

10.9 Functions of Matrices, the Exponential, and Linear Differential Equations

We have known since Chapter 2 what $p(\mathbf{A})$ means for a polynomial p . This section gives a meaning to $f(\mathbf{A})$ for a much larger class of functions — square roots, logarithms, sines, and above all $e^{\mathbf{A}}$ — and the Jordan form is what makes it possible. The definition needs a well-definedness proof, which turns out to be Hermite interpolation from Chapter 6: whatever f is, $f(\mathbf{A})$ is really $p(\mathbf{A})$ for a polynomial p that agrees with f to the right order at each eigenvalue. With $e^{t\mathbf{A}}$ in hand, every linear system of differential equations with constant coefficients is solved in one formula.

Throughout this section $F = \mathbb{C}$ (or $\mathbb{R}$, when we say so), so that every characteristic polynomial splits (Corollary 6.5.5) and Jordan forms exist.

10.9.1 What a polynomial does to a Jordan block

Start with what we already have. A Jordan block is a scalar plus a nilpotent matrix whose powers stop, and a polynomial applied to it can be read off from a Taylor expansion.

Lemma 10.9.1. A Polynomial of a Jordan Block

Let F have characteristic 0, let $k \geq 1$, $\lambda \in F$, and $p \in F[x]$. Then $p(\mathbf{J}_k(\lambda))$ is the upper triangular matrix whose $(r, r+j)$ -entry is

$$(p(\mathbf{J}_k(\lambda)))_{r,r+j} = \frac{p^{(j)}(\lambda)}{j!} \quad (0 \leq j \leq k-1, 1 \leq r \leq k-j),$$

with all entries below the diagonal equal to 0. In particular the value $p(\lambda)$ fills the diagonal, $p'(\lambda)$ the superdiagonal, and only the first k derivatives of p at λ are involved.

Proof. Write $\mathbf{J} = \mathbf{J}_k(\lambda) = \lambda \mathbf{I}_k + \mathbf{M}$ with $\mathbf{M} = \mathbf{J}_k(0)$ (Definition 10.3.1). Reading off the columns of $\mathbf{M}$, we have $\mathbf{M}\mathbf{e}_1 = \mathbf{0}$ and $\mathbf{M}\mathbf{e}_r = \mathbf{e}_{r-1}$ for $r \geq 2$, so by induction $\mathbf{M}^j\mathbf{e}_r = \mathbf{e}_{r-j}$ when $r > j$ and $\mathbf{M}^j\mathbf{e}_r = \mathbf{0}$ when $r \leq j$. Hence, for $0 \leq j \leq k-1$, the matrix $\mathbf{M}^j$ has 1 in each position $(r, r+j)$ and 0 elsewhere, and $\mathbf{M}^k = \mathbf{O}$. Let $d = \deg p$. By Theorem 6.4.7 (b), applied with $c = \lambda$ and $n = d$,

$$p(x) = \sum_{j=0}^d \frac{p^{(j)}(\lambda)}{j!} (x - \lambda)^j.$$

Substituting $\mathbf{J}$ for x — legitimate because evaluation at $\mathbf{J}$ is a ring homomorphism (Theorem 6.7.1) — and using $\mathbf{J} - \lambda \mathbf{I}_k = \mathbf{M}$,

$$p(\mathbf{J}) = \sum_{j=0}^d \frac{p^{(j)}(\lambda)}{j!} \mathbf{M}^j = \sum_{j=0}^{\min(d, k-1)} \frac{p^{(j)}(\lambda)}{j!} \mathbf{M}^j,$$

since $\mathbf{M}^j = \mathbf{O}$ for $j \geq k$. Each $\mathbf{M}^j$ contributes its coefficient to every position $(r, r+j)$ and to no other position, which is the claim.

For $p = x^m$ this reads $p^{(j)}(\lambda)/j! = \binom{m}{j}\lambda^{m-j}$, so

$$(\mathbf{J}_k(\lambda)^m)_{r,r+j} = \binom{m}{j}\lambda^{m-j},$$

which is the formula behind everything that follows, here and in the next section.

10.9.2 Functions of a matrix

Reading Lemma 10.9.1 backwards suggests the definition. The formula for $p(\mathbf{J}_k(\lambda))$ uses nothing about p except the numbers $p(\lambda), p'(\lambda), \dots, p^{(k-1)}(\lambda)$. Any function that has those derivatives can be inserted instead.

To apply f to a matrix, apply it to each Jordan block, putting $f^{(j)}(\lambda)/j!$ on the j -th superdiagonal.

Definition 10.9.1. Function of a Matrix

Let $\mathbf{A} \in M_n(\mathbb{C})$ with Jordan form $\mathbf{J} = \mathbf{J}_{k_1}(\lambda_1) \oplus \dots \oplus \mathbf{J}_{k_m}(\lambda_m)$ and $\mathbf{A} = \mathbf{PJP}^{-1}$. Let f be a complex-valued function that is defined and **at least $k_i - 1$ times differentiable at λ_i** for every i — for instance, a function given by a convergent power series on a disc around each eigenvalue. Define

$$f(\mathbf{J}_k(\lambda)) := \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \dots & \frac{f^{(k-1)}(\lambda)}{(k-1)!} \\ & f(\lambda) & f'(\lambda) & \ddots & \vdots \\ & & \ddots & \ddots & \frac{f''(\lambda)}{2!} \\ & & & f(\lambda) & f'(\lambda) \\ & & & & f(\lambda) \end{pmatrix} \in M_k(\mathbb{C}),$$

the upper triangular matrix with $f^{(j)}(\lambda)/j!$ in every position $(r, r+j)$, and

$$f(\mathbf{A}) := \mathbf{P} \left(f(\mathbf{J}_{k_1}(\lambda_1)) \oplus \dots \oplus f(\mathbf{J}_{k_m}(\lambda_m)) \right) \mathbf{P}^{-1}.$$

In words: the diagonal of $f(\mathbf{J})$ records the values $f(\lambda_i)$, and each further superdiagonal records the next derivative, divided by a factorial. For a diagonalizable $\mathbf{A}$ every block has size 1, no derivatives are needed, and the definition says the familiar thing: $f(\mathbf{A}) = \mathbf{P} \operatorname{diag}(f(\lambda_1), \dots, f(\lambda_n)) \mathbf{P}^{-1}$.

Well-definedness. Two choices were made: the invertible $\mathbf{P}$ and the order of the blocks. The Jordan form itself is unique only up to the order of the blocks (Theorem 10.3.1), and even for a fixed $\mathbf{J}$ there are many $\mathbf{P}$ with $\mathbf{A} = \mathbf{PJP}^{-1}$. The next theorem removes both worries at once, by identifying $f(\mathbf{A})$ with $p(\mathbf{A})$ for a suitable **polynomial** p , an expression in which no choice appears.

Theorem 10.9.2. Functions of a Matrix Are Polynomials in It

Let $\mathbf{A} \in M_n(\mathbb{C})$ with distinct eigenvalues $\mu_1, \dots, \mu_d$, and write

$$m_{\mathbf{A}} = (x - \mu_1)^{s_1} \dots (x - \mu_d)^{s_d},$$

so that s_l is the largest size of a μ_l -block of the Jordan form of $\mathbf{A}$ (Corollary 10.3.2 (c)). Let f be as in Definition 10.9.1. Then:

- (a) there is a polynomial $p \in \mathbb{C}[x]$ with $\deg p < \deg m_{\mathbf{A}}$ and

$$p^{(j)}(\mu_l) = f^{(j)}(\mu_l) \quad \text{for all } l = 1, \dots, d \text{ and } j = 0, \dots, s_l - 1,$$

and p is the **only** such polynomial of degree less than $\deg m_{\mathbf{A}}$;

- (b) for **every** polynomial $q \in \mathbb{C}[x]$ satisfying those $\deg m_{\mathbf{A}}$ equations, $q(\mathbf{A}) = p(\mathbf{A})$;
 (c) $f(\mathbf{A}) = p(\mathbf{A})$. In particular $f(\mathbf{A})$ does not depend on the choice of $\mathbf{P}$ or on the order of the blocks, and $f(\mathbf{A})$ is determined by the finitely many numbers $f^{(j)}(\mu_l)$.

The polynomial p of (a) is the **Hermite interpolating polynomial** of f on the spectrum of $\mathbf{A}$.

