Appendix D

Matrix Powers and Exponentials

The first order scalar linear differential equation $u'(t) = au(t)$ has the solution $u(t) = e^{at}u(0)$ and so

$$\begin{aligned} |u(t)| \to 0 \text{ as } t \to \infty \text{ for any } u(0) &\iff \operatorname{Re}(a) < 0, \\ |u(t)| \text{ remains bounded as } t \to \infty \text{ for any } u(0) &\iff \operatorname{Re}(a) \le 0, \\ |u(t)| \to \infty \text{ as } t \to \infty \text{ for any } u(0) \ne 0 &\iff \operatorname{Re}(a) > 0. \end{aligned} \tag{D.1}$$

The first order scalar linear difference equation $U^{n+1} = bU^n$ has the solution $U^n = b^n U^0$ and so

$$\begin{aligned} |U^n| \to 0 \text{ as } n \to \infty \text{ for any } U^0 &\iff |b| < 1, \\ |U^n| \text{ remains bounded as } n \to \infty \text{ for any } U^0 &\iff |b| \le 1, \\ |U^n| \to \infty \text{ as } n \to \infty \text{ for any } U^0 \ne 0 &\iff |b| > 1. \end{aligned} \tag{D.2}$$

For these first order scalar equations the behavior is very easy to predict. The purpose of this appendix is to review the extension of these results to the vector case, for linear differential equations $u'(t) = Au(t)$ and difference equations $U^{n+1} = BU^n$, where $u(t),\ U^n \in \mathbb{R}^m$ and $A,\ B \in \mathbb{R}^{m\times m}$, or more generally could be complex valued. This analysis relies on a good understanding of the material in Appendix C on eigendecompositions of matrices, and of the Jordan canonical form for the more interesting defective case.

In the scalar case there is a close relation between the results stated above in (D.1) and (D.2) for $u' = au$ and $U^{n+1} = bU^n$, respectively. If we introduce a time step $\Delta t = 1$ and let $U^n = u(n) = e^{an}u(0)$ be the solution to the ordinary differential equation (ODE) at discrete times, then $U^{n+1} = bU^n$ where $b = e^a$. Since the function e^z maps the left half-plane to the unit circle, we have

$$\begin{aligned} |b| < 1 &\iff \operatorname{Re}(a) < 0, \\ |b| = 1 &\iff \operatorname{Re}(a) = 0, \\ |b| > 1 &\iff \operatorname{Re}(a) > 0. \end{aligned} \tag{D.3}$$

Similarly, for the system cases we will see that boundedness of the solutions depends on eigenvalues of B lying inside the unit circle, or eigenvalues of A lying in the left half-plane, although the possibility of multiple eigenvalues makes the analysis somewhat more complicated. We will also see in Section D.4 that when the matrix involved is nonnormal the eigenvalues do not tell the full story—rapid transient growth in powers or the exponential can be observed even if there is eventual decay.

In the rest of this chapter we will use the symbol A as our general symbol in discussing both matrix powers and exponentials, but the reader should keep in mind that results on powers typically involve the unit circle, while results on exponentials concern the left half-plane.

D.1 The resolvent

Several of the bounds discussed below involve the resolvent of a matrix A, which is the complex matrix-valued function $(zI - A)^{-1}$. The domain of this function is the complex plane minus the eigenvalues of A, since the matrix $zI - A$ is noninvertible at these points. Bounds on the growth of A^n or e^{At} can be derived based on the size of $\|(zI - A)^{-1}\|$ over suitably chosen regions of the complex plane. These are generally obtained using the following representation of a function $f(A)$, the natural extension of the Cauchy integral formula to matrices.

Definition D.1. *If Γ is any closed contour in the complex plane that encloses the eigenvalues of A and if $f(z)$ is an analytic function within Γ, then the matrix $f(A)$ can be defined by*

$$f(A) = \frac{1}{2\pi i} \int_\Gamma f(z)(zI - A)^{-1}\, dz. \tag{D.4}$$

The value is independent of the choice of Γ.

For some functions $f(z)$, such as e^z, it may not be clear what $f(A)$ means for a matrix, and this can be used as the definition of $f(A)$. (Some other definitions of e^A are given in Section D.3, which are equivalent.) The representation (D.4) can also be useful computationally and is one approach to computing the matrix exponential; see, e.g., [53] and Section 11.6.1.

For the function $f(z) = z^n$ (or any polynomial in z) the meaning of $f(A)$ is clear since we know how to compute powers of a matrix, but (D.4) gives a different representation of the function that is sometimes useful in obtaining upper bounds.

D.2 Powers of matrices

Consider the linear difference equation

$$U^{n+1} = AU^n \tag{D.5}$$

for some iteration matrix A. The study of the asymptotic behavior of $\|U^n\|$ is important both in studying stability of finite difference methods and in studying convergence of iterative method for solving linear systems.

From (D.5) we obtain

$$\|U^{n+1}\| \le \|A\|\, \|U^n\|$$

in any norm and hence

$$\|U^n\| \le \|A\|^n \|U^0\|. \tag{D.6}$$

Alternatively, we can start with $U^{n+1} = A^n U^0$ to obtain

$$\|U^n\| \le \|A^n\| \, \|U^0\|. \tag{D.7}$$

Since $\|A^n\| \le \|A\|^n$ this again leads to (D.6), but in many cases $\|A^n\|$ is much smaller than $\|A\|^n$ and the form (D.7) may be more powerful.

If there exists any norm in which $\|A\| < 1$, then (D.6) shows that $\|U^n\| \to 0$ as $n \to \infty$. This is true not only in this particular norm but also in any other norm. Recall we are only considering matrix norms that are subordinate to some vector norm and that in a finite dimensional space of fixed dimension m all such norms are equivalent in the sense of (A.6). From these inequalities it follows that if $\|U^n\| \to 0$ in any norm, then it goes to zero in every equivalent norm, and similarly if $\|U^n\|$ blows up in some norm, then it blows up in every equivalent norm. Moreover, if $\|A\| = 1$ in some norm, then $\|U^n\|$ remains uniformly bounded as $n \to \infty$ in any equivalent norm.

From Theorem C.4 we thus obtain directly the following results.

Theorem D.1. (a) *Suppose $\rho(A) < 1$. Then $\|U^n\| \to 0$ as $n \to \infty$ in any vector norm.* (b) *Suppose $\rho(A) = 1$ and A has no defective eigenvalues of modulus 1. Then $\|U^n\|$ remains bounded as $n \to \infty$ in any vector norm, and there exists a norm in which $\|U^n\| \le \|U^0\|$.*

Note that result (a) holds even if A has defective eigenvalues of modulus $\rho(A)$ by choosing $\epsilon < 1 - \rho(A)$ in Theorem C.4(b).

If A is diagonalizable, then the results of Theorem D.1 can also be obtained directly from the eigendecomposition of A. Write $A = R\Lambda R^{-1}$, where

$$\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$$

is the diagonal matrix of eigenvalues of A and R is the matrix of right eigenvectors of A. Then the nth power of A is given by $A^n = R\Lambda^n R^{-1}$ and

$$\begin{aligned} \|U^n\| &\le \|A^n\| \, \|U^0\| \\ &\le \kappa(R)\rho(A)^n \, \|U^0\|, \end{aligned} \tag{D.8}$$

where $\kappa(R) = \|R\| \, \|R^{-1}\|$ is the condition number of R in whatever norm we are using. Note that if the value of $\kappa(R)$ is large, then $\|U^n\|$ could grow to large values before decaying, even in $\rho(A) < 1$ (see Example D.2).

If A is a normal matrix and we use the 2-norm, then $\kappa_2(R) = 1$ and we have the nice result that

$$\|U^n\|_2 \le \rho(A)^n \|U^0\|_2 \qquad \text{if } A \text{ is normal.} \tag{D.9}$$

For normal matrices, or for those close to normal, $\rho(A)$ gives a good indication of the behavior of $\|U^n\|_2$. For matrices that are far from being normal, the spectral radius may give a poor indication of how $\|U^n\|$ behaves. The nonnormal case is considered further in Section D.4.

More detailed information about the behavior of $\|U^n\|$ can be obtained by decomposing U^0 into eigencomponents. Still assuming A is diagonalizable, we can write

$$U^0 = W_1^0 r_1 + W_2^0 r_2 + \cdots + W_m^0 r_m = RW^0,$$

where $r_1, \ldots, r_m$ are the eigenvectors of A and the vector W^0 is given by $W^0 = R^{-1}U^0$. Multiplying U^0 by A multiplies each r_p by λ_p and so

$$U^n = A^n U^0 = W_1^0 \lambda_1^n r_1 + W_2^0 \lambda_2^n r_2 + \cdots + W_m^0 \lambda_m^n r_m = R\Lambda^n W^0. \tag{D.10}$$

For large n this is dominated by the terms corresponding to the largest eigenvalues, and hence the norm of this vector is proportional to $\rho(A)^n$, at least for generic initial data. This also shows that if $\rho(A) < 1$, then $U^n \to 0$, while if $\rho(A) > 1$ we expect U^n to blow up.

Note that for certain initial data $\|U^n\|$ may behave differently than $\rho(A)^n$, at least in exact arithmetic. For example, if U^0 is void in the dominant eigencomponents, then these terms will be missing from (D.10), and the asymptotic growth or decay rate will be different. In particular, it could happen that $\rho(A) > 1$ and yet $\|U^n\| \to 0$ for special data U^0 if some eigenvalues of A have modulus less than 1 and U^0 contains only these eigencomponents. However, this generally is not relevant for the stability and convergence issues considered in this book, where arbitrary initial data must be considered. Moreover, even if U^0 is void of some eigencomponents, rounding errors introduced computationally in each iteration $U^{n+1} = AU^n$ will typically be random and will contain all eigencomponents. The growing modes may start out at the level of rounding error, but if $\rho(A) > 1$ they will grow exponentially and eventually dominate the solution so that the asymptotic behavior will still be governed by $\rho(A)^n$.

