} \|\mathcal{L}_A^{-1}\|_\infty \,, \tag{4.9}$$

but give little information when n is large. A lower bound for $1 \leq p \leq \infty$ depending on $\sigma_1(A)$ and independent of n is given in Corollary 4.9.

We now relate $\|\mathcal{L}_A^{-1}\|_p$ to the distance from A to the set of unstable matrices as measured by its *radius of stability* [15].

DEFINITION 4.3. *For any stable matrix $A \in \mathbb{R}^{n\times n}$ define the radius of stability $r(A)$ by*

$$r(A) \equiv \min_{0\leq\theta\leq 2\pi} \|(e^{i\theta} I - A)^{-1}\|_\infty^{-1} = \min_{0\leq\theta\leq 2\pi} \|R(e^{i\theta}, A)\|_\infty^{-1} \,, \tag{4.10}$$

where the resolvent of A is $R(\lambda, A) = (\lambda I - A)^{-1}$.

If A is normal and stable, then $r(A) = 1 - |\lambda_1(A)|$. However, if A is nonnormal and if its eigenvalues are *sensitive* to perturbations, then $r(A) \ll 1 - |\lambda_1(A)|$. The sensitivity of the eigenvalues of A is most completely described by its *pseudospectrum* [21]. The radius of stability $r(A)$ is the largest value of ϵ such that the ϵ-pseudospectrum of A lies inside the unit circle; $r(A)$ being small indicates that the ϵ-pseudospectrum of A is close to the unit circle for small ϵ. The following theorem shows that when $r(A)$ is small, $\|\mathcal{L}_A^{-1}\|_p$ must be large.

THEOREM 4.4 (proven for $p = \infty$ in [7]). *Let A be a stable matrix. Then*

$$\|\mathcal{L}_A^{-1}\|_p \geq \frac{1}{2r(A) + r^2(A)} \,, \qquad 1 \leq p \leq \infty \,. \tag{4.11}$$

Proof. There exists a matrix $E \in \mathbb{R}^{n\times n}$ with $|\lambda_1(A + E)| = 1$ and $\|E\|_\infty = r(A)$. Therefore, there exists a vector x with $x^H x = 1$ such that $(A + E)x = e^{i\theta} x$ for some $0 \leq \theta \leq 2\pi$. Using $\|xx^H\|_p = 1$ and Lemma 2.1 gives

$$\begin{aligned} \|\mathcal{L}_A xx^H\|_p &= \| - Exx^H E^T + e^{i\theta} xx^H E^T + e^{-i\theta} Exx^H \| \\ &\leq \|Exx^H E^T\|_p + \|xx^H E^T\|_p + \|Exx^H\|_p \\ &\leq \|E\|_\infty^2 + 2\|E\|_\infty = r^2(A) + 2r(A) \,, \end{aligned} \tag{4.12}$$

and we have

$$\|\mathcal{L}_A^{-1}\|_p \geq \frac{\|\mathcal{L}_A^{-1}\mathcal{L}_A xx^H\|_p}{\|\mathcal{L}_A xx^H\|_p} = \frac{1}{\|\mathcal{L}_A xx^H\|_p} \geq \frac{1}{2r(A) + r^2(A)} \,. \tag{4.13}$$

∎

A consequence of Theorem 4.4 is the following lower bound for $r(A)$ in terms of $\|\mathcal{L}_A^{-1}\|_p$.

COROLLARY 4.5. *Let A be a stable matrix. Then*

$$r(A) \geq \frac{\|\mathcal{L}_A^{-1}\|_p^{-1}}{1 + \sqrt{1 + \|\mathcal{L}_A^{-1}\|_p^{-1}}} \,, \qquad 1 \leq p \leq \infty \,. \tag{4.14}$$

Bounds for $r(A)$ are useful in robust stability [12] and in the study of perturbations of the discrete algebraic Riccati equation (DARE) [19]. In [19, Lemma 2.2] the bound

$$r(A) \geq \frac{\|\mathcal{L}_A^{-1}\|_\infty^{-1}}{\sigma_1(A) + \sqrt{\sigma_1^2(A) + \|\mathcal{L}_A^{-1}\|_\infty^{-1}}} \tag{4.15}$$

was used to formulate conditions under which a perturbed DARE has a unique, symmetric, positive definite solution. Since the lower bound in (4.14) with $p = \infty$ is sharper than that in (4.15) when $\sigma_1(A) > 1$, it can be used to show existence of a unique, symmetric, positive definite solution of the perturbed DARE for a larger class of perturbations [19, Theorem 4.1].