Idea. The list of conditions “match f and its first $s_l - 1$ derivatives at μ_l ” is exactly a congruence modulo $(x - \mu_l)^{s_l}$, by the Taylor expansion; and the moduli $(x - \mu_l)^{s_l}$ are pairwise coprime. So the Chinese Remainder Theorem produces p and shows it is unique below the degree of the product $m_{\mathbf{A}}$ — this is Example (Values and Slopes by CRT) with more nodes and higher orders. Once p exists, Lemma 10.9.1 says that p and f do the same thing to every Jordan block of $\mathbf{A}$, because a μ_l -block has size at most s_l and only derivatives up to order $s_l - 1$ are consulted.

Proof.

- (a) For each l let $r_l := \sum_{j=0}^{s_l-1} \frac{f^{(j)}(\mu_l)}{j!} (x - \mu_l)^j \in \mathbb{C}[x]$. For $g \in \mathbb{C}[x]$, the Taylor expansion of Theorem 6.4.7 (b) at μ_l , together with the uniqueness of the remainder on division by $(x - \mu_l)^{s_l}$ (Theorem 6.1.1), gives

$$g \equiv r_l \pmod{(x - \mu_l)^{s_l}} \quad \Longleftrightarrow \quad g^{(j)}(\mu_l) = f^{(j)}(\mu_l) \text{ for } j = 0, \dots, s_l - 1,$$

exactly as in Example (Values and Slopes by CRT) (a) for $s_l = 2$: the remainder of g on division by $(x - \mu_l)^{s_l}$ is $\sum_{j < s_l} \frac{g^{(j)}(\mu_l)}{j!} (x - \mu_l)^j$, and two polynomials of degree less than s_l are equal precisely when their coefficients in the basis $(1, x - \mu_l, \dots, (x - \mu_l)^{s_l-1})$ agree (Theorem 6.4.7 (a)).

The moduli $(x - \mu_1)^{s_1}, \dots, (x - \mu_d)^{s_d}$ are pairwise coprime: for $l \neq l'$, any common divisor divides both, and a monic common divisor would be a power of $x - \mu_l$ and a power of $x - \mu_{l'}$ with $\mu_l \neq \mu_{l'}$, hence 1 (Lemma 9.5.2, Theorem 6.3.5). Their product is $m_{\mathbf{A}}$. So Theorem 6.6.4 (c) provides exactly one p with $\deg p < \deg m_{\mathbf{A}}$ and $p \equiv r_l \pmod{(x - \mu_l)^{s_l}}$ for every l , which by the displayed equivalence is (a).

- (b) Let q satisfy the same equations. By the equivalence again, $q \equiv r_l \equiv p \pmod{(x - \mu_l)^{s_l}}$ for every l , so Theorem 6.6.4 (b) gives $m_{\mathbf{A}} \mid q - p$. Writing $q - p = m_{\mathbf{A}} h$ and evaluating at $\mathbf{A}$, $q(\mathbf{A}) - p(\mathbf{A}) = m_{\mathbf{A}}(\mathbf{A})h(\mathbf{A}) = \mathbf{O}$ since $m_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}$ (Definition 9.5.1). Hence $q(\mathbf{A}) = p(\mathbf{A})$.
 (c) Let $\mathbf{J} = \mathbf{J}_{k_1}(\lambda_1) \oplus \dots \oplus \mathbf{J}_{k_m}(\lambda_m)$ be any Jordan form of $\mathbf{A}$ and $\mathbf{A} = \mathbf{PJP}^{-1}$. Fix i and let l be the index with $\lambda_i = \mu_l$. By Corollary 10.3.2 (c), $k_i \leq s_l$, so Lemma 10.9.1 and

the equations in (a) give

$$(p(\mathbf{J}_{k_i}(\lambda_i)))_{r,r+j} = \frac{p^{(j)}(\lambda_i)}{j!} = \frac{f^{(j)}(\lambda_i)}{j!} = (f(\mathbf{J}_{k_i}(\lambda_i)))_{r,r+j} \quad (0 \leq j \leq k_i - 1),$$

the derivatives involved having order at most $k_i - 1 \leq s_l - 1$. So $p(\mathbf{J}_{k_i}(\lambda_i)) = f(\mathbf{J}_{k_i}(\lambda_i))$ for every i . Since evaluating a polynomial blockwise gives the polynomial of the block diagonal matrix (Theorem 8.1.4 (a), (b)),

$$\begin{aligned} p(\mathbf{A}) &= p(\mathbf{PJP}^{-1}) = \mathbf{P} p(\mathbf{J}) \mathbf{P}^{-1} = \mathbf{P} \left(\bigoplus_i p(\mathbf{J}_{k_i}(\lambda_i)) \right) \mathbf{P}^{-1} \\ &= \mathbf{P} \left(\bigoplus_i f(\mathbf{J}_{k_i}(\lambda_i)) \right) \mathbf{P}^{-1} = f(\mathbf{A}), \end{aligned}$$

the second equality by Proposition 4.2.6 (c). The left-hand side mentions neither $\mathbf{P}$ nor the order of the blocks, so neither does the right-hand side. This proves the theorem. ■

So the definition is consistent with the old one: if f happens to be a polynomial, then f itself satisfies the interpolation conditions, and (b) gives $f(\mathbf{A})$ in the new sense = $f(\mathbf{A})$ in the old sense.

Warning. $f(\mathbf{A})$ is not f applied to the entries of $\mathbf{A}$. Even for $f(x) = x^2$ this fails: with $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, squaring each entry gives $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, while $\mathbf{A}^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. The definition is designed so that $f(\mathbf{A})$ is a polynomial in $\mathbf{A}$, and entrywise application is not.

Example. Square Roots of Two Matrices

Using Definition 10.9.1 with $f(x) = \sqrt{x}$, the branch with $\sqrt{4} = 2$ and $\sqrt{9} = 3$, compute a square root of

(a) $\mathbf{A} = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix};$

(b) $\mathbf{B} = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}.$

Solution.

- (a) $p_{\mathbf{A}} = (x-5)^2 - 16 = x^2 - 10x + 9 = (x-1)(x-9)$, so the eigenvalues are 1 and 9, distinct, and $\mathbf{A}$ is diagonalizable. Solving $(\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0}$ gives $\text{span}((1, -1))$ and $(\mathbf{A} - 9\mathbf{I})\mathbf{x} = \mathbf{0}$ gives $\text{span}((1, 1))$. With $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, $\det \mathbf{P} = 2$, and $\mathbf{P}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$, we have $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \text{diag}(1, 9)$ and

$$f(\mathbf{A}) = \mathbf{P} \text{diag}(1, 3) \mathbf{P}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 3 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Check. $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^2 = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix} = \mathbf{A}.$

(b) $\mathbf{B} = \mathbf{J}_2(4)$ is its own Jordan form, with $\mathbf{P} = \mathbf{I}$. Here $f(4) = 2$ and $f'(x) = \frac{1}{2\sqrt{x}}$, so $f'(4) = \frac{1}{4}$, and [Definition 10.9.1](#) gives

$$f(\mathbf{B}) = \begin{pmatrix} 2 & \frac{1}{4} \\ 0 & 2 \end{pmatrix}.$$

Check. $\begin{pmatrix} 2 & 1/4 \\ 0 & 2 \end{pmatrix}^2 = \begin{pmatrix} 4 & 2 \cdot \frac{1}{4} + \frac{1}{4} \cdot 2 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix} = \mathbf{B}.$

The Hermite polynomial. For (b), $m_{\mathbf{B}} = (x-4)^2$, and the interpolating polynomial of degree < 2 with $p(4) = 2$, $p'(4) = \frac{1}{4}$ is $p = 2 + \frac{1}{4}(x-4) = \frac{1}{4}x + 1$. Indeed $p(\mathbf{B}) = \frac{1}{4}\mathbf{B} + \mathbf{I} = \begin{pmatrix} 2 & 1/4 \\ 0 & 2 \end{pmatrix}$, as [Theorem 10.9.2](#) (c) predicts.

Quick check. A matrix $\mathbf{A} \in M_5(\mathbb{C})$ has Jordan form $\mathbf{J}_3(2) \oplus \mathbf{J}_1(2) \oplus \mathbf{J}_1(7)$. Which values of f and its derivatives does $f(\mathbf{A})$ depend on?