If A is defective, then we cannot express an arbitrary initial vector U^0 as a linear combination of eigenvectors. However, using the Jordan canonical form $A = RJR^{-1}$, we can still write $U^0 = RW^0$, where $W^0 = R^{-1}U^0$. The nonsingular matrix R now contains principal vectors as well as eigenvectors, as discussed in Section C.3.

Example D.1. Consider the iteration $U^{n+1} = AU^n$, where A is the 3×3 matrix from Example C.4, having a single eigenvalue λ with geometric multiplicity 1. If we decompose

$$U^0 = W_1^0 r_1 + W_2^0 r_2 + W_3^0 r_3,$$

then multiplying by A gives

$$\begin{aligned} U^1 &= W_1^0 \lambda r_1 + W_2^0 (r_1 + \lambda r_2) + W_3^0 (r_2 + \lambda r_3) \\ &= (W_1^0 \lambda + W_2^0) r_1 + (W_2^0 \lambda + W_3^0) r_2 + W_3^0 \lambda r_3. \end{aligned} \tag{D.11}$$

Repeating this shows that U^n has the form

$$U^n = A^n U^0 = p_1(\lambda) r_1 + p_2(\lambda) r_2 + W_3^0 \lambda^n r_3,$$

where $p_1(\lambda)$ and $p_2(\lambda)$ are polynomials in λ of degree n. It can be shown that

$$|p_1(\lambda)| \le n^2 \lambda^n \|W^0\|_2, \qquad |p_2(\lambda)| \le n\lambda^n \|W^0\|_2,$$

so that there are now algebraic terms (powers of n) in the asymptotic behavior in addition to the exponential terms (λ^n). More generally, as we will see below, a Jordan block of order k gives rise to terms of the form $n^{k-1}\lambda^n$.

If $|\lambda| > 1$, then the power n^{k-1} is swamped by the exponential growth and the algebraic term is unimportant. If $|\lambda| < 1$, then n^{k-1} grows algebraically but λ^n decays exponentially and the product decays, $n^{k-1}\lambda^n \to 0$ as $n \to \infty$ for any $k > 1$.

The borderline case $|\lambda| = 1$ is where this algebraic term makes a difference. In this case λ^n remains bounded but $n^{k-1}\lambda^n$ does not. Note how this relates to the results of Theorem D.1. If $\rho(A) < 1$, then $\|A^n\| \to 0$ even if A has defective eigenvalues of modulus $\rho(A)$, since the exponential decay overpowers the algebraic growth. However, if $\rho(A) = 1$, then the boundedness of $\|A^n\|$ depends on whether there are defective eigenvalues of modulus 1. If so, then $\|A^n\|$ grows algebraically (but not exponentially).

Recall also from Theorem C.4 that in this latter case we can find, for any $\epsilon > 0$, a norm in which $\|A\| < 1 + \epsilon$. This implies that $\|A^n\| < (1 + \epsilon)^n$. There may be growth, but we can make the exponential growth rate arbitrarily slow. This is consistent with the fact that in this case we have only algebraic growth. Exponential growth at rate $(1 + \epsilon)^n$ eventually dominates algebraic growth n^{k-1} no matter how small ϵ is, for any k.

To determine the algebraic growth factors for the general case of a defective matrix, we can use the fact that if $A = RJR^{-1}$ then $A^n = RJ^nR^{-1}$, and we can compute the nth power of the Jordan matrix J. Recall that J is a block diagonal matrix with Jordan blocks on the diagonal, and powers of J simply consist of powers of these blocks. For a single Jordan block (C.9) of order k,

$$J(\lambda, k) = \lambda I_k + S_k,$$

where S_k is the shift matrix (C.10). Powers of $J(\lambda, k)$ can be found using the binomial expansion and the fact that I_k and S_k commute,

$$\begin{aligned} J(\lambda,k)^n &= (\lambda I_k + S_k)^n \\ &\lambda^n I_k + n\lambda^{n-1}S_k + \binom{n}{2}\lambda^{n-2}S_k^2 + \binom{n}{3}\lambda^{n-3}S_k^3 \\ &+ \cdots + \binom{n}{n-1}\lambda S_k^{n-1} + S_k^n. \end{aligned} \tag{D.12}$$

Note that for $j < k$, S_k^j is the matrix with 1's along the jth superdiagonal and zeros everywhere else. For $j \geq k$, S_k^j is the zero matrix. So the series in expression (D.12) always terminates after at most k terms, and when $n > k$ reduces to

$$\begin{aligned} J(\lambda,k)^n = \lambda^n I_k + n\lambda^{n-1}S_k + \binom{n}{2}\lambda^{n-2}S_k^2 + \binom{n}{3}\lambda^{n-3}S_k^3 \\ + \cdots + \binom{n}{k-1}\lambda^{n-k+1}S_k^{k-1}. \end{aligned} \tag{D.13}$$

Since $\binom{n}{j} = O(n^j)$ as $n \to \infty$, we see that $J(\lambda,k)^n = P(n)\lambda^n$, where $P(n)$ is a matrix-valued polynomial of degree $k - 1$.

For example, returning to Example C.4, where $k = 3$, we have

$$J(\lambda, 3)^n = \lambda^n I_3 + n\lambda^{n-1} S_3 + \frac{n(n-1)}{2}\lambda^{n-2} S_3^2$$
$$= \begin{bmatrix} \lambda^n & n\lambda^{n-1} & \frac{n(n-1)}{2}\lambda^{n-2} \\ 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \lambda^n \end{bmatrix}. \tag{D.14}$$

This shows that $\|J^n\| \approx n^2\lambda^n$, the same result obtained in Example D.1,

If A is not normal, i.e., if $A^H A \neq AA^H$, then $\|A\|_2 > \rho(A)$ and $\rho(A)^n$ may not give a very good indication of the behavior of $\|A^n\|$. If A is diagonalizable, then we have

$$\|A^n\|_2 \leq \|R\|_2 \|\Lambda^n\|_2 \|R^{-1}\|_2 \leq \kappa_2(R)\rho(A)^n, \tag{D.15}$$

but if $\kappa_2(R)$ is large, then this may not be useful, particularly for smaller n. This does give information about the asymptotic behavior as $n \to \infty$, but in practice it may tell us little or nothing about how $\|A^n\|_2$ is behaving for the finite values of n we care about in a particular computation. See Section D.4 for more about the nonnormal case.

D.2.1 Solving linear difference equations

Matrix powers arise naturally when iterating with the first order linear difference equation (D.5). In studying linear multistep methods we need the general solution to an rth order linear difference equation of the form

$$a_0 V^n + a_1 V^{n+1} + a_2 V^{n+2} + \cdots + V^{n+r} = 0 \quad \text{for } n \geq 0. \tag{D.16}$$

We have normalized the equation by assuming $a_r = 1$. One approach to solving this is given in Section 6.4.1, and here we give an alternative based on converting this rth order equation into a first order system of r equations of the form (D.5) and then applying the results just found for matrix powers.

We can rewrite (D.16) as a system of equations by introducing

$$U_1^n = V^n, \quad U_2^n = V^{n+1}, \quad \ldots, \quad U_r^n = V^{n+r-1}. \tag{D.17}$$

Then (D.16) takes the form

$$U^{n+1} = CU^n, \tag{D.18}$$

where

$$C = \begin{bmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & 0 & 1 & & \\ & & & \ddots & \ddots & \\ & & & & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & & -a_{r-1} \end{bmatrix}. \tag{D.19}$$

This matrix is called the *companion matrix* for the difference equation. The general solution to (D.18) is $U^n = C^n U^0$, where

$$U^0 = [V^0,\ V^1,\ \ldots,\ V^{r-1}]^T$$

is the vector consisting of the r initial values required by (D.16). The solution can be expressed in terms of the eigenvalues of C.

It can be shown that the characteristic polynomial of C is

$$p(\lambda) = \det(\lambda I - C) = a_0 + a_1\lambda + \cdots + a_{r-1}\lambda^{r-1} + \lambda^r. \tag{D.20}$$

Call the roots of this polynomial $\zeta_1, \ldots, \zeta_r$, as in Section 6.4.1. Let $C = RJR^{-1}$ be the Jordan canonical form of the matrix C. Then $U^n = RJ^nR^{-1}U^0$ and each element of U^n (and in particular $V^n = U_1^n$) is a linear combination of elements of J^n.

If the roots are distinct, then C is diagonalizable, and so the general solution of (D.16) is

$$V^n = c_1\zeta_1^n + c_2\zeta_2^n + \cdots + c_r\zeta_r^n, \tag{D.21}$$

where the coefficients c_j depend on the initial data.

It can be shown that repeated eigenvalues of a companion matrix always have geometric multiplicity 1, regardless of their algebraic multiplicity. An eigenvalue ζ_j has a one-dimensional space of eigenvectors spanned by $[1, \zeta_j, \zeta_j^2, \ldots, \zeta_j^{r-1}]^T$. From the form (D.13) for powers of a Jordan block, we see that if ζ_j is a repeated root of algebraic multiplicity k, then the general solution of the difference equation (D.16) includes terms of the form $\zeta_j^n, n\zeta_j^n, \ldots, n^{k-1}\zeta_j^n$. We can conclude that the general solution to the rth order difference equation (D.16) will be bounded as $n \to \infty$ only if the roots of the characteristic polynomial all lie inside the unit circle, with no repeated roots of magnitude 1. This is known as the *root condition* and is used in the stability analysis of linear multistep methods; see Definition 6.2.

