Associating the Invariant Subspaces of a Non-Normal Matrix with Transient Effects in its Matrix Exponential or Matrix Powers

Matthew G. Reuter∗
Department of Applied Mathematics and Statistics & Institute for Advanced Computational Science, Stony Brook University, Stony Brook, New York 11794, United States

Abstract
It is well known that the matrix exponential of a non-normal matrix can exhibit transient growth even when all eigenvalues of the matrix have negative real part, and similarly for the powers of the matrix when all eigenvalues have magnitude less than 1. Established conditions for the existence of these transient effects depend on properties of the entire matrix, such as the Kreiss constant, and can be laborious to use in practice. In this work we develop a relationship between the invariant subspaces of the matrix and the existence of transient effects in the matrix exponential or matrix powers. Analytical results are obtained for two-dimensional invariant subspaces and Jordan subspaces, with the former causing transient effects when the angle between the subspace’s constituent eigenvectors is sufficiently small. In addition to providing a finer-grained understanding of transient effects in terms of specific invariant subspaces, this analysis also enables geometric interpretations for the transient effects.

Keywords: Matrix exponential, matrix powers, transient behavior, non-normal matrices, numerical range
AMS Codes: 15A16, 15A60

1. Introduction
The matrix exponential $e^{Mt}$, for $M \in \mathbb{C}^{N \times N}$ and $t \in \mathbb{R}$ (often $t \geq 0$), appears in many contexts [1, 2], notably when solving a system of first-order, linear ordinary differential equations. The matrix powers $M^n$ ($n \geq 0$) also have wide applicability [2]; for instance in Markov processes or in condensed matter physics [3]. When $M$ is normal (i.e., $M$ has a complete set of orthogonal eigenvectors), its eigenvalues completely determine the behavior of $e^{Mt}$ and of $M^n$ [2]. That is, if all eigenvalues of $M$ have negative real part, then $\|e^{Mt}\|$ decays monotonically with $t$; likewise, $\|M^n\|$ decays with $n$ when all eigenvalues of $M$ have magnitude less than 1.

The situation is more complicated when $M$ is not normal [2]. The eigenvalues of $M$ correctly predict asymptotic behaviors as $t \to \infty$ or $n \to \infty$, but can qualitatively fail for small and intermediate values of $t$ or $n$. For example, there can be transient growth in $e^{Mt}$ such that $\|e^{Mt}\| > 1$ for some $0 < t \leq \infty$ even though all eigenvalues of $M$ have negative real part. Similar effects can be observed in $\|M^n\|$. Numerous applications of these transient effects — i.e., $\|e^{Mt}\| > 1$ or $\|M^n\| > 1$ — are discussed in [2]. Note that this definition of transience, although taken from [2], is different from “transience” in some applications, where it may mean that the system has not yet reached asymptotic or steady-state conditions. Throughout this work, transience will mean that there is transient growth in the matrix exponential or matrix powers when none would be predicted from the matrix’s eigenvalues.

Pseudospectral analyses [2] help predict the existence of transient growth, but have two limitations. First, they rely on quantities that can be laborious to obtain in practice, such as the condition number of the eigenvector matrix of $M$ or the Kreiss constant of $M$ (see chapter 14 of [2]). Second, they do not
usually provide sharp conditions for when $M$ will or will not display transient effects. Furthermore, better predictions of transient effects typically require the more computationally intensive properties of $M$, e.g., the Kreiss constant.

The goal of this work is to determine more practical conditions for the existence of transient effects in the matrix exponential or matrix powers of $M$. Our approach differs significantly from previous studies. Rather than relating “holistic” properties of $M$ (such as the Kreiss constant) to transient effects, we examine the $k$-dimensional ($k \geq 2$) invariant subspaces of $M$ and show that transient effects can be associated with invariant subspaces that satisfy certain conditions. This ultimately provides a more fine-grained explanation of transient effects and potentially introduces a new line of inquiry for characterizing non-normal matrices in terms of their invariant subspaces (see Thm. 4).

The main results are stated in Thms. 1 and 2 for $e^{Mt}$ and $M^n$, respectively, and are proven in section 4. Most of these conditions come from analytical results for the 2-dimensional invariant subspaces of $M$. The layout of this paper is as follows. First, section 2 introduces our notation and reviews key concepts from matrix analysis. Section 3 then derives conditions for transient effects when $M \in \mathbb{C}^{2 \times 2}$. These results are generalized to $M \in \mathbb{C}^{N \times N}$ in section 4, where they apply to the 2-dimensional invariant subspaces of $M$. Results for higher-dimensional Jordan subspaces are also presented. Additionally, there are several geometric interpretations for these results, which are discussed in section 5. We finally offer concluding remarks and ideas for future studies in section 6.