Solution. On $f(2), f'(2), f''(2)$, and $f(7)$: four numbers. Indeed $m_{\mathbf{A}} = (x-2)^3(x-7)$ by Corollary 10.3.2 (c), so $s_1 = 3$ for $\mu_1 = 2$ and $s_2 = 1$ for $\mu_2 = 7$, and [Theorem 10.9.2](#) lists the derivatives of order 0, 1, 2 at 2 and order 0 at 7. Note $\deg m_{\mathbf{A}} = 4$, matching the number of conditions and the degree bound on the interpolating polynomial. Nothing about f away from $\{2, 7\}$ matters, and no third derivative is consulted even though $\dim = 5$.

10.9.3 The matrix exponential

The most important function of a matrix is the exponential, and it has a second description, as the sum of a series. We check that the series converges, and that its sum is the matrix $\exp(\mathbf{A})$ that [Definition 10.9.1](#) produces. Convergence is meant **entrywise**, as in Chapter 9: a sequence $\mathbf{M}_1, \mathbf{M}_2, \dots$ in $M_n(\mathbb{C})$ converges to $\mathbf{M}$ if for each position (r, s) the number sequence $(\mathbf{M}_m)_{rs}$ converges to $\mathbf{M}_{rs}$, and a series of matrices converges if its sequence of partial sums does. No notion of length of a matrix is used anywhere in this section.

Remark. What we assume from analysis. Everything below uses exactly the following facts about real and complex numbers, and nothing else. Chapter 16 will develop the tools that make such statements routine; here they are quoted.

- (A1) **The scalar exponential.** For every $z \in \mathbb{C}$ the series $\sum_{i \geq 0} z^i/i!$ converges; its sum is written e^z . It satisfies $e^{z+w} = e^z e^w$, $e^0 = 1$, and $e^{i\theta} = \cos \theta + i \sin \theta$ for real θ . For real $s \geq 0$, $e^s \geq 1$.
- (A2) **Termwise differentiation.** If a power series $\sum_{m \geq 0} c_m t^m$ with $c_m \in \mathbb{C}$ converges for **every** real t , then the function $t \mapsto \sum_m c_m t^m$ is differentiable on $\mathbb{R}$ with derivative $\sum_{m \geq 1} m c_m t^{m-1}$.
- (A3) **Limits and products.** Limits of convergent sequences of complex numbers respect

sums, scalar multiples, and products. The derivative of a product of two differentiable functions $\mathbb{R} \rightarrow \mathbb{C}$ obeys the product rule, and differentiation is linear.

(A4) **Zero derivative.** A differentiable function $\mathbb{R} \rightarrow \mathbb{C}$ whose derivative vanishes everywhere is constant.

(A5) **Decay.** For every integer $j \geq 0$ and every real $c < 0$, $t^j e^{ct} \rightarrow 0$ as $t \rightarrow \infty$.

Fact (A5) is used only in the last exercise of this section. Nothing else about $\exp$, $\cos$, or $\sin$ is assumed: the derivative $\frac{d}{dt} e^{zt} = z e^{zt}$ for $z \in \mathbb{C}$ is the 1×1 case of [Theorem 10.9.4](#) (b) below, proved from (A2), and the derivatives of $\cos$ and $\sin$ used in the examples follow from it through Euler's formula in (A1).

Definition 10.9.2. Matrix Exponential

Let $\mathbf{A} \in M_n(\mathbb{C})$. The **matrix exponential** of $\mathbf{A}$ is

$$e^{\mathbf{A}} := \sum_{m=0}^{\infty} \frac{\mathbf{A}^m}{m!} = \lim_{M \rightarrow \infty} \sum_{m=0}^M \frac{\mathbf{A}^m}{m!},$$

the limit taken entrywise, provided it exists; $\mathbf{A}^0 = \mathbf{I}$.

That the limit does exist, and what it is, is the content of the next theorem. The Jordan form reduces the question to a single block, where the series can be summed exactly.

Theorem 10.9.3. The Exponential Series Converges to the Jordan Definition

Let $\mathbf{A} \in M_n(\mathbb{C})$. Then the series $\sum_m \mathbf{A}^m/m!$ converges entrywise, and its sum is the matrix $\exp(\mathbf{A})$ given by [Definition 10.9.1](#) for the function $f = \exp$. In particular, for a single Jordan block,

$$e^{\mathbf{J}_k(\lambda)} = e^{\lambda} \begin{pmatrix} 1 & 1 & \frac{1}{2!} & \cdots & \frac{1}{(k-1)!} \\ & 1 & 1 & \ddots & \vdots \\ & & \ddots & \ddots & \frac{1}{2!} \\ & & & 1 & 1 \\ & & & & 1 \end{pmatrix},$$

the upper triangular matrix with $e^{\lambda}/j!$ in every position $(r, r+j)$.

Idea. For one block the entries of the partial sums are themselves partial sums of a scalar series, and that series can be recognized: the terms $\binom{m}{j} \lambda^{m-j}/m!$ simplify to $\lambda^{m-j}/(j!(m-j)!)$, which is $1/j!$ times the $(m-j)$ -th term of the series for e^{λ} . For general $\mathbf{A}$, conjugation by a fixed $\mathbf{P}$ turns each entry of a partial sum into a fixed finite linear combination of entries of the block partial sums, and limits pass through finite linear combinations.

Proof. Step 1. One block. Fix $k \geq 1$, $\lambda \in \mathbb{C}$, and a position $(r, r+j)$ with $0 \leq j \leq k-1$.

By Lemma 10.9.1 applied to $p = x^m$,

$$\left(\frac{\mathbf{J}_k(\lambda)^m}{m!}\right)_{r, r+j} = \frac{1}{m!} \binom{m}{j} \lambda^{m-j} = \begin{cases} \frac{\lambda^{m-j}}{j! (m-j)!}, & m \geq j, \\ 0, & m < j, \end{cases}$$

using $\binom{m}{j}/m! = 1/(j!(m-j)!)$ for $m \geq j$ and $\binom{m}{j} = 0$ for $m < j$. Hence the partial sums satisfy

$$\sum_{m=0}^{\mathbf{M}} \left(\frac{\mathbf{J}_k(\lambda)^m}{m!}\right)_{r, r+j} = \frac{1}{j!} \sum_{i=0}^{\mathbf{M}-j} \frac{\lambda^i}{i!} \quad (\mathbf{M} \geq j),$$

after the substitution $i = m - j$. By (A1) the right-hand side converges to $e^\lambda/j!$ as $\mathbf{M} \rightarrow \infty$. Entries with $j < 0$, that is entries below the diagonal, are 0 in every term. So the series converges entrywise and

$$e^{\mathbf{J}_k(\lambda)} = \left(\text{upper triangular with } \frac{e^\lambda}{j!} \text{ on the } j\text{-th superdiagonal}\right) = \exp(\mathbf{J}_k(\lambda)),$$

the last equality being Definition 10.9.1 for $f = \exp$, whose derivatives are $f^{(j)} = \exp$, so $f^{(j)}(\lambda)/j! = e^\lambda/j!$.

Step 2. A Jordan matrix. If $\mathbf{J} = \mathbf{J}_{k_1}(\lambda_1) \oplus \cdots \oplus \mathbf{J}_{k_m}(\lambda_m)$, then $\mathbf{J}^m/m!$ is the direct sum of the corresponding blocks (Theorem 8.1.4 (b)), so each entry of a partial sum of $\sum_m \mathbf{J}^m/m!$ is either identically 0 or an entry of a partial sum from one block. By Step 1 all of them converge, and $e^{\mathbf{J}} = \bigoplus_i e^{\mathbf{J}_{k_i}(\lambda_i)}$.