D.2.2 Resolvent estimates

In our analysis of powers of A we have used the eigenstructure of A. An alternative approach is to use the resolvent $(zI - A)^{-1}$ and the expression (D.4) for the function $f(z) = z^n$. For example, we can obtain an alternative proof of part (a) of Theorem D.1 as follows. If $\rho(A) < 1$, then we choose as our contour Γ a circle of radius $1-\epsilon$, where $\epsilon > 0$ is chosen small enough that $\rho(A) < 1 - \epsilon$. The eigenvalues of A then lie inside Γ and $zI - A$ is invertible for z on Γ and so $\|(zI-A)^{-1}\|$ is a bounded periodic function of z and hence attains some maximum value $C(A,\Gamma)$ on Γ. Now consider (D.4) with $f(z) = z^n$,

$$A^n = \frac{1}{2\pi i}\int_\Gamma z^n(zI - A)^{-1}\,dz. \tag{D.22}$$

Taking norms and using the fact that $|z^n| = (1-\epsilon)^n$ and $\|(zI - A)^{-1}\| \le C(A,\Gamma)$ for z on Γ, and that Γ has length $2\pi(1-\epsilon) < 2\pi$, we obtain

$$\begin{aligned}\|A^n\| &\le \frac{1}{2\pi}\int_\Gamma |z^n|\,\|(zI - A)^{-1}\|\,dz \\ &< C(A,\Gamma)(1-\epsilon)^n.\end{aligned} \tag{D.23}$$

Since $(1-\epsilon)^n \to 0$ as $n \to \infty$ this proves Theorem D.1(a).

Note that we also get a uniform bound on $\|A^n\|$ that holds for all n,

$$\|A^n\| < C(A,\Gamma).$$

If A is normal then in the 2-norm we have $\|A^n\|_2 \le 1$ from (D.9), but for nonnormal matrices there can be transient growth in $\|A^n\|$ before it eventually decays, and the resolvent gives one approach to bounding this potential growth

Note that the value $C(A,\Gamma)$ depends on the matrix A and may be very large if $\rho(A)$ is close to 1. The contour Γ must lie between the spectrum of A and the unit circle for the argument above, and $\|(zI - A)^{-1}\|$ is large near the spectrum of A. In the study of numerical methods we are often concerned not just with a single matrix A but with a family of matrices arising from different discretizations, and proving stability results often requires proving *uniform power boundedness* of the family, i.e., that there is a single constant C such that $\|A^n\| \le C$ for all matrices in the family. This is more difficult than proving that a single matrix is power bounded, and it is the subject of the Kreiss matrix theorem discussed in Section D.6.

Note that this resolvent proof does not extend directly to prove part (b) of Theorem D.1, the case where A has nondefective eigenvalues on the unit circle. In this case the contour Γ must lie outside the unit circle, at least near these eigenvalues, for (D.22) to hold. In this case it is necessary to investigate how the product

$$(|z| - 1)\|(zI - A)^{-1}\| \tag{D.24}$$

behaves for $|z| > 1$. Here the resolvent norm is multiplied by a factor that vanishes as $|z| \to 1$ and so there is hope that the product will be bounded even if the resolvent is blowing up.

In fact it can be shown that if $\rho(A) < 1$ or if $\rho(A) = 1$ with only nondefective eigenvalues on the unit circle (i.e., if either of the conditions of Theorem D.1 holds), then the product (D.24) will be uniformly bounded for all z outside the unit circle. The supremum is called the *Kreiss constant* for the matrix A, denoted by $\mathcal{K}(A)$,

$$\mathcal{K}(A) = \sup_{|z|>1} (|z| - 1)\|(zI - A)^{-1}\|. \tag{D.25}$$

An alternative definition based on the ϵ-pseudospectral radius is discussed in Section D.5.

For an idea of how this can be used, again consider the first line of (D.23) but now take Γ to be a circle of radius $1+\epsilon$ with $\epsilon > 0$. We have $\|(zI - A)^{-1}\| \le \mathcal{K}(A)/\epsilon$ on this circle, and so

$$\begin{aligned} \|A^n\| &\le \frac{1}{2\pi} \int_\Gamma |1+\epsilon|^n \, \frac{\mathcal{K}(A)}{\epsilon}\, dz \\ &= \frac{1}{\epsilon}(1+\epsilon)^{n+1}\mathcal{K}(A), \end{aligned} \tag{D.26}$$

since the circle has radius $2\pi(1+\epsilon)$. This bound holds for any $\epsilon > 0$. It appears to allows exponential growth for any fixed ϵ, but we are free to choose a different value of ϵ for each n, and taking $\epsilon = 1/(n+1)$, for example, gives

$$\|A^n\| \le (n+1)e\mathcal{K}(A) \qquad \text{for all } n \ge 0.$$

Unfortunately, this still allows algebraic growth and so does not prove the theorem.

It fact, it can be shown (by a more complicated argument based on the resolvent that will not be presented here) that a bound of the required form holds,

$$\|A^n\| \le me\mathcal{K}(A) \qquad \text{for all } n \ge 0, \tag{D.27}$$

where m is the dimension of the matrix. This shows that for a fixed matrix A all powers remain bounded provided its Kreiss constant is finite (which in turn is true if and only if $\rho(A) \le 1$ with no defective eigenvalues on the unit circle).

The bound in (D.27) is sharp in a sense made precise in Section 18 of [92] and was the end product of a long sequence of weaker bounds proved over the years (as recounted in [92]). This result was proved by Spijker [81] as a corollary to a more general result on the arclength of the image of the unit circle under a rational function.

Resolvent estimates can also be used to obtain a *lower bound* on the transient growth of $\|A^n\|$; see (D.46).

The bound (D.27) is a major part of the proof of the Kreiss matrix theorem (see Section D.6): a family of matrices is uniformly power bounded if (and only if) there is a uniform bound on the Kreiss constants $\mathcal{K}(A)$ of all matrices in the family.

D.3 Matrix exponentials

Now consider the linear system of m ordinary differential equations $u'(t) = Au(t)$, where $A \in \mathbb{R}^{m\times m}$ (or more generally $A \in \mathbb{C}^{m\times m}$). The nondiagonalizable (defective) case will be considered in Section D.4. When A is diagonalizable we can solve this system by changing variables to $v = R^{-1}u$ and multiplying both sides of the ODE by R^{-1} to obtain

$$R^{-1}u'(t) = R^{-1}AR \cdot R^{-1}u(t)$$

or

$$v'(t) = \Lambda v(t).$$

This is a decoupled set of m scalar equations $v_j'(t) = \lambda_j v_j(t)$ (for $j = 1, 2, \dots, m$) with solutions $v_j(t) = e^{\lambda_j t}v_j(0)$. Let $e^{\Lambda t}$ denote the matrix

$$e^{\Lambda t} = \text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_m t}). \tag{D.28}$$

Then we have $v(t) = e^{\Lambda t}v(0)$ and hence

$$u(t) = Rv(t) = Re^{\Lambda t}R^{-1}u(0),$$

so

$$u(t) = e^{At}u(0), \tag{D.29}$$

where

$$e^{At} = Re^{\Lambda t}R^{-1}. \tag{D.30}$$

This gives one way to define the matrix exponential e^{At}, at least in the diagonalizable case. The Cauchy integral of Definition D.1 is another. Yet another way to define it is by the Taylor series

$$e^{At} = I + At + \frac{1}{2!}A^2t^2 + \frac{1}{3!}A^3t^3 + \cdots = \sum_{j=0}^{\infty}\frac{1}{j!}A^jt^j. \tag{D.31}$$

This is often useful, particularly when t is small. The definitions (D.30) and (D.31) agree in the diagonalizable case since all powers A^j have the same eigenvector matrix R, and so (D.31) gives

$$\begin{aligned} e^{At} &= R\left[I + \Lambda t + \frac{1}{2!}\Lambda^2 t^2 + \frac{1}{3!}\Lambda^3 t^3 + \cdots\right] R^{-1} \\ &= \mathrm{R}\, e^{\Lambda t} R^{-1}, \end{aligned} \tag{D.32}$$

resulting in (D.30). To go from the first to the second line of (D.32), note that it is easy to verify that the Taylor series applied to the diagonal matrix Λ is a diagonal matrix of Taylor series, each of which converge to the corresponding diagonal element of $e^{\Lambda t}$, the value $e^{\lambda_j t}$.

Note that the matrix e^{At} has the same eigenvectors as A and its eigenvalues are $e^{\lambda_j t}$. To investigate the behavior of $u(t) = e^{At}u(0)$ as $t \to \infty$, we need only look at the real part of each eigenvalue λ_j. If none of these are greater than 0, then the solution will remain bounded as $t \to \infty$ (assuming still that A is diagonalizable) since $|e^{\lambda_j t}| \le 1$ for all j.

It is useful to introduce the *spectral abscissa* $\alpha(A)$, defined by

$$\alpha(A) = \max_{1\le j\le m} \mathrm{Re}(\lambda_j). \tag{D.33}$$

Then $u(t)$ remains bounded provided $\alpha(A) \le 0$ and $u(t) \to 0$ as $t \to \infty$ if $\alpha(A) < 0$.

Note that for integer values of $t = n$, we have $e^{An} = (e^A)^n$ and $\rho(e^A) = e^{\alpha(A)}$, so this result is consistent with what was found in the last section for matrix powers.