**Theorem 1** (Transient Effects in the Matrix Exponential). Let $M$ be a $N \times N$ complex matrix ($N \geq 2$) with all of its eigenvalues in the open left half of $\mathbb{C}$; that is, $\text{Re}(\lambda) < 0$ if $\lambda$ is an eigenvalue of $M$. Let $S$ be an invariant subspace of $M$. Then $e^{Mt}$ will have transient effects when at least one of the following conditions is satisfied.

1. $S$ is a $M$-dimensional ($M \geq 2$) Jordan subspace associated with defective eigenvalue $\lambda$ and
   \[
   \text{Re}(\lambda) > -\cos\left(\frac{\pi}{M+1}\right). \tag{1a}
   \]

2. $S$ is a 2-dimensional subspace spanned by eigenvectors associated with eigenvalues $\lambda_1$ and $\lambda_2$, $\lambda_1 \neq \lambda_2$, $\theta \in (0, \pi/2]$ is the angle between the eigenvectors [in the Hermitian sense of Eq. (3)], and
   \[
   \theta < \arctan\left(\frac{\left|\lambda_1 - \lambda_2\right|}{2\sqrt{\text{Re}(\lambda_1)\text{Re}(\lambda_2)}}\right). \tag{1b}
   \]

**Theorem 2** (Transient Effects in the Matrix Powers). Let $M$ be a nonsingular $N \times N$ complex matrix ($N \geq 2$) with all of its eigenvalues in the open unit circle of $\mathbb{C}$; that is, $0 < |\lambda| < 1$ if $\lambda$ is an eigenvalue of $M$. Let $S$ be an invariant subspace of $M$. Then $M^n$ will have transient effects when at least one of the following conditions is satisfied.

1. $S$ is a $M$-dimensional ($M \geq 2$) Jordan subspace.
2. $S$ is a 2-dimensional subspace spanned by eigenvectors associated with eigenvalues $\lambda_1$ and $\lambda_2$, $\lambda_1 \neq \lambda_2$, $\theta \in (0, \pi/2]$ is the angle between the eigenvectors [in the Hermitian sense of Eq. (3)], and
   \[
   \theta < \arctan\left(\frac{\ln(\lambda_1) - \ln(\lambda_2)}{2\sqrt{\ln|\lambda_1|\ln|\lambda_2|}}\right). \tag{2}
   \]

The branch of $\ln(z)$ can always be chosen to make the imaginary part of the numerator of Eq. (2) be in $[\pi, 2\pi)$.

2. Preliminaries and Notation

Several standard results from matrix analysis will be needed for the discussion and proofs of Thms. 1 and 2. We overview them in this section, which also introduces our notation. Throughout this section, we
assume $M \in \mathbb{C}^{N \times N}$. The angle between two vectors will be calculated in the Hermitian sense [5] with
\[
\cos(\theta) = \frac{|v_1 \cdot v_2|}{\|v_1\| \|v_2\|}
\] (3)
such that $\theta \in [0, \pi/2]$.

2.1. Schur Decomposition, Unitary Invariance, and Matrix Norm
Every $M$ is unitarily similar to an upper triangular matrix; that is, $M = QTQ^{-1}$, where $Q$ is unitary and $T$ is upper triangular. $T$ is called a Schur form of $M$ [6]. It is trivial to show that $M^n = QT^nQ^{-1}$ and, by use of the matrix exponential’s Taylor series, $e^{Mt} = Qe^{Tt}Q^{-1}$ [7]. Then, assuming the 2-norm or the Frobenius norm for the matrix norm, we get that $\|e^{Mt}\| = \|e^{Tt}\|$ and $\|M^n\| = \|T^n\|$ because these matrix norms are invariant to unitary transformations [6]. It is therefore sufficient to consider a Schur form of $M$ in our analyses.