Step 3. General A. By Theorem 10.3.1 there are a Jordan matrix $\mathbf{J}$ and an invertible $\mathbf{P}$ with $\mathbf{A} = \mathbf{PJP}^{-1}$. Then $\mathbf{A}^m = \mathbf{PJ}^m\mathbf{P}^{-1}$ (Proposition 4.2.6 (c)), so

$$\sum_{m=0}^{\mathbf{M}} \frac{\mathbf{A}^m}{m!} = \mathbf{P} \left(\sum_{m=0}^{\mathbf{M}} \frac{\mathbf{J}^m}{m!} \right) \mathbf{P}^{-1}.$$

Each entry of the left-hand side is a fixed finite linear combination — with coefficients built from the entries of $\mathbf{P}$ and $\mathbf{P}^{-1}$, independent of $\mathbf{M}$ — of the entries of the middle partial sum. By Step 2 those entries converge, so by (A3) the left-hand side converges entrywise, with

$$e^{\mathbf{A}} = \mathbf{P} e^{\mathbf{J}} \mathbf{P}^{-1} = \mathbf{P} \left(\bigoplus_i e^{\mathbf{J}_{k_i}(\lambda_i)} \right) \mathbf{P}^{-1} = \exp(\mathbf{A}),$$

the last equality by Definition 10.9.1. This proves the theorem. ■

From now on we write $e^{\mathbf{A}}$ for this matrix and use both descriptions freely. Note one consequence of Theorem 10.9.2: $e^{\mathbf{A}}$ is a **polynomial** in $\mathbf{A}$, namely $p(\mathbf{A})$ for the Hermite interpolant of $\exp$ on $\text{spec}(\mathbf{A})$. Everything that commutes with $\mathbf{A}$ therefore commutes with $e^{\mathbf{A}}$.

Theorem 10.9.4. Properties of the Matrix Exponential

Let $\mathbf{A}, \mathbf{B} \in M_n(\mathbb{C})$ and let $\mathbf{P} \in M_n(\mathbb{C})$ be invertible.

- (a) $e^{\mathbf{O}} = \mathbf{I}$ and $e^{\mathbf{P}^{-1}\mathbf{A}\mathbf{P}} = \mathbf{P}^{-1}e^{\mathbf{A}}\mathbf{P}$.

(b) The function $\mathbb{R} \rightarrow M_n(\mathbb{C})$, $t \mapsto e^{t\mathbf{A}}$, is differentiable entrywise, with

$$\frac{d}{dt} e^{t\mathbf{A}} = \mathbf{A}e^{t\mathbf{A}} = e^{t\mathbf{A}}\mathbf{A}.$$

(c) $e^{\mathbf{A}}$ is invertible, with $(e^{\mathbf{A}})^{-1} = e^{-\mathbf{A}}$.

(d) If $\mathbf{AB} = \mathbf{BA}$, then $e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}}e^{\mathbf{B}}$.

(e) $\det e^{\mathbf{A}} = e^{\text{tr } \mathbf{A}}$.

Idea. Part (b) is termwise differentiation, entry by entry. Everything after it is the same two-line trick: to prove $\mathbf{X}(t) = \mathbf{Y}(t)$ for two matrix functions, show that $\mathbf{Z}(t) = e^{-t\mathbf{C}}\mathbf{X}(t)$ has derivative $\mathbf{O}$ for a suitable $\mathbf{C}$, so $\mathbf{Z}$ is constant and equals $\mathbf{Z}(0)$. That is why (c) and (d) come after (b), even though the series would also give them. Part (e) is read off the Jordan form: $e^{\mathbf{J}}$ is triangular with e^{λ_i} on the diagonal.

Proof.

- (a) $\mathbf{O}^m = \mathbf{O}$ for $m \geq 1$ and $\mathbf{O}^0 = \mathbf{I}$, so the series is $\mathbf{I}$. For the second claim, Step 3 of the previous proof applies verbatim with $\mathbf{J}$ replaced by $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$: the partial sums of $\sum_m (\mathbf{P}^{-1}\mathbf{A}\mathbf{P})^m/m!$ are $\mathbf{P}^{-1}(\sum_m \mathbf{A}^m/m!)\mathbf{P}$, and (A3) lets the limit pass through.
- (b) Fix a position (r, s) . By Theorem 10.9.3 the series $\sum_m t^m (\mathbf{A}^m)_{rs}/m!$ converges for every real t , since $(t\mathbf{A})^m = t^m \mathbf{A}^m$. So (A2) applies with $c_m = (\mathbf{A}^m)_{rs}/m!$ and gives

$$\begin{aligned} \frac{d}{dt} (e^{t\mathbf{A}})_{rs} &= \sum_{m \geq 1} \frac{m t^{m-1}}{m!} (\mathbf{A}^m)_{rs} = \sum_{m \geq 1} \frac{t^{m-1}}{(m-1)!} (\mathbf{A} \cdot \mathbf{A}^{m-1})_{rs} \\ &= \left(\mathbf{A} \sum_{l \geq 0} \frac{t^l \mathbf{A}^l}{l!} \right)_{rs}, \end{aligned}$$

where the last step moves the constant matrix $\mathbf{A}$ outside the limit: each entry of $\mathbf{A}(\sum_{l \leq L} t^l \mathbf{A}^l/l!)$ is a fixed linear combination of entries of the partial sum, so (A3) applies. Hence $\frac{d}{dt} e^{t\mathbf{A}} = \mathbf{A}e^{t\mathbf{A}}$. The same computation with $\mathbf{A}^m = \mathbf{A}^{m-1}\mathbf{A}$ gives $e^{t\mathbf{A}}\mathbf{A}$.

- (c) Put $\mathbf{H}(t) = e^{t\mathbf{A}}e^{-t\mathbf{A}}$ for $t \in \mathbb{R}$. Each entry of $\mathbf{H}$ is a finite sum of products of differentiable functions, so by (A3) and (b),

$$\mathbf{H}'(t) = (\mathbf{A}e^{t\mathbf{A}})e^{-t\mathbf{A}} + e^{t\mathbf{A}}(-\mathbf{A}e^{-t\mathbf{A}}) = \mathbf{A}e^{t\mathbf{A}}e^{-t\mathbf{A}} - \mathbf{A}e^{t\mathbf{A}}e^{-t\mathbf{A}} = \mathbf{O},$$

where the second equality uses $e^{t\mathbf{A}}\mathbf{A} = \mathbf{A}e^{t\mathbf{A}}$ from (b). By (A4) every entry of $\mathbf{H}$ is constant, so $\mathbf{H}(t) = \mathbf{H}(0) = e^{\mathbf{O}}e^{\mathbf{O}} = \mathbf{I}$ for all t . At $t = 1$, $e^{\mathbf{A}}e^{-\mathbf{A}} = \mathbf{I}$, and Theorem 3.4.5 finishes it.

- (d) Suppose $\mathbf{AB} = \mathbf{BA}$. Then $\mathbf{B}$ commutes with every partial sum of $\sum_m t^m \mathbf{A}^m/m!$, so by (A3) it commutes with the limit: $\mathbf{B}e^{t\mathbf{A}} = e^{t\mathbf{A}}\mathbf{B}$ for every t . Put

$$\mathbf{G}(t) = e^{-t(\mathbf{A}+\mathbf{B})}e^{t\mathbf{A}}e^{t\mathbf{B}}.$$

Differentiating entrywise with the product rule (A3) and (b),

$$\mathbf{G}'(t) = -(\mathbf{A} + \mathbf{B})e^{-t(\mathbf{A}+\mathbf{B})}e^{t\mathbf{A}}e^{t\mathbf{B}} + e^{-t(\mathbf{A}+\mathbf{B})}\mathbf{A}e^{t\mathbf{A}}e^{t\mathbf{B}} + e^{-t(\mathbf{A}+\mathbf{B})}e^{t\mathbf{A}}\mathbf{B}e^{t\mathbf{B}}.$$

In the third term replace $e^{t\mathbf{A}}\mathbf{B}$ by $\mathbf{B}e^{t\mathbf{A}}$; in the first, use $(\mathbf{A} + \mathbf{B})e^{-t(\mathbf{A}+\mathbf{B})} = e^{-t(\mathbf{A}+\mathbf{B})}(\mathbf{A} + \mathbf{B})$ from (b). The three terms become

$$e^{-t(\mathbf{A}+\mathbf{B})}(-(\mathbf{A} + \mathbf{B}) + \mathbf{A} + \mathbf{B})e^{t\mathbf{A}}e^{t\mathbf{B}} = \mathbf{O}.$$

By (A4), $\mathbf{G}$ is constant, so $\mathbf{G}(t) = \mathbf{G}(0) = \mathbf{I}$. Multiplying on the left by $e^{t(\mathbf{A}+\mathbf{B})}$ and using (c), $e^{t\mathbf{A}}e^{t\mathbf{B}} = e^{t(\mathbf{A}+\mathbf{B})}$; put $t = 1$.