If A is not diagonalizable, then the case $\alpha(A) = 0$ is more subtle, as is the case $\rho(A) = 1$ for matrix powers. If A has a defective eigenvalue λ with $\mathrm{Re}(\lambda) = 0$, then the solution may still grow, although with polynomial growth in t rather than exponential growth.

When A is not diagonalizable, we can still write the solution to $u' = Au$ as $u(t) = e^{At}u(0)$, but we must reconsider the definition of e^{At}. In this case the Jordan canonical form $A = RJR^{-1}$ can be used, yielding

$$e^{At} = \mathrm{R}e^{Jt} R^{-1}.$$

If J has the block structure (C.12) then e^{Jt} is also block diagonal,

$$e^{Jt} = \begin{bmatrix} e^{J(\lambda_1,k_1)t} & & & \\ & e^{J(\lambda_2,k_2)t} & & \\ & & \ddots & \\ & & & e^{J(\lambda_s,k_s)t} \end{bmatrix}. \tag{D.34}$$

For a single Jordan block the Taylor series expansion (D.31) can be used in conjunction with the expansion (D.12) for powers of the Jordan block. We find that

$$
\begin{aligned}
e^{J(\lambda,k)t} &= \sum_{j=0}^{\infty} \frac{t^j}{j!}\left[\lambda^j I + j\lambda^{j-1}S_k + \binom{j}{2}\lambda^{j-2}S_k^2 + \cdots + \binom{j}{j-1}\lambda S_k^{j-1} + S_k^j\right] \\
&= \sum_{j=0}^{\infty} \frac{t^j}{j!}\lambda^j I + t\sum_{j=1}^{\infty}\frac{t^{j-1}}{(j-1)!}\lambda^{j-1}S_k + \frac{t^2}{2!}\sum_{j=2}^{\infty}\frac{t^{j-2}}{(j-2)!}\lambda^{j-2}S_k^2 + \cdots \\
&= e^{\lambda t}I + te^{\lambda t}S_k + \frac{t^2}{2!}e^{\lambda t}S_k^2 + \cdots + \frac{t^{(k-1)}}{(k-1)!}e^{\lambda t}S_k^{k-1} \\
&= \begin{bmatrix}
e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \cdots & & \frac{t^{(k-1)}}{(k-1)!}e^{\lambda t} \\
 & e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \cdots & \\
 & & e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \\
 & & & \ddots & \ddots & \\
 & & & & & \\
 & & & & & e^{\lambda t}
\end{bmatrix}.
\end{aligned}
\tag{D.35}
$$

Here we have used the fact that

$$\frac{1}{j!}\binom{j}{p} = \frac{1}{p!}\,\frac{1}{(j-p)!}$$

and the fact that $S_k^q = 0$ for $q \geq k$. The $k \times k$ matrix $e^{J(\lambda,k)t}$ is an upper triangular Toeplitz matrix with elements $d_0 = e^{\lambda t}$ on the diagonal and $d_j = \frac{t^j}{j!}e^{\lambda t}$ on the jth superdiagonal for $j = 1, 2, \ldots, k-1$.

We see that the situation regarding boundedness of e^{At} is exactly analogous to what we found for matrix powers A^n. If $\mathrm{Re}(\lambda) < 0$, then $t^j e^{\lambda t} \to 0$ despite the t^j factor, but if $\mathrm{Re}(\lambda) = 0$, then $t^j e^{\lambda t}$ grows algebraically. We obtain the following theorem, analogous to Theorem D.1.

Theorem D.2. *Let $A \in \mathbb{C}^{m\times m}$ be an arbitrary square matrix, and let $\alpha(A) = \max \mathrm{Re}(\lambda)$ be the spectral abscissa of A. Let $u(t) = e^{At}u(0)$ solve $u'(t) = Au(t)$. Then*

(a) *if $\alpha(A) < 0$, then $\|u(t)\| \to 0$ as $t \to \infty$ in any vector norm.*

(b) *if $\alpha(A) = 0$ and A has no defective eigenvalues with* $\mathrm{Re}(\lambda) = 0$*, then $\|u(t)\|$ remains bounded in any norm, and there exists a vector norm in which $\|u(t)\| \leq \|u(0)\|$ for all $t \geq 0$.*

If A is normal, then

$$\|e^{At}\|_2 = \|e^{\Lambda t}\|_2 = e^{\alpha(A)t} \tag{D.36}$$

since $\|R\|_2 = 1$. In this case the spectral abscissa gives precise information on the behavior of e^{At}. If A is nonnormal, then the behavior of $e^{\alpha(A)t}$ may not give a good indication of the behavior of e^{At} (except asymptotically for t sufficiently large), just as $\rho(A)^n$ may not give a good indication of how powers $\|A^n\|$ behave if A is not normal. See Section D.4 for some discussion of this case.

D.3.1 Solving linear differential equations

In Section D.2.1 we saw that the rth order linear difference equation (D.16) can be rewritten as a first order system (D.18) and solved using matrix powers. A similar approach can be used to convert the homogeneous rth order constant coefficient linear differential equation

$$a_0 v(t) + a_1 v'(t) + a_2 v''(t) + \cdots + a_{r-1} v^{(r-1)}(t) + v^{(r)}(t) = 0 \tag{D.37}$$

to a first order system of r equations and solve this using the matrix exponential. We follow the approach of Example 5.1 and introduce

$$u_1(t) = v(t), \quad u_2(t) = v'(t), \ldots, \; u_r(t) = v^{(r-1)}(t). \tag{D.38}$$

The equation (D.37) becomes $u'(t) = Cu(t)$, where C is the companion matrix (D.19). The solution is

$$u(t) = e^{Ct} u(0) = \mathrm{R}e^{Jt} R^{-1} u(0),$$

where $u(0)$ is the initial data (consisting of v and its first $r-1$ derivatives at time $t = 0$), and $C = RJR^{-1}$ is the Jordan canonical form of C. If the roots of the characteristic polynomial (D.20) are distinct, then $v(t)$ is a linear combination of the exponentials $e^{\zeta_j t}$. If there are repeated roots, then they are defective, and examining the expression (D.35) we see that a root of algebraic multiplicity k leads to terms of the form $e^{\zeta_j t}, \; te^{\zeta_j t}, \; \ldots, \; t^{k-1} e^{\zeta_j t}$. We can thus conclude that in general solutions to the linear differential equations (D.37) are bounded for all time only if the roots of the characteristic polynomial are in the left half-plane, with no repeated roots on the imaginary axis.

D.4 Nonnormal matrices

We have seen that if a matrix A has the Jordan form $A = RJR^{-1}$ (where J may be diagonal), then we can bound powers and the matrix exponential by the following expressions:

$$\begin{aligned} &\|A^n\| \le \kappa(R) \|J^n\| \;\text{ in general and} \\ &\|A^n\| \le \kappa(R) \rho(A)^n \;\text{ if } A \text{ is diagonalizable,} \end{aligned} \tag{D.39}$$

$$\begin{aligned} &\|e^{At}\| \le \kappa(R) \|e^{Jt}\| \;\text{ in general and} \\ &\|e^{At}\| \le \kappa(R) e^{\alpha(A)t} \;\text{ if } A \text{ is diagonalizable.} \end{aligned} \tag{D.40}$$

Here $\rho(A)$ and $\alpha(A)$ are the spectral radius and spectral abscissa, respectively. If A is defective, then J is not diagonal and algebraic growth terms can arise from the $\|J^n\|$ or $\|e^{Jt}\|$ factors.

If A is not normal, then even in the nondefective case the above bounds may not be very useful if $\kappa(R)$ is large. In particular, with nonnormal matrices matrix powers or exponentials can exhibit exponential *growth* during an initial transient phase (i.e., for n or t small enough), even if the bounds guarantee eventual exponential decay. Moreover, in these cases a small perturbation of the matrix may result in a matrix whose powers or exponential is not bounded.

This growth can be disastrous in terms of stability, particularly since in practice most interesting problems are nonlinear and often the matrix problems we consider are obtained from a local linearization of the problem. Transient growth or instability of perturbed problems can easily lead to nonlinear instabilities in the original problem.

A related problem is that in practice we often are dealing with coefficient variable problems, where the matrix changes in each iteration. This issue is discussed more in Section D.7. In this section we continue to consider a fixed matrix A and explore some upper and lower bounds on powers and exponentials in the nonnormal case.

D.4.1 Matrix powers

If A is a normal matrix, $A^H A = AA^H$, then A is diagonalizable and the eigenvector matrix R can be chosen as an unitary matrix, for which $R^H R = I$ and $\kappa(R) = 1$. (We assume the 2-norm is always used in this section.) In this case $\|A\| = \rho(A)$ and $\|A^n\| = (\rho(A))^n$, so the eigenvalues of A give precise information about the rate of growth or decay of $\|A^n\|$, and similarly for the matrix exponential.

If A is not normal, then $\|A\| > \rho(A)$ and $(\rho(A))^n$ may not give a very good indication of the behavior of $\|A^n\|$ even in the diagonalizable case. From (D.39) we know that $\|A^n\|$ eventually decays at worst like $(\rho(A))^n$ for large enough n, but if $\kappa(R)$ is huge, then there can be enormous growth of $\|A^n\|$ before decay sets in. This is easily demonstrated with a simple example.