We generalize to Schatten p-norms the conjecture of [7] proven in [20] for the Frobenius norm.

THEOREM 4.6. *Let A be a stable matrix. Then*

$$\|\mathcal{L}_A^{-1}\|_p \le \frac{1}{r^2(A)}\,, \quad 1 \le p \le \infty\,. \tag{4.16}$$

Proof. $\mathcal{L}_A^{-1} I$ can be expressed as [20, 13]

$$\mathcal{L}_A^{-1} I = \frac{1}{2\pi}\int_0^{2\pi} R(e^{i\theta}, A) R(e^{i\theta}, A)^H \, d\theta. \tag{4.17}$$

Therefore, from Lemma 2.4,

$$\|\mathcal{L}_A^{-1}\|_\infty = \|\mathcal{L}_A^{-1} I\|_\infty \le \frac{1}{2\pi}\int_0^{2\pi} \|R(e^{i\theta}, A)\|_\infty^2 \, d\theta \le \frac{1}{r^2(A)}\,. \tag{4.18}$$

The inequality (4.16) for $p = 1$ follows from $\|\mathcal{L}_A^{-1}\|_1 = \|\mathcal{L}_{A^T}^{-1} I\|_\infty$ and $r(A) = r(A^T)$. The theorem follows from Lemma 2.2. ∎

As a consequence, any solution of the DALE can be used to obtain an upper bound for $r(A)$.

COROLLARY 4.7. *Let A be a stable matrix and let $\mathcal{L}_A P = Q$. Then*

$$r^2(A) \le \frac{\|Q\|_p}{\|P\|_p}\,, \qquad 1 \le p \le \infty\,. \tag{4.19}$$

Theorem 4.6 can be combined with any lower bound for $\|\mathcal{L}_A^{-1}\|_p$ to obtain an upper bound for $r(A)$. For instance, from Theorem 4.2 we get the following upper bound.

COROLLARY 4.8. *Let A be a stable matrix. Then*

$$r^2(A) \le \frac{1}{1 + \sigma_1^2(A)}\,. \tag{4.20}$$

Combining Corollary 4.8 and Theorem 4.4 gives a lower bound for $\|\mathcal{L}_A^{-1}\|_p$.

COROLLARY 4.9. *Let A be a stable matrix. Then*

$$\|\mathcal{L}_A^{-1}\|_p \ge \frac{1 + \sigma_1^2(A)}{1 + 2\sqrt{1 + \sigma_1^2(A)}}\,, \qquad 1 \le p \le \infty\,. \tag{4.21}$$

5. Examples. We present three examples that illustrate how ill-conditioning of $\mathcal{L}_A$ leads to low-rank approximate solutions of the DALE.

Example 1. *Almost unit eigenvalues.* Take $A = \lambda z z^T$, where λ and z are real, $0 < \lambda < 1$, and $z^T z = 1$. The matrix A is symmetric and $\mathcal{L}_A$ is self-adjoint. The eigenvalues of A are $(\lambda, 0, \ldots, 0)$. The operator $\mathcal{L}_A$ has singular values (and eigenvalues) $(1, \ldots, 1, 1 - \lambda^2)$. Therefore, $\|\mathcal{L}_A\|_2 = 1$ and $1 \le \|\mathcal{L}_A\|_p \le 1 + \lambda^2$ from Theorem 3.1. The norm of the inverse Lyapunov operator is

$$\|\mathcal{L}_A^{-1}\|_p = \frac{1}{1 - \lambda^2}\,, \qquad 1 \le p \le \infty\,, \tag{5.1}$$

according to Theorem 4.1. As the eigenvalue λ approaches the unit circle, $\mathcal{L}_A$ is increasingly poorly conditioned. The solution of the DALE for this choice of A is

$$P = \frac{\lambda^2}{1-\lambda^2}\left(z^T Q z\right) z z^T + Q\,. \tag{5.2}$$

A "natural" rank-1 approximation $\tilde{P}$ of P is $\tilde{P} = \lambda^2(1-\lambda^2)^{-1}(z^TQz)zz^T$. As the eigenvalue λ approaches the unit circle, if (z^TQz) is nonzero, P is increasingly well approximated by $\tilde{P}$ in the sense that $\|P-\tilde{P}\|_p/\|P\|_p$ approaches zero.