(e) Write $\mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1}$ with $\mathbf{J} = \bigoplus_i \mathbf{J}_{k_i}(\lambda_i)$ a Jordan matrix. By (a) and [Theorem 10.9.3](#),

$$\det e^{\mathbf{A}} = \det(\mathbf{P}e^{\mathbf{J}}\mathbf{P}^{-1}) = \det e^{\mathbf{J}}$$

(Corollary 7.5.7), and $e^{\mathbf{J}} = \bigoplus_i e^{\mathbf{J}_{k_i}(\lambda_i)}$ is upper triangular with e^{λ_i} repeated k_i times along the diagonal. By Theorem 7.4.6 and (A1),

$$\det e^{\mathbf{J}} = \prod_i (e^{\lambda_i})^{k_i} = e^{\sum_i k_i \lambda_i} = e^{\text{tr } \mathbf{J}} = e^{\text{tr } \mathbf{A}},$$

the last equality because similar matrices have equal traces (Theorem 4.2.5). This proves the theorem. ■

Part (e) says something striking: $\det e^{\mathbf{A}} \neq 0$ always, which re-proves (c), and if $\text{tr } \mathbf{A} = 0$ then $e^{\mathbf{A}} \in \text{SL}_n(\mathbb{C})$.

The following corollary is the formula actually used in computations. Note that $t\mathbf{J}_k(\lambda)$ is **not** a Jordan block for $k \geq 2$ and $t \neq 1$, since its superdiagonal entries are t ; the commuting law does the work instead.

Corollary 10.9.5. Exponential of a Jordan Block, with a Parameter

Let $k \geq 1$, $\lambda \in \mathbb{C}$, and $t \in \mathbb{R}$. Then $e^{t\mathbf{J}_k(\lambda)}$ is the upper triangular matrix with

$$(e^{t\mathbf{J}_k(\lambda)})_{r, r+j} = \frac{t^j e^{\lambda t}}{j!} \quad (0 \leq j \leq k-1)$$

and zeros below the diagonal. Consequently, if $\mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1}$ with $\mathbf{J} = \mathbf{J}_{k_1}(\lambda_1) \oplus \cdots \oplus \mathbf{J}_{k_m}(\lambda_m)$, then

$$e^{t\mathbf{A}} = \mathbf{P} \left(e^{t\mathbf{J}_{k_1}(\lambda_1)} \oplus \cdots \oplus e^{t\mathbf{J}_{k_m}(\lambda_m)} \right) \mathbf{P}^{-1}.$$

In particular, for a diagonal $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$, $e^{t\mathbf{D}} = \text{diag}(e^{td_1}, \dots, e^{td_n})$.

Proof. Write $t\mathbf{J}_k(\lambda) = (t\lambda)\mathbf{I}_k + t\mathbf{M}$ with $\mathbf{M} = \mathbf{J}_k(0)$. The two summands commute, since $\mathbf{I}_k$ commutes with everything, so [Theorem 10.9.4](#) (d) gives $e^{t\mathbf{J}_k(\lambda)} = e^{(t\lambda)\mathbf{I}_k} e^{t\mathbf{M}}$. For the first factor, $((t\lambda)\mathbf{I}_k)^m = (t\lambda)^m \mathbf{I}_k$, so the partial sums are $(\sum_{m \leq M} (t\lambda)^m / m!) \mathbf{I}_k$ and $e^{(t\lambda)\mathbf{I}_k} = e^{\lambda t} \mathbf{I}_k$ by (A1). For the second, $\mathbf{M}^j = \mathbf{O}$ for $j \geq k$, so the series terminates:

$$e^{t\mathbf{M}} = \sum_{j=0}^{k-1} \frac{t^j}{j!} \mathbf{M}^j,$$

and $\mathbf{M}^j$ has 1 in each position $(r, r+j)$, as computed in the proof of [Lemma 10.9.1](#). Multiply-

ing by the scalar $e^{\lambda t}$ gives the entry formula. The second display follows from $t\mathbf{A} = \mathbf{P}(t\mathbf{J})\mathbf{P}^{-1}$ with [Theorem 10.9.4](#) (a) and Theorem 8.1.4 (b), exactly as in Step 2 of the proof of [Theorem 10.9.3](#). The diagonal case is the case in which every block has size 1. ■

Warning. $e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}}e^{\mathbf{B}}$ fails when $\mathbf{A}$ and $\mathbf{B}$ do not commute. Take

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Both square to $\mathbf{O}$, so the series stop after two terms: $e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $e^{\mathbf{B}} = \mathbf{I} + \mathbf{B} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Then

$$e^{\mathbf{A}}e^{\mathbf{B}} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = e^{\mathbf{B}}e^{\mathbf{A}}.$$

Since $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$, at most one of the two products can equal $e^{\mathbf{A}+\mathbf{B}}$, so the identity must fail for one of the pairs $(\mathbf{A}, \mathbf{B})$, $(\mathbf{B}, \mathbf{A})$. (In fact it fails for both: $(\mathbf{A} + \mathbf{B})^2 = \mathbf{I}$, so $e^{\mathbf{A}+\mathbf{B}}$ has all four entries nonzero and none of them equal to 2.) The moral: check $\mathbf{AB} = \mathbf{BA}$ before splitting an exponential.

10.9.4 Linear systems of differential equations

A system of n linear differential equations with constant coefficients,

$$x'_1 = a_{11}x_1 + \cdots + a_{1n}x_n, \quad \dots, \quad x'_n = a_{n1}x_1 + \cdots + a_{nn}x_n,$$

is the single equation $\mathbf{x}' = \mathbf{A}\mathbf{x}$ for a vector-valued function $\mathbf{x}: \mathbb{R} \rightarrow \mathbb{C}^n$, differentiated entrywise. In one dimension the answer is $x(t) = e^{ta}x_0$; the matrix exponential makes the general answer look exactly the same.

Theorem 10.9.6. Solution of a Linear System

Let $\mathbf{A} \in M_n(\mathbb{C})$ and $\mathbf{x}_0 \in \mathbb{C}^n$. Then the function

$$\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}_0 \quad (t \in \mathbb{R})$$

is differentiable and satisfies $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$ for all t and $\mathbf{x}(0) = \mathbf{x}_0$; and it is the **only** differentiable function $\mathbb{R} \rightarrow \mathbb{C}^n$ with these two properties.

Idea. Existence is [Theorem 10.9.4](#) (b) applied to a fixed vector. For uniqueness, take any solution $\mathbf{y}$ and multiply it by $e^{-t\mathbf{A}}$, the one factor that can undo the growth: the product has derivative $\mathbf{O}$, hence is constant. This is the same move as in the proofs of parts (c) and (d) above.

Proof. Existence. Each entry of $\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}_0$ is a linear combination of entries of $e^{t\mathbf{A}}$ with constant coefficients, so $\mathbf{x}$ is differentiable with $\mathbf{x}'(t) = (\frac{d}{dt}e^{t\mathbf{A}})\mathbf{x}_0$ by the linearity of differentiation (A3). By Theorem 10.9.4 (b), $\mathbf{x}'(t) = \mathbf{A}e^{t\mathbf{A}}\mathbf{x}_0 = \mathbf{A}\mathbf{x}(t)$. And $\mathbf{x}(0) = e^{\mathbf{0}}\mathbf{x}_0 = \mathbf{x}_0$ by Theorem 10.9.4 (a).

Uniqueness. Let $\mathbf{y}: \mathbb{R} \rightarrow \mathbb{C}^n$ be differentiable with $\mathbf{y}' = \mathbf{A}\mathbf{y}$ and $\mathbf{y}(0) = \mathbf{x}_0$. Put $\mathbf{z}(t) = e^{-t\mathbf{A}}\mathbf{y}(t)$. By the product rule (A3) and Theorem 10.9.4 (b) applied to $-\mathbf{A}$,

$$\mathbf{z}'(t) = (-\mathbf{A}e^{-t\mathbf{A}})\mathbf{y}(t) + e^{-t\mathbf{A}}\mathbf{y}'(t) = -\mathbf{A}e^{-t\mathbf{A}}\mathbf{y}(t) + e^{-t\mathbf{A}}\mathbf{A}\mathbf{y}(t) = \mathbf{0},$$

where the last step uses $e^{-t\mathbf{A}}\mathbf{A} = \mathbf{A}e^{-t\mathbf{A}}$ from Theorem 10.9.4 (b). By (A4) each entry of $\mathbf{z}$ is constant, so $\mathbf{z}(t) = \mathbf{z}(0) = \mathbf{x}_0$ for all t . Multiplying by $e^{t\mathbf{A}}$ and using Theorem 10.9.4 (c),

$$\mathbf{y}(t) = e^{t\mathbf{A}}\mathbf{z}(t) = e^{t\mathbf{A}}\mathbf{x}_0 = \mathbf{x}(t).$$

This proves the theorem. ■

The formula is only as useful as our ability to compute $e^{t\mathbf{A}}$, and the Jordan form does that. Two shapes of answer occur, and they are worth seeing side by side.