Example D.2. Consider the nonnormal matrix

$$A = \begin{bmatrix} 0.8 & 100 \\ 0 & 0.9 \end{bmatrix}. \tag{D.41}$$

This matrix is diagonalizable and the spectral radius is $\rho(A) = 0.9$. We expect $\|A^n\| \sim C(0.9)^n$ for large n, but for smaller n we observe considerable growth before the norm begins to decay. For example, starting with $U^0 = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ and computing $U^n = A^n U^0$ for $n = 1,\ 2,\ \dots$ we find

$$U^0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad U^1 = \begin{bmatrix} 100 \\ 0.9 \end{bmatrix}, \quad U^2 = \begin{bmatrix} 170 \\ 0.81 \end{bmatrix}, \quad U^3 = \begin{bmatrix} 217 \\ 0.729 \end{bmatrix}, \quad \dots .$$

We have the bound $\|U^n\|_2 \le \kappa_2(R)(0.9)^n \|U^0\|_2$ but in this case

$$R = \begin{bmatrix} 1 & 1 \\ 0 & 0.001 \end{bmatrix}, \qquad R^{-1} = \begin{bmatrix} 1 & -1000 \\ 0 & 1000 \end{bmatrix},$$

so $\kappa_2(R) = 2000$. Figure D.1 shows $\|U^n\|_2$ for $n = 1,\ \dots,\ 30$ along with the bound.

Clearly this example could be made much more extreme by replacing $a_{22} = 100$ by a larger value. Larger matrices can exhibit similar growth before decay even if all the elements of the matrix are modest.

To gain some insight into the behavior of $\|A^n\|$ for a general matrix A, recall that the 2-norm of A is

$$\|A\| = \sqrt{\rho(A^H A)},$$

Hence

$$\begin{aligned} \|A^n\| &= [\rho((A^n)^H A^n)]^{1/2} \\ &= [\rho(A^H A^H \cdots A^H AA \cdots A)]^{1/2}. \end{aligned} \tag{D.42}$$

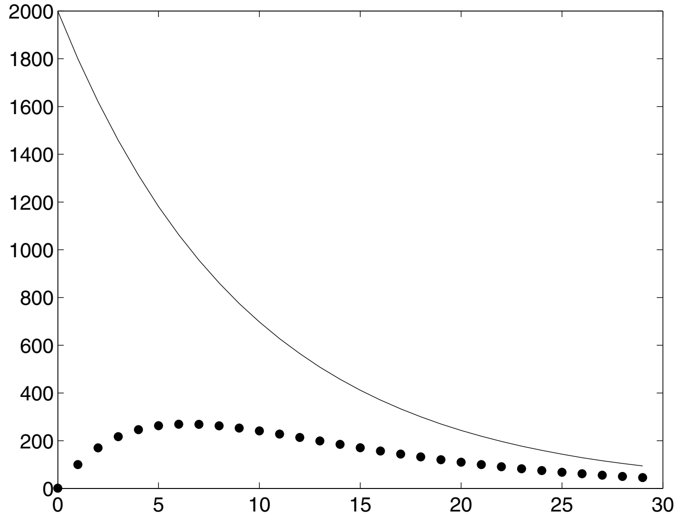

Figure D.1. *The points show* $\|U^n\|_2$ *for Example* D.2 *and the line shows the upper bound* $2000(0.9)^n$.

If A is normal, then $A^H A = AA^H$ and the terms in this product of $2n$ matrices can be rearranged to give

$$\|A^n\| = [\rho((A^H A)^n)]^{1/2} = (\rho(A^H A))^{n/2} = (\rho(A))^n = \|A\|^n. \tag{D.43}$$

If A is not normal, then we cannot rearrange the product, and in general we have

$$(\rho(A))^n \le \|A^n\| \le \|A\|^n. \tag{D.44}$$

If $\rho(A) < 1 < \|A\|$, as is often the case for nonnormal matrices of interest, then the lower bound is decaying exponentially to zero while the upper bound is growing exponentially. If we expect to see transient growth followed by decay, as illustrated, for example, in Figure D.1, then neither of these bounds tells us anything about how much growth is expected before the decay sets in. We would like to have lower and upper bounds on

$$\mathcal{P}(A) = \sup_{n\ge 0} \|A^n\|. \tag{D.45}$$

The matrix A is said to be *power bounded* if $\mathcal{P}(A) < \infty$, and of course a necessary condition for this is that the eigenvalues of A lie in the unit disk, with no defective eigenvalues on the disk.

Lower and upper bounds on $\mathcal{P}(A)$ can be written very concisely in terms of the Kreiss constant (D.25):

$$\mathcal{K}(A) \le \mathcal{P}(A) \le em\mathcal{K}(A) \tag{D.46}$$

for any $A \in \mathbb{C}^{m\times m}$. The upper bound has already been discussed in Section D.1. The lower bound is easier to obtain; it says that if $\|A^n\| \le C$ for all n, then

$$(|z| - 1)\|(zI - A)^{-1}\| \le C.$$

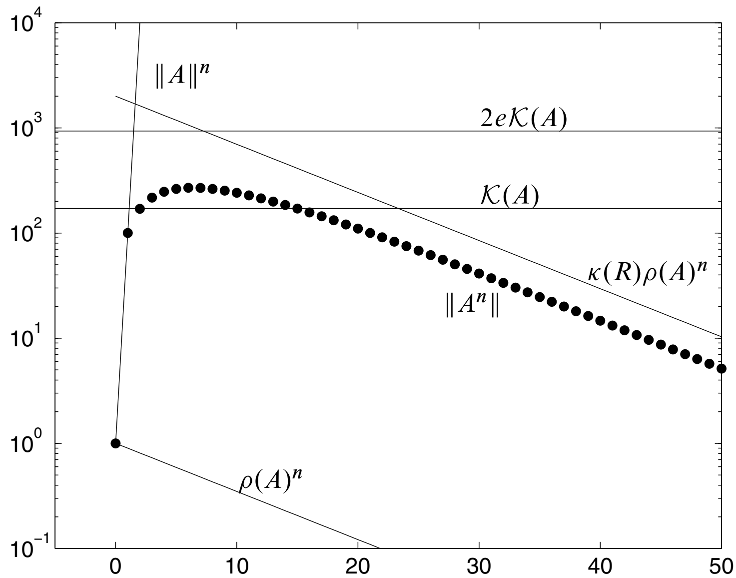

Figure D.2. *The norm of the matrix power, $\|A^n\|$ plotted on a logarithmic scale for the nonnormal matrix* (D.41). *Also shown are the lower bound $\rho(A)^n$ and the upper bounds $\|A\|^n$ and $\kappa(R)\rho(A)^n$, as well as the value of the Kreiss constant $\mathcal{K}$ and $2e\mathcal{K}$ that give lower and upper bounds on* $\sup_{t\geq 0}\|A^n\|$. *Note that $\|A^n\|$ initially grows like $\|A\|^n$ and asymptotically decays like $\rho(A)^n$.*

To prove this note that $(zI - A)^{-1}$ has the series expansion

$$(zI - A)^{-1} = z^{-1}(I - z^{-1}A)^{-1} = z^{-1}\left(I + (z^{-1}A) + (z^{-1}A)^2 + (z^{-1}A)^3 + \cdots\right).$$

Taking norms and using $\|A^n\| \leq C$ gives

$$\|(zI - A)^{-1}\| \leq |z^{-1}|\left(1 + |z^{-1}| + |z^{-1}|^2 + |z^{-1}|^3 + \cdots\right)C = \frac{C}{|z| - 1}.$$

Figure D.2 shows the various bounds discussed above along with $\|A^n\|$ for the nonnormal matrix (D.41), this time on a logarithmic scale. For this matrix $\rho(A) = 0.9$, $\|A\| \approx 100$, and the Kreiss constant is $\mathcal{K}(A) = 171.5$.

D.4.2 Matrix exponentials

For the matrix exponential there are similar bounds (Theorem 18.5 in [92]):

$$\mathcal{K}_e(A) \leq \sup_{t\geq 0}\|e^{At}\| \leq em\mathcal{K}_e(A), \tag{D.47}$$

where the Kreiss constant with respect to the matrix exponential $\mathcal{K}_e(A)$ is defined by

$$\mathcal{K}_e(A) = \sup_{\mathrm{Re}(z)>0} \mathrm{Re}(z)\|(zI - A)^{-1}\|. \tag{D.48}$$

This measures how the resolvent of A behaves near the imaginary axis, which is the stability boundary for the matrix exponential.

It is also possible to derive upper and lower bounds on the norm of the matrix exponential as functions of t, analogous to (D.44) for powers of the matrix. For an arbitrary matrix A these take the form

$$e^{\alpha(A)t} \le \|e^{At}\| \le e^{\omega(A)t} \qquad \text{for all } t \ge 0, \tag{D.49}$$

where $\alpha(t)$ is the spectral abscissa (D.33) and $\omega(A)$ is the *numerical abscissa* defined by

$$\omega(A) = \frac{1}{2}\rho(A^H + A). \tag{D.50}$$

Note that $A^H + A$ is always hermitian and has real eigenvalues, but they may be positive even if $\alpha(A) < 0$. In general,

$$\alpha(A) \le \omega(A). \tag{D.51}$$

Also note that for any vector u,

$$\begin{aligned} \mathrm{Re}(u^H Au) &= \frac{1}{2}(u^H Au + u^H A^H u) \\ &= u^H \left(\frac{1}{2}(A^H + A)\right) u \\ &\le \omega(A) u^H u, \end{aligned} \tag{D.52}$$

since the Rayleigh quotient $u^H Bu/u^H u$ is always bounded by $\rho(B)$ for any hermitian matrix B. Another way to characterize $\omega(A)$ is as the maximum value that the real part of $u^H Au$ can take over any unit vector u. The set

$$W(A) = \{z \in \mathbb{C} :\ z = u^H Au \text{ for some } u \text{ with } \|u\|_2 = 1\} \tag{D.53}$$

is called the *numerical range* or *field of values* of the matrix A, and

$$\omega(A) = \max_{z \in W(A)} \mathrm{Re}(z). \tag{D.54}$$

For a normal matrix $W(A)$ is the convex hull of the eigenvalues, but for a nonnormal matrix it may be larger.