Example 2. *Large singular values.* Take $A = \sigma y z^T$, where $\sigma > 0$ and y and z are real unit n-vectors. The matrix A has at most one nonzero eigenvalue, namely, $\lambda = \sigma(y^Tz)$, taken to be less than one in absolute value. The sensitivity s of the eigenvalue λ is the cosine of the angle between y and z, i.e., $s = \lambda/\sigma$ for $\lambda \neq 0$, indicating that λ is sensitive to perturbations to A when σ is large [8].

Theorem 3.1 gives that $1+\sigma^2 \geq \|\mathcal{L}_A\|_p \geq |1-\sigma^2|$, showing that $\|\mathcal{L}_A\|_p$ is large when σ is large. From Lemmas 2.3 and 2.4,

$$\|\mathcal{L}_A^{-1}\|_1 = \|\mathcal{L}_A^{-1}\|_\infty = 1 + \frac{\sigma^2}{1-\lambda^2}\,, \tag{5.3}$$

and it follows from Lemma 2.2 that $\|\mathcal{L}_A^{-1}\|_p \leq 1+\sigma^2/(1-\lambda^2)$. A lower bound for the $(p=2)$-norm is

$$\|\mathcal{L}_A^{-1}\|_2 \geq \|\mathcal{L}_A^{-1} z z^T\|_2 = \sqrt{1 + 2\frac{\lambda^2}{1-\lambda^2} + \frac{\sigma^4}{(1-\lambda^2)^2}}\,. \tag{5.4}$$

The matrix A is near an unstable matrix either when $|\lambda|$ is near unity or when σ is large since

$$\left\|\left(e^{i\theta}I - \sigma y z^T\right)^{-1}\right\|_\infty = \left\|e^{-i\theta}I + \frac{\sigma e^{-2i\theta}}{1-\lambda e^{-i\theta}} y z^T\right\|_\infty \geq 1 + \frac{2|\lambda|}{1-|\lambda|} + \frac{\sigma^2}{(1-|\lambda|)^2}\,. \tag{5.5}$$

Therefore, $r(A) \leq (1-|\lambda|)/\sigma$ and a lower bound on $\|\mathcal{L}_A^{-1}\|_p$ follows from Theorem 4.4. When either $|\lambda|$ is close to unity or σ is large, $r(A)$ is small and $\kappa_p(\mathcal{L}_A)$ is large.

The solution of the DALE is

$$P = \frac{\sigma^2}{1-\lambda^2}\left(z^T Q z\right) y y^T + Q\,. \tag{5.6}$$

When $\mathcal{L}_A$ is ill-conditioned and $\left(z^TQz\right) \neq 0$, the rank-1 matrix $\tilde{P} = \sigma^2(1-\lambda^2)^{-1}\times (z^TQz)yy^T$ is a good approximation of P in the sense that $\|P-\tilde{P}\|_p/\|P\|_p$ is small.

Example 3. *Sensitive eigenvalues.* Consider the dynamics arising from the one-dimensional advection equation, $w_t + w_x = 0$ for $0 \leq x \leq n$, with boundary condition $w(0,t) = 0$. The matrix A that advances the n-vector $w(x = 1,2,\ldots,n, t = t_0)$ to $w(x = 1,2,\ldots,n, t = t_0+1)$ is the $n \times n$ matrix with ones on the subdiagonal and zero elsewhere, i.e., the transpose of an $n \times n$ Jordan block with zero eigenvalue. Adding stochastic forcing with covariance Q at unit time intervals leads to the DALE, $\mathcal{L}_A P = Q$, where P is the steady-state covariance of w.