Example. Two Two-Dimensional Systems

Solve $\mathbf{x}' = \mathbf{A}\mathbf{x}$ with $\mathbf{x}(0) = (1, 0)$ for

- (a) $\mathbf{A} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$ (complex eigenvalues);
- (b) $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ (one eigenvalue, one block of size 2).

Solution.

- (a) $p_{\mathbf{A}} = (x-1)^2 + 4$, with roots $\lambda = 1 \pm 2i$, distinct, so $\mathbf{A}$ is diagonalizable over $\mathbb{C}$ and every block has size 1. Solving $(\mathbf{A} - (1+2i)\mathbf{I})\mathbf{x} = \mathbf{0}$: the first row reads $-2ix_1 - 2x_2 = 0$, so $x_2 = -ix_1$ and $(1, -i)$ is an eigenvector; $(1, i)$ is one for $1-2i$. With $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$ we get $\det \mathbf{P} = 2i$ and $\mathbf{P}^{-1} = \frac{1}{2i} \begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$, so $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \text{diag}(1+2i, 1-2i)$. By Corollary 10.9.5,

$$\begin{aligned} e^{t\mathbf{A}} &= \mathbf{P} \text{diag}(e^{(1+2i)t}, e^{(1-2i)t}) \mathbf{P}^{-1} \\ &= \frac{1}{2} \begin{pmatrix} e^{(1+2i)t} + e^{(1-2i)t} & i(e^{(1+2i)t} - e^{(1-2i)t}) \\ -i(e^{(1+2i)t} - e^{(1-2i)t}) & e^{(1+2i)t} + e^{(1-2i)t} \end{pmatrix}. \end{aligned}$$

By (A1), $e^{(1\pm 2i)t} = e^t(\cos 2t \pm i \sin 2t)$, so $e^{(1+2i)t} + e^{(1-2i)t} = 2e^t \cos 2t$ and $e^{(1+2i)t} -$

$e^{(1-2i)t} = 2ie^t \sin 2t$. Substituting, and using $i \cdot 2i = -2$,

$$e^{t\mathbf{A}} = e^t \begin{pmatrix} \cos 2t & -\sin 2t \\ \sin 2t & \cos 2t \end{pmatrix}, \quad \mathbf{x}(t) = e^{t\mathbf{A}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e^t \begin{pmatrix} \cos 2t \\ \sin 2t \end{pmatrix}.$$

The solution spirals outwards: rotation at angular speed 2, growth at rate 1. The entries are real, as they must be for a real $\mathbf{A}$ and a real $\mathbf{x}_0$.

Check. $\mathbf{x}'(t) = e^t(\cos 2t - 2\sin 2t, \sin 2t + 2\cos 2t)$, and $\mathbf{A}\mathbf{x}(t) = e^t(\cos 2t - 2\sin 2t, 2\cos 2t + \sin 2t)$. They agree, and $\mathbf{x}(0) = (1, 0)$.

- (b) $\mathbf{A} = \mathbf{J}_2(2)$ is its own Jordan form. Writing $\mathbf{A} = 2\mathbf{I} + \mathbf{M}$ with $\mathbf{M} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, the two summands commute and $\mathbf{M}^2 = \mathbf{O}$, so [Corollary 10.9.5](#) with $k = 2$ gives

$$e^{t\mathbf{A}} = e^{2t}(\mathbf{I} + t\mathbf{M}) = e^{2t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

So $\mathbf{x}(t) = e^{2t}(1, 0)$. A different initial vector shows the new phenomenon: $\mathbf{x}(0) = (0, 1)$ gives $\mathbf{x}(t) = e^{2t}(t, 1)$, whose first entry is te^{2t} . **A repeated eigenvalue with a block of size k produces solutions containing $t^j e^{\lambda t}$ for $j < k$.** This is the same $t^j \lambda^m$ pattern that Chapter 9 met for recurrences with repeated roots, now in continuous time.

A scalar equation of higher order is a system in disguise, and the translation is the companion matrix of Section 5.

Example. A Second-Order Equation with a Repeated Root

Solve $y'' - 4y' + 4y = 0$ with $y(0) = 1$ and $y'(0) = 3$.

Solution. Make it a system. Put $\mathbf{u} = (y, y')$. Then $\mathbf{u}' = (y', y'') = (y', 4y' - 4y)$, so $\mathbf{u}' = \mathbf{M}\mathbf{u}$ with

$$\mathbf{M} = \begin{pmatrix} 0 & 1 \\ -4 & 4 \end{pmatrix}, \quad \mathbf{u}(0) = (1, 3).$$

Note that $p_{\mathbf{M}} = x(x - 4) + 4 = x^2 - 4x + 4$ is the polynomial of the equation, and that $\mathbf{M} = \mathbf{C}(p_{\mathbf{M}})^{\top}$ for the companion matrix $\mathbf{C}(p_{\mathbf{M}}) = \begin{pmatrix} 0 & -4 \\ 1 & 4 \end{pmatrix}$ of Definition 10.5.3; by Corollary 10.7.7 the two are similar, so which of them one uses is immaterial.

Exponentiate. $p_{\mathbf{M}} = (x - 2)^2$. Put $\mathbf{N} = \mathbf{M} - 2\mathbf{I} = \begin{pmatrix} -2 & 1 \\ -4 & 2 \end{pmatrix}$; then $\mathbf{N}^2 = \begin{pmatrix} 4 - 4 & -2 + 2 \\ 8 - 8 & -4 + 4 \end{pmatrix} = \mathbf{O}$, so $2\mathbf{I}$ and $\mathbf{N}$ commute and [Theorem 10.9.4](#) (d) gives, as in the previous example,

$$e^{t\mathbf{M}} = e^{2t}(\mathbf{I} + t\mathbf{N}) = e^{2t} \begin{pmatrix} 1 - 2t & t \\ -4t & 1 + 2t \end{pmatrix}.$$

(No basis change was needed: $\mathbf{M}$ is a scalar plus a nilpotent matrix on the nose.)

Read off the answer.

$$\mathbf{u}(t) = e^{t\mathbf{M}} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = e^{2t} \begin{pmatrix} (1-2t) + 3t \\ -4t + 3(1+2t) \end{pmatrix} = e^{2t} \begin{pmatrix} 1+t \\ 3+2t \end{pmatrix},$$

so $y(t) = (1+t)e^{2t}$.

Check. $y' = e^{2t} + 2(1+t)e^{2t} = (3+2t)e^{2t}$, which is the second entry of $\mathbf{u}$, as it should be, and $y'(0) = 3$. Then $y'' = 2e^{2t} + 2(3+2t)e^{2t} = (8+4t)e^{2t}$, and

$$y'' - 4y' + 4y = [(8+4t) - 4(3+2t) + 4(1+t)]e^{2t} = 0.$$

The repeated root 2 produces the term te^{2t} , exactly as the size-2 block predicts.

Chapter 16 will return to $e^{\mathbf{A}}$ with norms available, which give quantitative bounds instead of exact formulas. The next section stays with exact formulas and asks the discrete-time question: what happens to $\mathbf{A}^m$?

10.9.5 Exercises

10.9.5.1 A. Check your understanding

Exercise. A1

- State the formula for $p(\mathbf{J}_k(\lambda))$ and say where the factorials come from.
- Define $f(\mathbf{A})$ for $\mathbf{A} \in M_n(\mathbb{C})$ with a given Jordan form, and list the two choices whose irrelevance has to be checked.
- Which values of f and its derivatives does $f(\mathbf{A})$ depend on? How many numbers is that?
- State the five properties of $e^{\mathbf{A}}$ proved in [Theorem 10.9.4](#), with their hypotheses.
- True or false: $e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}}e^{\mathbf{B}}$ for all $\mathbf{A}, \mathbf{B} \in M_n(\mathbb{C})$. Justify your answer.
- True or false: $e^{\mathbf{A}}$ can be singular. Justify your answer.

Solution.