2.2. Invariant Subspaces and Restrictions
A subspace $S \subseteq \mathbb{C}^N$ is an invariant subspace of $M$ if $Mx \in S$ for every $x \in S$ [4]. The invariant subspaces of $M$ are closely tied to the Jordan decomposition of $M$; every invariant subspace of $M$ is spanned by a set of eigenvectors and generalized eigenvectors of $M$. In this way, each invariant subspace of $M$ can be associated with eigenvalues of $M$. An invariant subspace of $M$ spanned by eigenvectors and generalized eigenvectors that all belong to the same Jordan chain is called a Jordan subspace of $M$. A 2- or higher-dimensional Jordan subspace of $M$ is thus associated with a sole defective eigenvalue of $M$.

We will also need the restriction of $M$ onto an invariant subspace $S$, denoted $M|_S : S \to S$ and defined by
\[
M|_S x = Mx
\]
for $x \in S$. Thus, $M|_S$ can be regarded as a $\dim(S) \times \dim(S)$ matrix. If $f$ is a well-behaved function, such as the exponential, then the Jordan form of $M$ and its relation to $S$ shows that [7]
\[
f(M)|_S = f(M|_S).
\]

2.3. Numerical Range and Numerical Abscissa
The numerical range of $M$ is the compact, convex subset of $\mathbb{C}$ given by [8]
\[
W(M) = \{x \cdot Mx : x \in \mathbb{C}^N, \|x\| = 1\}.
\] (4)
The numerical abscissa,
\[
\omega(M) = \max \{\text{Re}(z) : z \in W(M)\},
\] (5)
comes from the boundary of $W(M)$ and relates to transient effects in $e^{Mt}$ [2]. An example of the numerical range and numerical abscissa is displayed in Figure 1.

2.4. Transient Effects in the Matrix Exponential and Matrix Powers
Transient effects in the matrix exponential $e^{Mt}$ occur when all eigenvalues of $M$ have negative real part and $\|e^{Mt}\| > 1$ for some $t > 0$. Clearly, $\|e^{Mt}\| = 1$ when $t = 0$ such that a sufficient condition for transient behavior is
\[
\frac{d}{dt} \|e^{Mt}\| \bigg|_{t=0^+} > 0.
\]
From Eq. (14.2) in [2],
\[
\frac{d}{dt} \|e^{Mt}\| \bigg|_{t=0^+} = \omega(M),
\]
Thus, a sufficient condition for the existence of transient effects in $e^{Mt}$ is

$$\omega(M) > 0.$$  \hfill (6)$$

Transient effects in the matrix exponential are consequently related to the boundary of $W(M)$.

Other bounds for predicting the existence of transient effects in $e^{Mt}$ can be found in chapter 14 of [2], and involve (for example) the pseudospectra of $M$ or the Kreiss constant of $M$. These quantities are “holistic”, meaning they examine $M$ in its entirety. Our analyses in sections 3.1 and 3.2 will look at the relationship between specific invariant subspaces of $M$ and transient effects, thereby providing a more fine-grained analysis of transient effects.

Transient effects in $M^n$ are similar. In this case the eigenvalues of $M$ are less than one in magnitude and $\|M^n\| > 1$ for some $n > 0$. As before, a discussion of transient effects in the matrix powers can be found in chapter 14 of [2], which includes bounds that depend on holistic properties of $M$.

3. Analytical Results on $2 \times 2$ Matrices

Many of the conditions in Thms. 1 and 2 come from analytical results on transient effects in $2 \times 2$ matrices. This section develops these results. First in section 3.1 we discuss the numerical range of $2 \times 2$ matrices [8, 9]. Section 3.2 then parameterizes $\mathbb{C}^{2\times2}$ in a way that facilitates our analyses. The key parameter is the angle between the matrix’s eigenvectors (assuming the matrix is not defective). Finally, sections 3.3 and 3.4 derive sufficient conditions for the matrix exponential and matrix powers of a $2 \times 2$ matrix, respectively, to display transient behavior.

3.1. The Numerical Range

The numerical range of a $2 \times 2$ matrix $M$ is a (possibly-degenerate) ellipse with boundary given by Thm. 3 [8, 9]. The center of the ellipse is $\text{tr}(M)/2$, the eigenvalues of $M$ are the foci of the ellipse, and the major
and minor axes are stated in Lemma 1. Many of the ensuing results will come from analyzing the geometry of this ellipse, for which an example is displayed in Figure 1.