Since $\sigma_1(A) = 1$, Theorem 3.1 yields $1 \leq \|\mathcal{L}_A\|_p \leq 2$. Further, since $\|\mathcal{L}_A\|_1 \geq \|\mathcal{L}_A e_1 e_1^T\|_1 = \|e_1 e_1^T - e_2 e_2^T\|_1 = 2$, where e_j is the jth column of the identity matrix,

$\|\mathcal{L}_A\|_1 = 2$. A similar argument with $\mathcal{L}_{A^T}$ gives $\|\mathcal{L}_A\|_\infty = 2$. Calculating $\mathcal{L}_A^{-1} I$ and $\mathcal{L}_{A^T}^{-1} I$ gives $\|\mathcal{L}_A^{-1}\|_\infty = \|\mathcal{L}_A^{-1}\|_1 = n$. Therefore, using Lemma 2.2, $\|\mathcal{L}_A^{-1}\|_p \le n$. Also,

$$\|\mathcal{L}_A^{-1}\|_2 \ge \frac{\|\mathcal{L}_A^{-1} e_1 e_1^T\|_2}{\|e_1 e_1^T\|_2} = \sqrt{n}\,. \tag{5.7}$$

A direct calculation shows that

$$\|(e^{i\theta} I - A)^{-1}\|_2^2 = \left\| \sum_{k=0}^{n-1} A^k e^{-i(k+1)\theta} \right\|_2^2 = \frac{n(n+1)}{2} \tag{5.8}$$

for any real θ. Since $\sqrt{n}\|(e^{i\theta} I - A)^{-1}\|_\infty \ge \|(e^{i\theta} I - A)^{-1}\|_2$, we have $r^2(A) \le 2/(n+1)$. Theorem 4.4 then gives a lower bound for $\|\mathcal{L}_A^{-1}\|_p$, $1 \le p \le \infty$. Thus as n becomes large; that is, as the domain becomes large with respect to the advection length scale, $\mathcal{L}_A$ is increasingly ill-conditioned.

The elements P_{ij} of the solution P of the DALE are

$$P_{ij} = e_i^T P e_j = \sum_{k=0}^{n-1} e_i^T A^k Q (A^T)^k e_j = \sum_{k=0}^{\min(i-1, j-1)} Q_{i-k, j-k}\,. \tag{5.9}$$

Therefore, if $Q = Q^T > 0$, a "natural" rank-m approximation of P is the matrix $\tilde{P}$ defined by

$$\tilde{P}_{i,j} = \begin{cases} P_{i,j}\,, & n - m < i, j \le n, \\ 0 & \text{otherwise.} \end{cases} \tag{5.10}$$

When Q is diagonal, P is also diagonal and

$$P_{ii} = \sum_{k=1}^{i} Q_{kk}\,. \tag{5.11}$$

In this case, each $Q_{kk} > 0$ and $\tilde{P}$ is the best rank-m approximation of P in the sense of minimizing $\|P - \tilde{P}\|_p$. We note that $\tilde{P}$ is associated with the left-most part of the domain $0 \le x \le n$.

6. Concluding remarks. Results about $\|\mathcal{L}_A^{-1}\|_p$ translate into bounds for solutions of the DALE. For instance, the solution P of the DALE for $Q = Q^T \ge 0$ satisfies

$$\operatorname{tr} P \le \|\mathcal{L}_A^{-1}\|_1 \operatorname{tr} Q\,, \tag{6.1}$$

and the upper bound is achieved for $Q = w_1 w_1^T$, where w_1 is the leading eigenvector of $\mathcal{L}_{A^T}^{-1} I$. In the $(p = \infty)$-norm, $\mathcal{L}_A^{-1}$ achieves its norm on the identity. In the $(p = 2)$-norm, $\mathcal{L}_A^{-1}$ does not in general achieve its norm on the identity, and the question arises whether it achieves its norm on any symmetric, positive semidefinite matrix. The forward operator $\mathcal{L}_A$ does not in general assume its norm on a symmetric, positive semidefinite matrix. The following theorem states that $\mathcal{L}_A^{-1}$ does achieve its $(p = 2)$-norm on a symmetric, positive semidefinite matrix.