- $p(\mathbf{J}_k(\lambda))$ is upper triangular with $p^{(j)}(\lambda)/j!$ in every position $(r, r+j)$ ([Lemma 10.9.1](#)). The factorials come from the Taylor expansion $p = \sum_j \frac{p^{(j)}(\lambda)}{j!}(x-\lambda)^j$ (Theorem 6.4.7), whose j -th term contributes $\mathbf{J}_k(0)^j$, the matrix with ones on the j -th superdiagonal.
- $f(\mathbf{A}) = \mathbf{P}(\oplus_i f(\mathbf{J}_{k_i}(\lambda_i)))\mathbf{P}^{-1}$ with $f(\mathbf{J}_k(\lambda))$ as in (a) ([Definition 10.9.1](#)). The choices are the matrix $\mathbf{P}$ and the order in which the blocks are listed; both are irrelevant because $f(\mathbf{A}) = p(\mathbf{A})$ for the Hermite interpolating polynomial p ([Theorem 10.9.2](#)).
- On $f^{(j)}(\mu)$ for each eigenvalue μ and each $j < s_\mu$, where s_μ is the exponent of $x - \mu$ in $m_{\mathbf{A}}$, that is, the largest μ -block size. The total number is $\sum_\mu s_\mu = \deg m_{\mathbf{A}}$ ([Theorem 10.9.2](#)).
- $e^{\mathbf{P}^{-1}\mathbf{A}\mathbf{P}} = \mathbf{P}^{-1}e^{\mathbf{A}}\mathbf{P}$ for invertible $\mathbf{P}$; $\frac{d}{dt}e^{t\mathbf{A}} = \mathbf{A}e^{t\mathbf{A}} = e^{t\mathbf{A}}\mathbf{A}$; $(e^{\mathbf{A}})^{-1} = e^{-\mathbf{A}}$; $e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}}e^{\mathbf{B}}$ provided $\mathbf{AB} = \mathbf{BA}$; and $\det e^{\mathbf{A}} = e^{\text{tr } \mathbf{A}}$ with no hypothesis.
- False. With $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\mathbf{B} = \mathbf{A}^\top$, the matrices $e^{\mathbf{A}}e^{\mathbf{B}}$ and $e^{\mathbf{B}}e^{\mathbf{A}}$ differ, so the

identity cannot hold for both orders; see the warning after [Theorem 10.9.4](#).

- (f) False. $\det e^{\mathbf{A}} = e^{\text{tr } \mathbf{A}} \neq 0$ by [Theorem 10.9.4](#) (e) and (A1), so $e^{\mathbf{A}}$ is always invertible (Theorem 7.5.2).

10.9.5.2 B. Practice

Exercise. B1: Computing exponentials

Compute $e^{t\mathbf{A}}$ for each of the following.

- (a) $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$.
 (b) $\mathbf{A} = \begin{pmatrix} 5 & 1 \\ 0 & 5 \end{pmatrix}$.
 (c) $\mathbf{A} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix} = \mathbf{J}_3(3)$.

Solution.

- (a) $p_{\mathbf{A}} = (x-1)^2 - 9$, with roots 4 and -2 . Eigenvectors: $(\mathbf{A} - 4\mathbf{I})\mathbf{x} = \mathbf{0}$ gives $\text{span}((1, 1))$; $(\mathbf{A} + 2\mathbf{I})\mathbf{x} = \mathbf{0}$ gives $\text{span}((1, -1))$. With $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ we have $\det \mathbf{P} = -2$, $\mathbf{P}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, and $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \text{diag}(4, -2)$, so by [Corollary 10.9.5](#)

$$e^{t\mathbf{A}} = \mathbf{P} \text{diag}(e^{4t}, e^{-2t}) \mathbf{P}^{-1} = \frac{1}{2} \begin{pmatrix} e^{4t} + e^{-2t} & e^{4t} - e^{-2t} \\ e^{4t} - e^{-2t} & e^{4t} + e^{-2t} \end{pmatrix}.$$

Check at $t = 0$: both diagonal entries give 1 and both off-diagonal ones 0, so $e^{\mathbf{0}} = \mathbf{I}$.

- (b) $\mathbf{A} = \mathbf{J}_2(5)$, so [Corollary 10.9.5](#) with $k = 2$ gives

$$e^{t\mathbf{A}} = e^{5t}(\mathbf{I} + t\mathbf{M}) = e^{5t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

- (c) By [Corollary 10.9.5](#) with $k = 3$ and $\lambda = 3$, or directly from $\mathbf{A} = 3\mathbf{I} + \mathbf{M}$ with $\mathbf{M} = \mathbf{J}_3(0)$ and $\mathbf{M}^3 = \mathbf{0}$, giving $e^{t\mathbf{A}} = e^{3t}(\mathbf{I} + t\mathbf{M} + \frac{t^2}{2}\mathbf{M}^2)$:

$$e^{t\mathbf{A}} = e^{3t} \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

This is [Definition 10.9.1](#) for $f(x) = e^{tx}$: the j -th superdiagonal carries $f^{(j)}(3)/j! = t^j e^{3t}/j!$.

Exercise. B2: Solving systems

Solve each initial value problem, and verify your answer by differentiating.

- (a) $x' = x + 3y$, $y' = 3x + y$, with $x(0) = 2$, $y(0) = 0$.
 (b) $\mathbf{x}' = \begin{pmatrix} 0 & -3 \\ 3 & 0 \end{pmatrix} \mathbf{x}$ with $\mathbf{x}(0) = (0, 1)$.
 (c) $y''' = 0$ with $y(0) = 1$, $y'(0) = 0$, $y''(0) = 2$, by turning it into a system.

Solution.

- (a) This is $\mathbf{x}' = \mathbf{A}\mathbf{x}$ for the matrix of [Exercise \(B1: Computing exponentials\)](#) (a). By [Theorem 10.9.6](#),

$$\mathbf{x}(t) = e^{t\mathbf{A}} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} e^{4t} + e^{-2t} \\ e^{4t} - e^{-2t} \end{pmatrix},$$

that is, $x = e^{4t} + e^{-2t}$ and $y = e^{4t} - e^{-2t}$. *Check.* $x' = 4e^{4t} - 2e^{-2t}$ and $x + 3y = e^{4t} + e^{-2t} + 3e^{4t} - 3e^{-2t} = 4e^{4t} - 2e^{-2t}$; similarly for y' . And $x(0) = 2$, $y(0) = 0$.

- (b) Write $\mathbf{A} = 3\mathbf{R}$ with $\mathbf{R} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, whose eigenvalues are $\pm i$ (Example (A Real Matrix with Non-Real Eigenvalues)), so those of $\mathbf{A}$ are $\pm 3i$, distinct. Repeating the computation of [Example \(Two Two-Dimensional Systems\)](#) (a) with $1 \pm 2i$ replaced by $\pm 3i$ — eigenvectors $(1, -i)$ for $3i$ and $(1, i)$ for $-3i$ — gives

$$e^{t\mathbf{A}} = \begin{pmatrix} \cos 3t & -\sin 3t \\ \sin 3t & \cos 3t \end{pmatrix}, \quad \mathbf{x}(t) = \begin{pmatrix} -\sin 3t \\ \cos 3t \end{pmatrix}.$$

Check. $\mathbf{x}'(t) = (-3\cos 3t, -3\sin 3t)$ and $\mathbf{A}\mathbf{x}(t) = (-3\cos 3t, -3\sin 3t)$; and $\mathbf{x}(0) = (0, 1)$. The solution runs around the unit circle: no growth, since the eigenvalues are purely imaginary.

- (c) Put $\mathbf{u} = (y, y', y'')$. Then $\mathbf{u}' = (y', y'', 0) = \mathbf{M}\mathbf{u}$ with $\mathbf{M} = \mathbf{J}_3(0)$, and $\mathbf{u}(0) = (1, 0, 2)$. By [Corollary 10.9.5](#) with $\lambda = 0$ and $k = 3$,

$$e^{t\mathbf{M}} = \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{u}(t) = e^{t\mathbf{M}} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 + t^2 \\ 2t \\ 2 \end{pmatrix}.$$

So $y(t) = 1 + t^2$. *Check.* $y' = 2t$, $y'' = 2$, $y''' = 0$, and the three initial values are 1, 0, 2. The eigenvalue 0 with one block of size 3 gives polynomial solutions $t^j e^{0t} = t^j$ for $j < 3$.