**Theorem 3** (Boundary of the Numerical Range of a $2 \times 2$ Matrix, Taken from Thm. 2 in [9]). In the complex plane $\mathbb{C} \cong \mathbb{R}^2$, the boundary curve of the numerical range of $M \in \mathbb{C}^{2 \times 2}$ is

$$
\begin{bmatrix}
  x - \text{Re} \left( \frac{\text{tr}(M)}{2} \right) \\
  y - \text{Im} \left( \frac{\text{tr}(M)}{2} \right)
\end{bmatrix}
S
\begin{bmatrix}
  x - \text{Re} \left( \frac{\text{tr}(M)}{2} \right) \\
  y - \text{Im} \left( \frac{\text{tr}(M)}{2} \right)
\end{bmatrix}
= \frac{\det(S)}{4},
$$

(7)

where $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$, $S = \begin{bmatrix} \|M_0\|_F^2 + 2 \text{Re}(\det(M_0)) \\ 2 \text{Im}(\det(M_0)) \end{bmatrix}$, and $\cdot$ is the Frobenius norm.

**Lemma 1** (Major and Minor Axes of the Numerical Range of a $2 \times 2$ Matrix, Taken from Thm. 1 in [9]). The major and minor axes of the elliptical numerical range of $M \in \mathbb{C}^{2 \times 2}$ are

$$
2\sqrt{\|M_0\|_F^2 + 2|\det(M_0)|} 
$$

(8a)

and

$$
2\sqrt{\|M_0\|_F^2 - 2|\det(M_0)|},
$$

(8b)

respectively, where $M_0 = M - (\text{tr}(M)/2)I$, and $\| \cdot \|_F$ is the Frobenius norm.

### 3.2. Parameterizing $2 \times 2$ Non-Normal Matrices

Before developing the results for $2 \times 2$ matrices, we first need to parameterize $\mathbb{C}^{2 \times 2}$ using quantities that will facilitate our analysis. Appealing to the Jordan form, there are three cases.

First is when the matrix is diagonalizable and has a degenerate eigenvalue. In this case the matrix is a scalar multiple of the identity matrix; it is normal and not of interest to this discussion. We will not consider this case further.

Second is when the matrix is diagonalizable but with distinct eigenvalues, $\lambda_1$ and $\lambda_2$. We will denote this matrix by $A$. From section 2.1 we can assume without loss of generality that $A$ is upper triangular:

$$
A = \begin{bmatrix}
  \lambda_1 & a \\
  0 & \lambda_2
\end{bmatrix},
$$

where $a, \lambda_1, \lambda_2 \in \mathbb{C}$ and $\lambda_1 \neq \lambda_2$. It is straightforward to see that $\begin{bmatrix} 1 & 0 \end{bmatrix}^T$ is a right eigenvector associated with $\lambda_1$ and that $\begin{bmatrix} a & \lambda_2 - \lambda_1 \end{bmatrix}^T$ is a right eigenvector associated with $\lambda_2$. Then, the angle between the eigenvectors [Eq. (3)] is

$$
\cos(\theta) = \frac{|a|}{\sqrt{|a|^2 + |\lambda_2 - \lambda_1|^2}},
$$

(9)

which yields

$$
|a| = |\lambda_2 - \lambda_1| \cot(\theta).
$$

Then, for $0 \leq \varphi < 2\pi$, we get

$$
A = \begin{bmatrix}
  \lambda_1 & e^{i\varphi} |\lambda_2 - \lambda_1| \cot(\theta) \\
  0 & \lambda_2
\end{bmatrix}.
$$

(10)

The angle between the eigenvectors can thus be used to parameterize the non-normality of $A$. This idea of examining the angle between eigenvectors comes from the condition number of an eigenvalue [10], which essentially considers the angle between the eigenvalue’s left and right eigenvectors.
Third is when the matrix is defective, with sole eigenvalue $\lambda$. This matrix will be called $D$. The Schur form is the $2 \times 2$ Jordan block because the generalized eigenvector can always be chosen to be orthogonal to the eigenvector. Thus,

$$D = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}. \quad (11)$$

Some useful properties of $A$ and $D$ are stated in the following Lemmas.

**Lemma 2.** The major and minor axes of $W(A)$ are

$$2 \sqrt{(|\lambda_1| + |\lambda_2|)^2 + |\lambda_1 - \lambda_2|^2 \cot^2(\theta)}$$

and

$$2 \sqrt{(|\lambda_1| - |\lambda_2|)^2 + |\lambda_1 - \lambda_2|^2 \cot^2(\theta)},$$

respectively.