THEOREM 6.1. *There exists a matrix $S = S^T \ge 0$ such that $\|\mathcal{L}_A^{-1} S\|_2 / \|S\|_2 = \|\mathcal{L}_A^{-1}\|_2$.*

Proof. Theorem 8 of [4] states that the inverse of the stable, continuous-time Lyapunov operator achieves its $(p = 2)$-norm on a symmetric matrix. The proof is easily adapted to give that $\mathcal{L}_A^{-1}$ achieves its $(p = 2)$-norm on a symmetric matrix. We now show that if $\mathcal{L}_A^{-1}$ achieves its $(p = 2)$-norm on a symmetric matrix, it does so on a symmetric, positive semidefinite matrix. Suppose that $\|\mathcal{L}_A^{-1}S\|_2/\|S\|_2 = \|\mathcal{L}_A^{-1}\|_2$ and S is symmetric with Schur decomposition $S = UDU^T$. Define the symmetric, positive semidefinite matrix $S^+ = U|D|U^T$. Then $\|S\|_2 = \|S^+\|_2$ and $-S^+ \leq S \leq S^+$. The positiveness of the stable, discrete-time inverse Lyapunov operator mapping implies that $-\mathcal{L}_A^{-1}S^+ \leq \mathcal{L}_A^{-1}S \leq \mathcal{L}_A^{-1}S^+$, which implies that $\|\mathcal{L}_A^{-1}S\|_2 \leq \|\mathcal{L}_A^{-1}S^+\|_2$. Therefore,

$$\frac{\|\mathcal{L}_A^{-1}S\|_2}{\|S\|_2} = \frac{\|\mathcal{L}_A^{-1}S\|_2}{\|S^+\|_2} \leq \frac{\|\mathcal{L}_A^{-1}S^+\|_2}{\|S^+\|_2}\,. \tag{6.2}$$ ∎

Additional information about the leading singular vectors of $\mathcal{L}_A^{-1}$ could be useful for determining low-rank approximations of P. The power method can be applied to $\mathcal{L}_{A^T}^{-1}\mathcal{L}_A^{-1}$ to calculate the leading right singular vector and singular value of $\mathcal{L}_A^{-1}$ [7]. However, this approach requires solving two DALEs at each iteration, which may be impractical for large n. If it is practical to store P and to apply $\mathcal{L}_A$ and $\mathcal{L}_{A^T}$, a Lanczos method could be used to compute the trailing eigenvectors of $\mathcal{L}_A\mathcal{L}_{A^T}$ while avoiding the cost of solving any DALEs.

Appendix. *Proof of Theorem* 1.1. By definition, $\|P\|_p \leq \|\mathcal{L}_A^{-1}\|_p \|Q\|_p$, and it remains to show that $\|P^{-1}\|_\infty \leq \|\mathcal{L}_A\|_\infty \|Q^{-1}\|_\infty$. Since $P = P^T > 0$, there is a nonzero $x \in \mathbb{R}^n$ such that

$$\|P^{-1}\|_\infty = \frac{1}{\lambda_n(P)} = \frac{x^T x}{x^T\left(\mathcal{L}_A^{-1}Q\right)x} = \frac{\operatorname{tr} xx^T}{\operatorname{tr}\left(\mathcal{L}_A^{-1}Q\right)xx^T} = \frac{\operatorname{tr} xx^T}{\operatorname{tr}\left(\left(\mathcal{L}_{A^T}\right)^{-1}xx^T\right)Q}\,. \tag{A.1}$$

Let $B = \mathcal{L}_{A^T}^{-1}(xx^T)$ and note $B = B^T \geq 0$. Then using Lemma 2.3 and $\operatorname{tr} BQ \geq \lambda_n(Q)\operatorname{tr} B$ gives

$$\|P^{-1}\|_\infty = \frac{\operatorname{tr}\mathcal{L}_{A^T}B}{\operatorname{tr} BQ} \leq \frac{\operatorname{tr}\mathcal{L}_{A^T}B}{\operatorname{tr} B}\frac{1}{\lambda_n(Q)} \leq \|\mathcal{L}_{A^T}\|_1\|Q^{-1}\|_\infty = \|\mathcal{L}_A\|_\infty\|Q^{-1}\|_\infty\,. \tag{A.2}$$ ∎

Theorem 1.1 holds for $1 \leq p \leq \infty$ given some restrictions on A. From [16], $\lambda_i(P) \geq \lambda_i(Q) + \sigma_n^2(A)\lambda_n(P)$, and it follows that $\|P^{-1}\|_p \leq \|Q^{-1}\|_p$ for $1 \leq p \leq \infty$. From Theorem 3.1, $\|\mathcal{L}_A\|_p \geq 1$ if either A is singular or $\sigma_1^2(A) \geq 2$. Therefore, if either A is singular or $\sigma_1^2(A) \geq 2$,

$$\|P^{-1}\|_p \leq \|\mathcal{L}_A\|_p\|Q^{-1}\|_p\,, \qquad 1 \leq p \leq \infty\,. \tag{A.3}$$

Acknowledgments. The authors thank Greg Gaspari for valuable observations and notation suggestions and the reviewer for useful comments.