The bound (D.52) is of direct interest in studying the matrix exponential and can be used to prove the upper bound in (D.49) as follows. Suppose $u'(t) = Au(t)$ and consider

$$\begin{aligned} \frac{d}{dt}(u^H u) &= (u')^H u + u^H u' \\ &= u^H A^H u + u^H Au \\ &= 2\mathrm{Re}(u^H Au) \\ &\le 2\omega(A) u^H u. \end{aligned} \tag{D.55}$$

In other words, using the 2-norm,

$$\frac{d}{dt}\left(\|u(t)\|^2\right) \le 2\omega(A)\|u(t)\|^2, \tag{D.56}$$

which in turn implies

$$\frac{d}{dt}\|u(t)\| \le \omega(A)\|u(t)\|, \tag{D.57}$$

or, dividing by $\|u(t)\|$,

$$\frac{d}{dt}\log(\|u(t)\|) \le \omega(A). \tag{D.58}$$

Integrating gives

$$\|u(t)\| \le e^{\omega(A)t}\|u(0)\|. \tag{D.59}$$

Since $u(t) = e^{At}u(0)$, we have

$$\|e^{At}u(0)\| \le e^{\omega(A)t}\|u(0)\| \tag{D.60}$$

for any vector $u(0)$ and hence the matrix norm of e^{At} is bounded as in (D.49).

For a normal matrix A, $\alpha(A) = \omega(A)$ and (D.49) reduces to

$$\|e^{At}\| = e^{\alpha(A)t} = e^{\omega(A)t} \qquad \text{for all } t \ge 0. \tag{D.61}$$

In the scalar case, for a complex number A, the spectral abscissa and numerical abscissa are both equal to the real part of A, so each can be viewed as a generalization of the real part.

Finally, it is sometimes useful to investigate the initial transient growth of $\|e^{At}\|$ at $t = 0$. This can be determined by differentiating $\|e^{At}\|$ with respect to t and evaluating at $t = 0$. The result is

$$\begin{aligned} \frac{d}{dt}\|e^{At}\|\bigg|_{t=0} &= \lim_{k\to 0}\frac{\|e^{Ak}\| - 1}{k} \\ &= \lim_{k\to 0}\frac{\|I + Ak\| - 1}{k}. \end{aligned} \tag{D.62}$$

We can compute

$$\begin{aligned} \|I + kA\| &= \left[\rho\left((I + A^H k)(I + Ak)\right)\right]^{1/2} \\ &= \left[\rho(I + (A + A^H)k + A^H A k^2)\right]^{1/2} \\ &= \left[1 + \rho(A + A^H)k + O(k^2)\right]^{1/2} \\ &= 1 + \frac{1}{2}\rho(A + A^H)k + O(k^2), \end{aligned} \tag{D.63}$$

and so

$$\frac{d}{dt}\|e^{At}\|\bigg|_{t=0} = \lim_{k\to 0}\frac{\|I + Ak\| - 1}{k} = \frac{1}{2}\rho(A + A^H) = \omega(A). \tag{D.64}$$

Hence we expect

$$\|e^{At}\| = e^{\omega(A)t} + o(t) \quad \text{as } t \to 0. \tag{D.65}$$

From (D.49) we know that $e^{\omega(A)t}$ is an upper bound on the norm, but this shows that for small t we will observe transient growth at this rate.

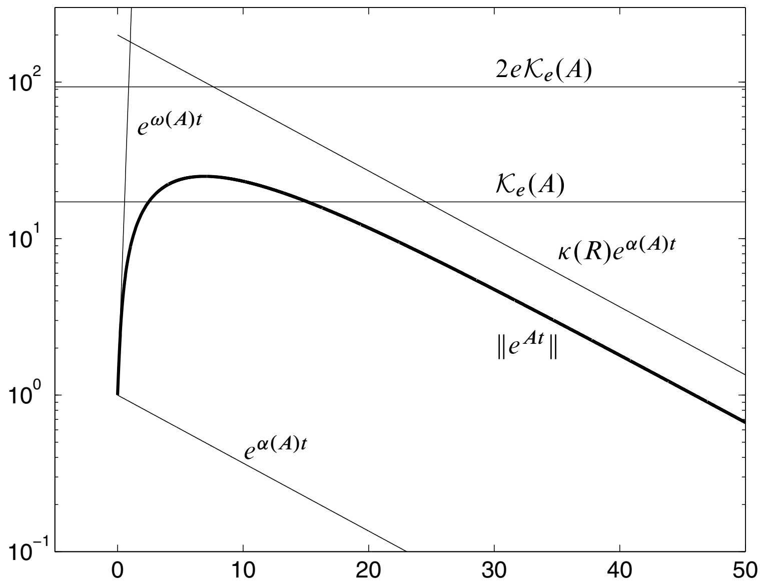

Figure D.3. *The norm of the matrix exponential* $\|e^{At}\|$ *plotted on a logarithmic scale for the nonnormal matrix* (D.66). *Also shown are the lower bound* $e^{\alpha(A)t}$ *and the upper bounds* $e^{\omega(A)t}$ *and* $\kappa(R)e^{\alpha(A)t}$*, as well as the value of the Kreiss constant* $\mathcal{K}_e$ *and* $2e\mathcal{K}_e$ *that give lower and upper bounds on* $\sup_{t\geq 0}\|e^{At}\|$. *Note that* $\|e^{At}\|$ *initially grows like* $e^{\omega(A)t}$ *and asymptotically decays like* $e^{\alpha(A)t}$.

For example, consider the nonnormal matrix

$$A = \begin{bmatrix} -0.2 & 10 \\ 0 & -0.1 \end{bmatrix}. \tag{D.66}$$

Figure D.3 shows $\|e^{At}\|$ as a function of t on a logarithmic scale. The initial growth has slope $\omega(A) = 5.15$ and the eventual decay has slope $\alpha(A) = -0.1$, which follows from (D.40). The Kreiss constant for this matrix is $\mathcal{K}_e(A) = 17.17$.

The quantity $\omega(A)$ is sometimes defined as

$$\omega(A) = \lim_{k\to 0} \frac{\|I + Ak\| - 1}{k} \tag{D.67}$$

and in the ODE literature often goes by the name of the *logarithmic norm of A*. Of course it is not really a norm since it can be negative, but this terminology is motivated by inequalities such as (D.58).

D.5 Pseudospectra

Various tools have been developed to better understand the behavior of matrix powers or exponentials in the case of nonnormal matrices. One powerful approach is to investigate the *pseudospectra* of the matrix. Roughly speaking, this is the set of eigenvalues of all "nearby" matrices. For a highly nonnormal matrix a small perturbation to the matrix can give a very large change in the eigenvalues of the matrix. For example, perturbing the matrix (D.41) to

$$\tilde{A} = \begin{bmatrix} 0.8 & 100 \\ 0.001 & 0.9 \end{bmatrix} \tag{D.68}$$

changes the eigenvalues from $\{0.8,\ 0.9\}$ to $\{0.53,\ 1.17\}$. A perturbation to A of magnitude 10^{-3} leads to an eigenvalue that is greater than 1. Since A is so close to a matrix $\tilde{A}$ for which $\tilde{A}^n$ blows up as $n \to \infty$, it is perhaps not so surprising that A^n exhibits initial growth before decaying. We say that the value $z = 1.17$ lies in the ϵ-pseudospectrum σ_ϵ of A for $\epsilon = 10^{-3}$.

The eigenvalues of A are isolated points in the complex plane where $(zI - A)$ is singular. We know that if any of these points lies outside the unit circle, then A^n blows up. The idea of pseudospectral analysis is to expand these isolated points to larger regions, the pseudospectra σ_ϵ, for some small ϵ, and see whether these pseudospectra extend beyond the unit circle.

There are various equivalent ways to define the ϵ-pseudospectrum of a matrix. Here are three.

Definition D.3. *For each $\epsilon \geq 0$, the ϵ-pseudospectrum $\sigma_\epsilon(A)$ of A is the set of numbers $z \in \mathbb{C}$ satisfying any one of the following equivalent conditions:*

(a) *z is an eigenvalue of $A + E$ for some matrix E with $\|E\| < \epsilon$,*
(b) *$\|Au - zu\| \leq \epsilon$ for some vector u with $\|u\| = 1$, or*
(c) *$\|(zI - A)^{-1}\| \geq \epsilon^{-1}$.*

In all these conditions the 2-norm is used (although the ideas can be extended to a general Banach space). Condition (a) is the easiest to understand and the one already illustrated above by example: z is an ϵ-pseudoeigenvalue of A if it is a genuine eigenvalue of some ϵ-sized perturbation of A. Condition (b) says that z is an ϵ-pseudoeigenvalue if there is a unit vector that is almost an eigenvector for this z. Condition (c) relates pseudoeigenvalues to the resolvent $(zI - A)^{-1}$, which we have already seen plays a role in obtaining bounds on the behavior of matrix powers and exponentials. The value z is an ϵ-pseudoeigenvalue of A if the resolvent is sufficiently large at z. This fits with the notion of expanding the singular points λ_i into regions σ_ϵ where $zI - A$ is nearly singular. (Note that by convention we set $\|(zI - A)^{-1}\| = \infty$ if z is an eigenvalue of A.)