**Proof.** The proof is trivial following from Eq. (10) and Lemma 1.

**Lemma 3.** The numerical abscissa of $A$ is

$$\omega(A) = \sqrt{|\lambda_1 - \lambda_2|^2 \left(1 + 2 \cot^2(\theta)\right) + \frac{\text{Re}(\lambda_1 + \lambda_2)}{2}}. \quad (13)$$

**Proof.** $S$ in Thm. 3 is

$$S = \begin{bmatrix} |\lambda_1 - \lambda_2|^2 \left(\frac{1}{2} + \cot^2(\theta)\right) - \frac{1}{2} \text{Re}(\lambda_1 - \lambda_2)^2 & -\frac{1}{2} \text{Im}(\lambda_1 - \lambda_2)^2 \\
-\frac{1}{2} \text{Im}(\lambda_1 - \lambda_2)^2 & |\lambda_1 - \lambda_2|^2 \left(\frac{1}{2} + \cot^2(\theta)\right) + \frac{1}{2} \text{Re}(\lambda_1 - \lambda_2)^2 \end{bmatrix}. \quad (13)$$

Then, recast Eq. (5) as a constrained optimization problem: Maximize $x$ subject to Eq. (7). This can be solved using standard techniques, e.g., Lagrange multipliers, which produces the result.

**Lemma 4.** The numerical abscissa of $D$ is $\omega(D) = \text{Re}(\lambda) + 1/2$.

**Proof.** The proof is very similar to that of Lemma 3. The matrix $S$ from Thm. 3 is the identity matrix such that the boundary of $W(D)$ is

$$(x - \text{Re}(\lambda))^2 + (y - \text{Im}(\lambda))^2 = \frac{1}{4}. \quad (13)$$

The right-most point of the boundary curve is trivially $(\text{Re}(\lambda) + 1/2, \text{Im}(\lambda))$; thus, $\omega(D) = \text{Re}(\lambda) + 1/2$. Note that a circular numerical range is expected for nilpotent operators [11] and $W(D)$ having a radius of $1/2$ is consistent with [12, 13]; see Thm. 5 below.

### 3.3. Matrix Exponential ($2 \times 2$ Matrices)

As discussed in section 2.4, $e^{Mt}$ has transient behavior when $\omega(M) > 0$. We can then examine the non-defective matrix $A$ in Eq. (10) and the defective matrix $D$ in Eq. (11) using this condition.

**Lemma 5.** Let $A$ be given as in Eq. (10) with $\lambda_1 \neq \lambda_2$ and $\text{Re}(\lambda_1), \text{Re}(\lambda_2) < 0$. Then, $e^{At}$ will have transient effects when

$$\theta < \arctan \left( \frac{|\lambda_1 - \lambda_2|}{2 \sqrt{\text{Re}(\lambda_1) \text{Re}(\lambda_2)}} \right).$$

**Proof.** Use Lemma 3 to solve $\omega(A) > 0$ for $\theta$. 

6
Corollary 1. Let $\mathbf{M} \in \mathbb{C}^{2 \times 2}$ have the form (Schur form)

$$\mathbf{M} = \begin{bmatrix} \lambda_1 & a \\ 0 & \lambda_2 \end{bmatrix},$$

with $\lambda_1 \neq \lambda_2$ and $\text{Re}(\lambda_1), \text{Re}(\lambda_2) < 0$. Then, $e^{\mathbf{Mt}}$ will have transient effects when $|a| > 2\sqrt{\text{Re}(\lambda_1)\text{Re}(\lambda_2)}$.

Proof. The earlier discussion [Eq. (9)] relates $a$ to $\theta$, the angle between the eigenvectors of $\mathbf{M}$. Using trigonometry,

$$\theta = \arctan\left(\frac{|\lambda_2 - \lambda_1|}{|a|}\right),$$

which provides the result when combined with Lemma 5.

Lemma 6. Let $\mathbf{D}$ be given as in Eq. (11) with $\text{Re}(\lambda) < 0$. Then, $e^{\mathbf{Dt}}$ will have transient effects when $\text{Re}(\lambda) > -1/2$.

Proof. Use Lemma 4 to solve $\omega(\mathbf{D}) > 0$ for $\text{Re}(\lambda)$.

3.4. Matrix Powers ($2 \times 2$ Matrices)

Let us now turn our attention to matrix powers $\mathbf{M}^n$. Because the existence of $\mathbf{M}^n$ is complicated for singular $\mathbf{M}$, we will restrict our attention to nonsingular $\mathbf{M}$. The results here follow straightforwardly from Lemmas 3 and 4 for the matrix exponential.