REFERENCES

[1] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.
[2] R. Bhatia, *A note on the Lyapunov equation*, Linear Algebra Appl., 259 (1997), pp. 71–76.

[3] K. M. Butler and B. F. Farrell, *Three-dimensional optimal perturbations in viscous shear flow*, Phys. Fluids A, 4 (1992), pp. 1637–1650.
[4] R. Byers and S. Nash, *On the singular "vectors" of the Lyapunov operator*, SIAM J. Alg. Discrete Methods, 8 (1987), pp. 59–66.
[5] S. E. Cohn and R. Todling, *Approximate data assimilation schemes for stable and unstable dynamics*, J. Meteor. Soc. Japan, 74 (1996), pp. 63–75.
[6] B. F. Farrell and P. J. Ioannou, *Generalized stability theory. Part* I: *Autonomous operators*, J. Atmospheric Sci., 53 (1996), pp. 2025–2040.
[7] P. M. Gahinet, A. J. Laub, C. S. Kenney, and G. A. Hewer, *Sensitivity of the stable discrete-time Lyapunov equation*, IEEE Trans. Automat. Control, 35 (1990), pp. 1209–1217.
[8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, MD, 1996.
[9] P. L. Houtekamer and H. L. Mitchell, *Data assimilation using an ensemble Kalman filter technique*, Monthly Weather Rev., 126 (1998), pp. 796–811.
[10] I. M. Jaimoukha and E. M. Kasenally, *Krylov subspace methods for solving large Lyapunov equations*, SIAM J. Numer. Anal., 31 (1994), pp. 227–251.
[11] R. E. Kalman and J. E. Bertram, *Control system analysis and design via the "second method" of Lyapunov.* II. *Discrete-time systems*, Trans. ASME Ser. D. J. Basic Engrg., 82 (1960), pp. 394–400.
[12] S. R. Kolla, *Improved stability robustness bounds for digital control systems in state-space models*, Internat. J. Control, 64 (1996), pp. 991–994.
[13] P. Lancaster, *Explicit solutions of linear matrix equations*, SIAM Rev., 12 (1970), pp. 544–566.
[14] C. C. MacDuffee, *The Theory of Matrices*, Chelsea, New York, 1956.
[15] T. Mori, *On the relationship between the spectral radius and the stability radius of discrete systems*, IEEE Trans. Automat. Control, 35 (1990), p. 835.
[16] T. Mori, N. Fukuma, and M. Kuwahara, *Upper and lower bounds for the solution to the discrete Lyapunov matrix equation*, Internat. J. Control, 36 (1982), pp. 889–892.
[17] A. Ostrowski, *A quantitative formulation of Sylvester's law of inertia*, Proc. Natl. Acad. Sci. USA, 45 (1959), pp. 740–744.
[18] K. M. Przyluski, *The Lyapunov equation and the problem of stability for linear bounded discrete-time systems in Hilbert space*, Appl. Math. Optim., 6 (1980), pp. 97–112.
[19] J.-G. Sun, *Perturbation theory for algebraic Riccati equations*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 39–65.
[20] M. K. Tippett and D. Marchesin, *Upper bounds for the solution of the discrete algebraic Lyapunov equation*, Automatica, 35 (1999), pp. 1485–1489.
[21] L. N. Trefethen, *Pseudospectra of linear operators*, SIAM Rev., 39 (1997), pp. 383–406.
[22] L. N. Trefethen, A. E. Trefethen, and S. C. Reddy, *Hydrodynamic stability without eigenvalues*, Science, 261 (1993), pp. 578–584.
[23] J. S. Whitaker and P. D. Sardeshmukh, *A linear theory of extratropical synoptic Eddy statistics*, J. Atmospheric Sci., 55 (1998), pp. 237–258.