The ϵ-pseudospectral radius $\rho_\epsilon(A)$ and ϵ-pseudospectral abscissa $\alpha_\epsilon(A)$ can be defined in the natural way as the maximum absolute value and maximum real part of any ϵ-pseudoeigenvalue of A, respectively. The Kreiss constants (D.25) and (D.48) can then be expressed in terms of pseudospectra as

$$\mathcal{K}(A) = \sup_{\epsilon > 0} \frac{\rho_\epsilon(A) - 1}{\epsilon}, \qquad \mathcal{K}_e(A) = \sup_{\epsilon > 0} \frac{\alpha_\epsilon(A)}{\epsilon}. \tag{D.69}$$

Hence the Kreiss constants can be viewed as a measure of how far the pseudospectra of A extend outside the unit circle or into the right half-plane.

The MATLAB package `eigtool` developed by Wright [104] provides tools for computing and plotting the pseudospectra of matrices and also quantities such as the ϵ-pseudospectral radius and ϵ-pseudospectral abscissa. See the book by Trefethen and Embree [92] for an in-depth discussion of pseudospectra with many examples of their use.

D.5.1 Nonnormality of a Jordan block

In Section D.2 we found an explicit expression for powers of a Jordan block, and we see that in addition to terms of the form λ^n, a block of order k has terms of order $n^k\lambda^n$ in its nth power. This clearly exhibits transient growth even in $\lambda < 1$. It is interesting to further investigate the Jordan block as an example of a highly nonnormal matrix.

For this discussion, let

$$J_\epsilon = \begin{bmatrix} c & 1 & & & \\ & c & 1 & & \\ & & \ddots & & \\ & & & c & 1 \\ \epsilon & & & & c \end{bmatrix} \in \mathbb{R}^{k\times k} \tag{D.70}$$

with the entries not shown all equal to 0, so that for $\epsilon = 0$, J_0 is a Jordan block of the form (C.9) with all k of its eigenvalues at c.

If $\epsilon > 0$ on the other hand, the characteristic equation is

$$(\lambda - c)^k - \epsilon = 0$$

and the eigenvalues are

$$\lambda_p = c + \epsilon^{1/k} e^{2\pi i p/k}, \qquad p = 1,\ 2,\ \ldots,\ k. \tag{D.71}$$

The eigenvalues are now equally spaced around a circle of radius $\epsilon^{1/k}$ centered at $z = c$ in the complex plane.

Note that if k is large, $\epsilon^{1/k}$ will be close to 1 even for very small ϵ. For example, if $k = 1000$ and $\epsilon = 10^{-16}$, then $\epsilon^{1/k} \approx 0.96$. So although the eigenvalues of J_0 are all at c, a perturbation on the order of the machine round-off will blast them apart to a circle of radius nearly 1 about this point. For large k the ϵ-pseudospectrum of J_0 tends to fill up this circle, even for very small ϵ.

A similar matrix arises when studying the upwind method for advection, in which case k corresponds to the number of grid points and can easily be large. An implication of this nonnormality in stability analysis is explored in Section 10.12.1.

D.6 Stable families of matrices and the Kreiss matrix theorem

So far we have studied the behavior of powers of a single matrix A. We have seen that if the eigenvalues of A are inside the unit circle, then the powers of A are uniformly bounded,

$$\|A^n\| \le C \quad \text{for all } n, \tag{D.72}$$

for some constant C. If A is normal, then it fact $\|A^n\| \leq \|A\|$. Otherwise $\|A^n\|$ may initially grow, perhaps to a very large value if the deviation from normality is large, but will eventually decay and hence some bound of the form (D.72) holds.

In studying the stability of discretizations of differential equations, we often need to consider not just a single matrix but an entire family of matrices. A particular discretization with mesh width h and/or time step k leads to a particular matrix A, but to study stability and prove convergence we need to let $h, k \to 0$ and study the whole family of resulting matrices. This is quite difficult to study in general because typically the dimensions of the matrices involved is growing as we refine the grid. However, at least in simple cases we can use von Neumann analysis to decouple the system into Fourier modes, each of which leads to a system of fixed dimension (the number of equations in the original differential equation). As we refine the grid we obtain more modes and the matrices involved may depend explicitly on h and k as well as on the wave number, but the matrices all have fixed dimension and it is this case that we consider here. (In Sections 9.6 and 10.5 we consider von Neumann analysis applied to scalar problems, in which case proving stability only requires studying powers of the scalar amplification factor g for each wave number, and powers of a scalar are uniformly bounded if and only if $|g| \leq 1$. The considerations of this section come into play if von Neumann analysis is applied to a system of differential equations.)

Let $\mathcal{F}$ represent a family of matrices, say all the amplification matrices for different wave numbers that arise from discretizing a particular differential equation with different mesh widths. We say that $\mathcal{F}$ is *uniformly power bounded* if there is a constant $C > 0$ such that (D.72) holds for all matrices $A \in \mathcal{F}$. The bound must be uniform in both A and n, i.e., a single constant for all matrices in the family and all powers.

If $\mathcal{F}$ consists of only normal matrices and if $\rho(A) \leq 1$ for all $A \in \mathcal{F}$, then the family is uniformly power bounded and (D.72) holds in general with $C = 1$.

When the matrices are not normal it can be more difficult to establish such a bound. Obviously a necessary condition is that $\rho(A) \leq 1$ for all $A \in \mathcal{F}$ and that any eigenvalues of modulus 1 must be nondefective. If this condition fails for any $A \in \mathcal{F}$, then that particular matrix will fail to be power bounded and so the family cannot be. However, this condition is not sufficient—even if each matrix is power bounded they may not be uniformly so. For example, the infinite family of matrices

$$A_\epsilon = \begin{bmatrix} 1-\epsilon & 1 \\ 0 & 1-\epsilon \end{bmatrix} \tag{D.73}$$

for $\epsilon > 0$ are all individually power bounded but not uniformly power bounded. We have

$$A_\epsilon^n = \begin{bmatrix} (1-\epsilon)^n & n(1-\epsilon)^{n-1} \\ 0 & (1-\epsilon)^n \end{bmatrix},$$

and the off-diagonal term can be made arbitrarily large for large n by choosing ϵ small enough.

One fundamental result on power boundedness of matrix families is the Kreiss matrix theorem.

Theorem D.4. *The following conditions on a family $\mathcal{F}$ of matrices are equivalent:*

(a) *There exists a constant C such that $\|A^n\| \le C$ for all n and for all $A \in \mathcal{F}$. (The family is power bounded.)*

(b) *There exists a constant C_1 such that, for all $A \in \mathcal{F}$ and all $z \in \mathbb{C}$ with $|z| > 1$, the resolvent $(zI - A)^{-1}$ exists and is bounded by*

$$\|(zI - A)^{-1}\| \le \frac{C_1}{|z| - 1}. \tag{D.74}$$

In other words, the $\mathcal{K}(A) \le C_1$ for all $A \in \mathcal{F}$, where $\mathcal{K}(A)$ is the Kreiss constant (D.25).

(c) *There exist constants C_2 and C_3 such that for each $A \in \mathcal{F}$ a nonsingular matrix S exists such that*

(i) $\|S\| \le C_2, \quad \|S^{-1}\| \le C_2,$
(ii) $B = S^{-1}AS$ *is upper triangular with off-diagonal elements bounded by*

$$|b_{ij}| \le C_3 \min(1 - |b_{ii}|,\ 1 - |b_{jj}|). \tag{D.75}$$

(Note that the diagonal elements of b are the eigenvalues of A.)

(d) *There exists a constant C_4 such that for each $A \in \mathcal{F}$ a positive definite matrix G exists with*

$$\begin{aligned} &C_4^{-1} I \le G \le C_4 I, \\ &A^H G A \le G. \end{aligned} \tag{D.76}$$

In condition (d) we say that two Hermitian matrices A and B satisfy $A \le B$ if $u^H Au \le u^H Bu$ for any vector u. This condition can be rewritten in a more familiar form as follows:

(d′) *There exists a constant C_5 so that for each $A \in \mathcal{F}$ there is a nonsingular matrix T such that*

$$\|A\|_T \le 1 \ \text{ and } \ \kappa(T) \le C_5. \tag{D.77}$$

Here the T-norm of A is defined as in (C.35) in terms of the 2-norm,

$$\|A\|_T = \|T^{-1}AT\|.$$

Condition (d′) is related to (d) by setting $G = T^{-H}T^{-1}$, and (d′) states that we can define a set of norms, one for each $A \in \mathcal{F}$, for which the norm of A is less than 1 and therefore

$$\|A^n\|_T \le 1 \quad \text{for all } n \ge 0.$$

From this we can obtain uniform power boundedness by noting that

$$\|A^n\| \le \kappa(T)\|A^n\|_T \le C_5.$$

To make sense of condition (c), consider the case where all matrices $A \in \mathcal{F}$ are normal. Then each A can be diagonalized by a unitary similarity transformation and so

(c) holds with $\|S\|_2 = \|S^{-1}\|_2 = 1$ and $b_{ij} = 0$ for $i \neq j$. More generally, condition (c) requires that we can bring all $A \in \mathcal{F}$ to upper triangular form by uniformly well-conditioned similarity transformations, and with a uniform bound on the off-diagonals that is related to how close the diagonal elements (which are the eigenvalues of A) are to the unit circle.

Several other equivalent conditions have been identified and are sometimes more useful in practice; see Richtmyer and Morton [75] or the more recent paper of Strikwerda and Wade [85].

The equivalence of the conditions in Theorem D.4 can be proved by showing that (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c) $\Longrightarrow$ (d) $\Longrightarrow$ (d$'$) $\Longrightarrow$ (a). For a more complete discussion and proofs see [75] or [85].

The equivalence of (a) and (b) also follows directly from (D.46) and a proof of this can be found in [92]. It is this equivalence that is the most interesting part of the theorem and that has received the most attention in subsequent work, to the point where the term "Kreiss matrix theorem" is often applied to inequalities of the form (D.46).