Lemma 7. Let $\mathbf{A}$ be nonsingular and given as in Eq. (10) with $\lambda_1 \neq \lambda_2$ and $0 < |\lambda_1|, |\lambda_2| < 1$. Then, $\mathbf{A}^n$ will have transient effects when

$$\theta < \arctan\left(\frac{|\ln(\lambda_1) - \ln(\lambda_2)|}{2\sqrt{|\ln(\lambda_1)| |\ln(\lambda_2)|}}\right).$$

Proof. Because $\mathbf{A}$ is not singular, we can write $\mathbf{A}^n = e^{n \ln(\mathbf{A})}$ by choosing the logarithmic branch cut to avoid both eigenvalues [7]. The result then trivially follows from Lemma 6 by using $\ln(\lambda_1)$ and $\ln(\lambda_2)$ as eigenvalues. Note that the logarithmic branch cut can always be chosen to make the imaginary part of $\ln(\lambda_1) - \ln(\lambda_2)$ be in $[-\pi, \pi]$.

Lemma 8. Let $\mathbf{D}$ be nonsingular and given as in Eq. (11) with $0 < |\lambda| < 1$. Then, $\mathbf{D}^n$ will have transient effects.

Proof. The proof is similar to that of Lemmas 4 and 7. Because $\lambda \neq 0$, $\mathbf{D}^n = e^{n \ln(\mathbf{D})}$. It can be shown that

$$\ln(\mathbf{D}) = \begin{bmatrix} \ln(\lambda) & \lambda^{-1} \\ 0 & \ln(\lambda) \end{bmatrix},$$

which leads to $\mathbf{S} = \mathbf{I}/|\lambda|^2$ in Thm. 3. Noting that $\text{Re}(\ln(\lambda)) = \ln |\lambda|$, the boundary of $W(\ln(\mathbf{D}))$ is $(x - \ln |\lambda|)^2 + (y - \text{arg}(\lambda))^2 = 1/(4|\lambda|^2)$. Thus,

$$\omega(\ln(\mathbf{D})) = \ln |\lambda| + \frac{1}{2|\lambda|}.$$ 

One can verify that $\omega(\ln(\mathbf{D})) > 0$ for $0 < |\lambda| < 1$, implying that there will always be transient effects.

Remarks. Although not proven, numerical experiments suggest that the conditions in Lemmas 5, 6, and 7 are also necessary for transient effects in $e^{\mathbf{Mt}}$ or $\mathbf{M}^n$ when $\mathbf{M} \in \mathbb{C}^{2 \times 2}$. Future work should be performed to confirm or refute this conjecture.
4. Higher-Dimensional Matrices (Proofs of Thms. 1 and 2)

The conditions for transient effects in 2 × 2 matrices that we developed in the previous section readily generalize to N × N matrices by considering the 2-dimensional invariant subspaces of \( M \in \mathbb{C}^{N \times N} \). Theorem 4 states the relationship.

**Theorem 4.** Let \( M \in \mathbb{C}^{N \times N} \) and let \( S \subseteq \mathbb{C}^N \) be an invariant subspace of \( M \). If \( e^{M|_S}t \) admits transient behavior, then so will \( e^{Mt} \). A similar statement holds for \((M|_S)^n\) and \(M^n\).

**Proof.** Transient effects in \( e^{M|_S}t \) require \( \|e^{M|_S}t\| > 1 \) for some \( t > 0 \). Because \( M|_S \) admits transient behavior, we know that \( 1 < \|e^{M|_S}t\| \) for some \( t > 0 \). Then, using the definition of the matrix norm,

\[
1 < \|e^{M|_S}t\| = \max_{\|x\|=1} \|e^{M|_S}tx\| = \max_{x \in S, \|x\|=1} \|e^{Mt}x\| \leq \max_{\|x\|=1} \|e^{Mt}x\| = \|e^{Mt}\|.
\]

Thus, if \( M|_S \) admits transient behavior in the matrix exponential, so must \( M \). Similar logic is used to show that if \((M|_S)^n\) has transient effects, then \( M^n \) will as well. \( \square \)

Using Thm. 4, Lemma 5 proves the second condition of Thm. 1 and Lemmas 8 and 7 prove the first and second conditions of Thm. 2, respectively. Lemma 6 shows the first condition of Thm. 1 for 2-dimensional Jordan subspaces. [12, 13] allow us to consider Jordan subspaces of arbitrary dimension because the numerical range is analytically known for these cases.

**Theorem 5 (Numerical Range of a Jordan Block [12, 13]).** Let \( J \) be a \( M \times M \) Jordan block \( (M \geq 2) \) with eigenvalue \( \lambda \). Then \( W(J) \) is all points \( z \in \mathbb{C} \) satisfying

\[
|z - \lambda| \leq \cos \left( \frac{\pi}{M + 1} \right).
\]

Consequently,

\[
\omega(J) = \Re(\lambda) + \cos \left( \frac{\pi}{M + 1} \right).
\]

The first condition in Thm. 1 trivially follows from Eq. (6) and Thm. 5. This completes the proofs of Thms. 1 and 2.

5. Geometric Interpretations

The results in section 3 largely came from analyzing a matrix \( M \in \mathbb{C}^{2 \times 2} \) in terms of the angle between its eigenvectors (assuming \( M \) is not defective). This alone leads to new geometric interpretations for transient effects in the matrix exponential and matrix powers of \( M \), but two other geometric observations can be made as well.

First regards the numerical detection of a defective eigenvalue. Due to numerical instabilities in computing the eigenvalues of a non-normal matrix [2] — especially a defective eigenvalue — it can be very difficult to distinguish (e.g.) two eigenvalues that are close to each other from a defective eigenvalue. A recent discussion of this issue appears in [14]. Let \( S \) be the 2-dimensional invariant subspace of \( M \) corresponding to these two eigenvalues. An examination of \( W(M|_S) \) provides insight into the situation. Without loss of generality [9], assume that \( \tr(M|_S) = 0 \) such that \( \lambda \) and \( -\lambda \) are the eigenvalues of \( M|_S \). From Lemma 2, the major and minor axes of \( W(M|_S) \) are \( 4|\lambda| \csc(\theta) \) and \( 4|\lambda| \cot(\theta) \), respectively. As \( \theta \to 0^+ \), meaning \( M|_S \) approaches a defective matrix, \( W(M|_S) \) will tend toward a circle. Examining the shape of the restricted numerical range for the invariant subspace may help determine if it is a Jordan subspace or not.

Second is a relation to antieigenvalue analyses [15]. When \( M \) is a self-adjoint, positive-definite matrix (and thus normal), its antieigenvector is the vector that is rotated the most by \( M \). The antieigenvalue is the
corresponding angle of rotation, which will be in $[0, \pi/2)$. If $M$ is $2 \times 2$, self-adjoint, and negative-definite with eigenvalues $\lambda_1$ and $\lambda_2$, its antieigenvalue is
\[
\arcsin \left( \frac{|\lambda_1 - \lambda_2|}{|\lambda_1 + \lambda_2|} \right),
\]
where the angle of rotation is measured in the Hermitian sense of Eq. (3). Compare this against the critical angle where $\omega(A) = 0$ in Lemma 5 which is
\[
\arctan \left( \frac{|\lambda_1 - \lambda_2|}{2\sqrt{\lambda_1 \lambda_2}} \right)
\]
when $\lambda_1, \lambda_2 \in \mathbb{R}$ and $\lambda_1, \lambda_2 < 0$. Some basic trigonometry reveals that these two angles are the same; the critical angle for transient effects in $e^{Mt}$ is exactly the antieigenvalue of a normal matrix with the same real eigenvalues. That is, transient effects appear in $e^{Mt}$ when the normal matrix with the same eigenvalues can rotate vectors more than the angle between the eigenvectors of $M$. Future work may consider generalizing this connection to matrices with complex eigenvalues, for which an antieigenvalue analysis is not well developed.

6. Concluding Remarks

In this work we have established a relationship (Thm. 4) between the invariant subspaces of a non-normal matrix $M \in \mathbb{C}^{N \times N}$ and transient behavior in $e^{Mt}$ ($t > 0$) and $M^n$ ($n > 0$). This then allowed us to derive sufficient conditions for transient effects in the exponential and powers of $M$, as stated in Thms. 1 and 2, respectively. Most of the conditions came from analytic results for 2-dimensional invariant subspaces of $M$, where the angle between the two associated eigenvectors is a key parameter. We should note that the conditions in Thms. 1 and 2 are sufficient but not necessary for transient effects; it is possible that transient effects can only be associated with higher-dimensional invariant subspaces.