D.7 Variable coefficient problems

So far we have only considered solving equations of the form $U^{n+1} = AU^n$ or $u'(t) = Au(t)$, where the matrix A is constant (independent of n or t), and the solution can be written in terms of powers or matrix exponentials. In most applications, however, the matrix changes with time. Often A represents the Jacobian matrix for a nonlinear system and so it certainly varies with time as the solution changes. Adding this complication makes it considerably more difficult to analyze the behavior of solutions. Often a study of the "frozen coefficient" problem where A is frozen at a particular value as we solve forward in time is valuable, however, to gain some information about issues such as boundedness of the solution, and the theory presented earlier in this appendix will be useful in many contexts. However, new issues can come into play when the matrices vary, particularly if they vary rapidly, or more accurately, particularly if the *eigenvectors* of the matrix vary rapidly in time. We will not discuss this in detail—just give a brief introduction to this topic.

We first consider a discrete time iteration of the form

$$U^{n+1} = A_n U^n, \tag{D.78}$$

where A_n may vary with n. The solution is

$$U^j = A_{j-1} A_{j-2} \cdots A_1 A_0 U^0. \tag{D.79}$$

If $A_n \equiv A$ for all n, then this reduces to $U^j = A^j U^0$, but more generally the matrix product is harder to analyze than powers of a single matrix.

One case is relatively simple: suppose all the matrices A_n have the same eigenvectors, although possibly different eigenvalues, so

$$A_n = R\Lambda_n R^{-1} \tag{D.80}$$

for some fixed matrix R. In this case we say the A_n are *simultaneously diagonalizable*. Then the product in (D.79) reduces to

$$U^j = R\Lambda_{j-1}\Lambda_{j-2}\cdots\Lambda_1\Lambda_0 R^{-1}U^0$$

and $\Lambda_{j-1}\Lambda_{j-2}\cdots\Lambda_1\Lambda_0$ is a diagonal matrix whose ith diagonal element is the product of the ith eigenvalue of each of the matrices $A_0, A_1, \ldots, A_{j-1}$. Then we clearly have, for example, that if $\rho(A_n) \le 1$ for all n, then $\|U^n\|$ is uniformly bounded as $n \to \infty$. In fact we don't need $\rho(A_n) \le 1$ for all n; it is sufficient to have

$$\rho(A_n) \le 1 + \gamma_n \tag{D.81}$$

for some sequence of values γ_n satisfying

$$\sum_{j=0}^{\infty} \gamma_j < \infty. \tag{D.82}$$

From (D.81) it follows that $\rho(A_n) \le e^{\gamma_n}$ and so

$$\begin{aligned} \rho(R\Lambda_{j-1}\Lambda_{j-2}\cdots\Lambda_1\Lambda_0 R^{-1}) &\le \prod_{n=0}^{j-1}\rho(A_n) \\ &\le \prod_{n=0}^{j-1} e^{\gamma_n} \\ &\le \exp\left(\sum_{n=0}^{j-1}\gamma_n\right), \end{aligned} \tag{D.83}$$

and hence $\|U^n\|$ is uniformly bounded.

If the eigenvectors vary with n, however, then it happens that $\|U^n\|$ will grow without bound even if $\rho(A_n) < 1$ for all n.

Example D.3. As a simple example, consider

$$A_0 = \begin{bmatrix} 0 & 0 \\ 2 & 0.1 \end{bmatrix}, \qquad A_1 = \begin{bmatrix} 0.1 & 2 \\ 0 & 0 \end{bmatrix} \tag{D.84}$$

and then let A_n alternate between these two matrices for larger n, so $A_{2i} = A_0$ and $A_{2i+1} = A_1$. Then $\rho(A_n) = 0.1$ for all n. However, we see that after an even number of steps

$$U^{2i} = (A_1A_0)^i U^0$$

and

$$A_1A_0 = \begin{bmatrix} 4 & 0.2 \\ 0 & 0 \end{bmatrix},$$

so $\|U^j\|$ grows like 2^j.

If we iterated with either A_0 or A_1 alone, then U^n would go rapidly to zero. But note that these matrices are nonnormal and in either case there can be transient growth before decay sets in. For example,

$$U^0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \implies A_0 U^0 = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \implies A_0^2 U^0 = \begin{bmatrix} 0 \\ 0.2 \end{bmatrix}$$
$$\implies A_0^3 U^0 = \begin{bmatrix} 0 \\ 0.02 \end{bmatrix} \implies \text{etc.}$$

Beyond the first iteration there is exponential decay by a factor 0.1 each iteration since $A_0 U^0$ is in the eigenspace of A_0 corresponding to $\lambda = 0.1$. But if we apply A_0 only once to U^0 and then apply A_1, we instead obtain

$$U^0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \implies A_0 U^0 = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \implies A_1 A_0 U^0 = \begin{bmatrix} 4 \\ 0 \end{bmatrix}.$$

Instead of decay we see amplification by another factor of 2. Moreover, the vector has been moved by A_1 out of the eigenspace it was in and into a vector that again suffers transient growth when A_0 is next applied. This couldn't happen if A_0 and A_1 shared the same eigenspaces.

One way to try to guarantee that the vectors U^n generated by the process (D.78) are uniformly bounded is to look for a single norm $\|\cdot\|$ in which

$$\|A_n\| \le 1 + \gamma_n \tag{D.85}$$

with (D.82) holding. In the simultaneously diagonalizable case considered above we can base the norm on the joint eigenvector matrix R using Theorem C.4. Another case in which we have stability is if the matrices A_n are all normal and satisfy (D.81), for then we can use the 2-norm.

But even if the matrices are not normal and the eigenvectors vary, if they do so slowly enough we may be able to prove boundedness using the following theorem. Here $\|\cdot\|_{T_n}$ is the T_n-norm defined by (C.35) in terms of some fixed norm $\|\cdot\|$, and we may be able to use the eigenvector matrix R_n of each A_n for these norms, although the theorem allows more flexibility.

Theorem D.5. *Suppose that for the difference equation* (D.78) *we can find a sequence of nonsingular matrices T_n such that*

1. $\|A_n\|_{T_n} \le 1 + \gamma_n$ *with* $\sum_{j=0}^{\infty} \gamma_j < \infty$,
2. $\|T_n\| \le C$, *a constant independent of n, and*
3. $\|T_n^{-1} T_{n-1}\| \le 1 + \beta_n$ *with* $\sum_{j=0}^{\infty} \beta_j < \infty$.

Then $\|U^n\|$ is uniformly bounded for all n.

Note that the third condition is the requirement that the norm vary sufficiently slowly. The proof can be found in [23].

Similar considerations apply to solutions to the variable coefficient ODE

$$u'(t) = A(t)u(t). \tag{D.86}$$

If the $A(t)$ are simultaneously diagonalizable for all t, then this reduces to $v'(t) = \Lambda(t)v(t)$, where $v(t) = R^{-1}u(t)$. Then

$$v_i(t) = \exp\left(\int_0^t \lambda_i(\tau)\,d\tau\right)$$

and $\|u(t)\|$ is uniformly bounded in t if $\int_0^t \alpha(A(\tau))\,d\tau$ is uniformly bounded, where $\alpha(A(t))$ is the spectral abscissa (D.33). (For the constant case $A(t) \equiv A$, this requires $\alpha(A) \le 0$.)

If the $A(t)$ are not simultaneously diagonalizable, then there are examples, similar to the Example D.3, where $\|u(t)\|$ may grow without bound even if all the matrices $A(t)$ have eigenvalues only in the left half-plane.

For a general function $A(t)$ we have (D.57),

$$\frac{d}{dt}\|u(t)\| \le \omega(A(t))\|u(t)\|,$$

where $\omega(A)$ is the numerical abscissa (D.50). Dividing by $\|u(t)\|$ gives

$$\frac{d}{dt}\log(\|u(t)\|) \le \omega(A(t))$$

and integrating yields

$$\|u(t)\| \le \exp\left(\int_0^t \omega(A(t))\,d\tau\right)\|u(0)\|. \tag{D.87}$$

This shows that the solution $u(t)$ is bounded in norm for all t provided that $\int_0^t \omega(A(t))\,d\tau$ is bounded above uniformly in t.

Recall, however, that $\omega(A)$ may be positive even when the eigenvalues of A all have negative real part (in the nonnormal case). So requiring, for example, $\omega(A(t)) \le 0$ for all t is akin to requiring $\|A_n\| \le 1$ in the same norm for all matrices A_n in the difference equation (D.78). As in the case of the difference equation this requirement can be relaxed by introducing the notion of a logarithmic T-norm that varies with time, and a theorem similar to Theorem D.5 obtained for differential equations if the eigenvector matrix of $A(t)$ is not varying too rapidly; see [23].

For a hint of what's involved, suppose the matrices are all diagonalizable, $A(t) = R(t)\Lambda(t)R^{-1}(t)$, and that $R(t)$ is differentiable. Note that $(R^{-1})'(t) = -R^{-1}(t)R'(t)R^{-1}(t)$, obtained by differentiating $RR^{-1} = I$. If we set $v(t) = R^{-1}(t)u(t)$ we find that

$$\begin{aligned} v' &= R^{-1}u' + (R^{-1})'u \\ &= R^{-1}ARv + (R^{-1})'Rv \\ &= (\Lambda - R^{-1}R')v. \end{aligned} \tag{D.88}$$

So the boundedness of $v(t)$ depends on the matrices $\Lambda(t) - R^{-1}(t)R'(t)$, and if the eigenvectors vary rapidly, then the latter term can lead to unbounded growth even if the eigenvalues are all in the left half-plane.