Consequently, one immediate next step is to analyze higher-dimensional invariant subspaces for conditions on transient effects. Although not as straightforward as the case for $2 \times 2$ matrices, the numerical range of $3 \times 3$ matrices is discussed in [16]. Their results can be used with Eq. (6) to develop appropriate conditions. Along these lines, the Jordan principal angles [5] may be useful when generalizing the angle between eigenvectors to angles between invariant subspaces of arbitrary dimensions.

In a broader sense such investigations are examining structure within $W(M)$ [17, 18]. For instance, what is the lowest-dimensional invariant subspace $S$ of $M$ that satisfies $\omega(M|_S) > 0$, thereby inducing transient behavior? Ideas on the restricted numerical range [19, 20] may help answer this question. Finally, we end with a practical comment. All of our results make use of the Jordan decomposition of a non-normal $M$, which can be difficult to compute accurately [2]. Although the results are theoretically useful, additional work needs to be performed to establish their numerical utility.

Acknowledgments

I thank Jay Bardhan, Thorsten Hansen, Christopher DeGrendele, and Jonathan Kazakov for helpful conversations.

References

[1] C. Moler, C. Van Loan, Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later, SIAM Rev. 45 (1) (2003) 3–49.
[2] L. N. Trefethen, M. Embree, Spectra and Pseudospectra, The Behavior of Nonnormal Matrices and Operators, Princeton University Press, Princeton, NJ, USA, 2005.
[3] M. G. Reuter, A Unified Perspective of Complex Band Structure: Interpretations, Formulations, and Applications, J. Phys.: Condens. Matter 29 (2017) 033001.
[4] I. Gohberg, P. Lancaster, L. Rodman, Invariant Subspaces of Matrices with Applications, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2006.

[5] A. Galántai, C. J. Hegedűs, Jordan’s Principal Angles in Complex Vector Spaces, Numer. Lin. Alg. Appl. 13 (7) (2006) 589–598.

[6] G. H. Golub, C. F. Van Loan, Matrix Computations, 4th Edition, The Johns Hopkins University Press, Baltimore, MD, USA, 2013.

[7] N. J. Higham, Functions of Matrices, Theory and Computation, SIAM, Philadelphia, PA, 2008.

[8] R. A. Horn, C. R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, UK, 1991.

[9] F. Uhlig, The Field of Values of a Complex Matrix, An Explicit Description in the $2 \times 2$ Case, SIAM J. Alg. Disc. Math. 6 (4) (1985) 541–545.

[10] M. T. Heath, Scientific Computing: An Introductory Survey, 2nd Edition, McGraw-Hill, New York, NY, USA, 2002.

[11] M. T. Karaev, The Numerical Range of a Nilpotent Operator on a Hilbert Space, Proc. Am. Math. Soc. 132 (8) (2004) 2321–2326.

[12] U. Haagerup, P. de la Harpe, The Numerical Radius of a Nilpotent Operator on a Hilbert Space, Proc. Am. Math. Soc. 115 (2) (1992) 371–379.

[13] P. Y. Wu, A Numerical Range Characterization of Jordan Blocks, Lin. Multilin. Alg. 43 (1998) 351–361.

[14] Z. Zeng, Sensitivity and Computation of a Defective Eigenvalue, SIAM J. Matrix Anal. Appl. 37 (2) (2016) 798–817.

[15] K. Gustafson, Antieigenvalue Analysis, World Scientific, Hackensack, NJ, USA, 2012.

[16] D. S. Keeler, L. Rodman, I. M. Spitkovsky, The Numerical Range of $3 \times 3$ Matrices, Lin. Alg. Appl. 252 (1997) 115–139.

[17] F. Uhlig, An Inverse Field of Values Problem, Inv. Prob. 24 (5) (2008) 055019.

[18] C. F. Dunkl, P. Gawron, J. A. Holbrook, Z. Puchała, K. Życzkowski, Numerical Shadows: Measures and Densities on the Numerical Range, Lin. Alg. Appl. 434 (9) (2011) 2042–2080.

[19] K. C. Das, S. Majumdar, B. Sims, Restricted Numerical Range and Weak Convergence, J. Math. Phys. Sci. 21 (1) (1987) 35–42.

[20] P. Gawron, Z. Puchała, J. A. Miszczak, L. Skowronek, K. Życzkowski, Restricted Numerical Range: A Versatile Tool in the Theory of Quantum Information, J. Math. Phys. 51 (10) (2010) 102204.