Exercise. B3: A square root and a logarithm

- (a) Let $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Compute $f(\mathbf{A})$ for $f(x) = \sqrt{x}$ (positive branch) and verify that $f(\mathbf{A})^2 = \mathbf{A}$.
 (b) Let $\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Compute $f(\mathbf{B})$ for $f(x) = \log x$ with $\log 1 = 0$, and verify that

$$e^{f(\mathbf{B})} = \mathbf{B}.$$

- (c) Explain why the same method produces **no** square root of $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Solution.

- (a) $p_{\mathbf{A}} = (x - 2)^2 - 1 = (x - 1)(x - 3)$, eigenvalues 1, 3. Eigenvectors $(1, -1)$ for 1 and $(1, 1)$ for 3. With $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ and $\mathbf{P}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$,

$$f(\mathbf{A}) = \mathbf{P} \operatorname{diag}(1, \sqrt{3}) \mathbf{P}^{-1} = \frac{1}{2} \begin{pmatrix} 1 + \sqrt{3} & -1 + \sqrt{3} \\ -1 + \sqrt{3} & 1 + \sqrt{3} \end{pmatrix}.$$

Squaring, the diagonal entry is $\frac{1}{4}[(1 + \sqrt{3})^2 + (\sqrt{3} - 1)^2] = \frac{1}{4}(4 + 2\sqrt{3} + 4 - 2\sqrt{3}) = 2$, and the off-diagonal entry is $\frac{1}{4} \cdot 2(1 + \sqrt{3})(\sqrt{3} - 1) = \frac{1}{2}(3 - 1) = 1$. So $f(\mathbf{A})^2 = \mathbf{A}$.

- (b) Here $m_{\mathbf{B}} = (x - 1)^2$, since $(\mathbf{B} - \mathbf{I})^2 = \mathbf{O} \neq \mathbf{B} - \mathbf{I}$, so by [Theorem 10.9.2](#) the answer is $p(\mathbf{B})$ for the polynomial of degree < 2 with $p(1) = f(1) = 0$ and $p'(1) = f'(1) = 1$, namely $p = x - 1$. Hence

$$f(\mathbf{B}) = \mathbf{B} - \mathbf{I} = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}.$$

Its square is $\mathbf{O}$, so $e^{f(\mathbf{B})} = \mathbf{I} + f(\mathbf{B}) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \mathbf{B}$. The Jordan route gives the

same thing: with $\mathbf{Q} = \operatorname{diag}(2, 1)$ we get $\mathbf{Q}^{-1}\mathbf{B}\mathbf{Q} = \mathbf{J}_2(1)$, and $f(\mathbf{J}_2(1)) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, so

$$f(\mathbf{B}) = \mathbf{Q}f(\mathbf{J}_2(1))\mathbf{Q}^{-1} = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}.$$

- (c) The matrix is $\mathbf{J}_2(0)$, with the single eigenvalue 0 and a block of size 2. [Definition 10.9.1](#) needs f and f' at 0, and $f(x) = \sqrt{x}$ is not differentiable at 0 — the entry $f'(0)$ does not exist. In fact $\mathbf{J}_2(0)$ has no square root at all: if $\mathbf{X}^2 = \mathbf{J}_2(0)$, then $\mathbf{X}^4 = \mathbf{J}_2(0)^2 = \mathbf{O}$, so $\mathbf{X}$ is nilpotent, so $\mathbf{X}^2 = \mathbf{O}$ by Proposition 10.2.1 (a) with $n = 2$, contradicting $\mathbf{X}^2 = \mathbf{J}_2(0) \neq \mathbf{O}$.

10.9.5.3 C. Going deeper

Exercise. C1: How badly the exponential law fails

Let $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ in $M_2(\mathbb{R})$.

- Compute $e^{\mathbf{A}}$, $e^{\mathbf{B}}$, $e^{\mathbf{A}}e^{\mathbf{B}}$, and $e^{\mathbf{B}}e^{\mathbf{A}}$.
- Compute $e^{\mathbf{A}+\mathbf{B}}$ in closed form. *Hint:* $(\mathbf{A} + \mathbf{B})^2 = \mathbf{I}$, so split the series into even and odd terms.
- Verify $\det e^{\mathbf{A}+\mathbf{B}} = e^{\operatorname{tr}(\mathbf{A}+\mathbf{B})}$ and $\det(e^{\mathbf{A}}e^{\mathbf{B}}) = e^{\operatorname{tr}\mathbf{A}}e^{\operatorname{tr}\mathbf{B}}$, and explain why equal determinants do not rescue the identity.
- Show that $e^{\mathbf{A}}e^{\mathbf{B}} = e^{\mathbf{C}}$ for **some** $\mathbf{C} \in M_2(\mathbb{R})$, and find $\operatorname{tr} \mathbf{C}$.

Solution.

(a) $\mathbf{A}^2 = \mathbf{B}^2 = \mathbf{O}$, so $e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $e^{\mathbf{B}} = \mathbf{I} + \mathbf{B} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Then

$$e^{\mathbf{A}}e^{\mathbf{B}} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad e^{\mathbf{B}}e^{\mathbf{A}} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

(b) Let $\mathbf{K} = \mathbf{A} + \mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then $\mathbf{K}^2 = \mathbf{I}$, so $\mathbf{K}^{2l} = \mathbf{I}$ and $\mathbf{K}^{2l+1} = \mathbf{K}$. Splitting the partial sums by parity, which is legitimate because each entry sequence is a sum of two convergent sequences (A3),

$$\begin{aligned} e^{\mathbf{K}} &= \left(\sum_{l \geq 0} \frac{1}{(2l)!} \right) \mathbf{I} + \left(\sum_{l \geq 0} \frac{1}{(2l+1)!} \right) \mathbf{K} \\ &= \frac{e + e^{-1}}{2} \mathbf{I} + \frac{e - e^{-1}}{2} \mathbf{K} = \begin{pmatrix} \frac{e+e^{-1}}{2} & \frac{e-e^{-1}}{2} \\ \frac{e-e^{-1}}{2} & \frac{e+e^{-1}}{2} \end{pmatrix}, \end{aligned}$$

where the two scalar sums were identified by adding and subtracting the series for e^1 and e^{-1} term by term (A1). Numerically the diagonal entries are about 1.543, not 2, so $e^{\mathbf{A}+\mathbf{B}} \neq e^{\mathbf{A}}e^{\mathbf{B}}$.

(c) $\text{tr}(\mathbf{A} + \mathbf{B}) = 0$ and $\det e^{\mathbf{A}+\mathbf{B}} = \left(\frac{e+e^{-1}}{2} \right)^2 - \left(\frac{e-e^{-1}}{2} \right)^2 = \frac{1}{4}[(e+e^{-1})^2 - (e-e^{-1})^2] = \frac{1}{4} \cdot 4 = 1 = e^0$. Also $\det(e^{\mathbf{A}}e^{\mathbf{B}}) = 2 - 1 = 1 = e^0e^0$. The determinants agree because both matrices are products of exponentials of trace-zero matrices; a determinant is one number and cannot distinguish two matrices, so it gives no information about the identity either way. The entries do.

(d) $\mathbf{X} = e^{\mathbf{A}}e^{\mathbf{B}} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ is symmetric with $\det \mathbf{X} = 1$ and $\text{tr} \mathbf{X} = 3$, so $p_{\mathbf{X}} = x^2 - 3x + 1$,

whose roots $\frac{3 \pm \sqrt{5}}{2}$ are real, distinct, and positive. So $\mathbf{X} = \mathbf{P} \text{diag}(\alpha, \beta) \mathbf{P}^{-1}$ with $\alpha\beta = 1$, $\alpha, \beta > 0$. Put $\mathbf{C} = \mathbf{P} \text{diag}(\log \alpha, \log \beta) \mathbf{P}^{-1}$, a real matrix because $\mathbf{P}$ can be taken real (the eigenvalues are real) and the logarithms are real. By Definition 10.9.1 applied to $f = \exp$, $e^{\mathbf{C}} = \mathbf{P} \text{diag}(\alpha, \beta) \mathbf{P}^{-1} = \mathbf{X}$. Finally $\text{tr} \mathbf{C} = \log \alpha + \log \beta = \log(\alpha\beta) = \log 1 = 0$, as Theorem 10.9.4 (e) requires, since $\det \mathbf{X} = 1$.

Exercise. C2: When does every solution die out?

Let $\mathbf{A} \in M_n(\mathbb{C})$. Prove that the following are equivalent:

- (i) every solution of $\mathbf{x}' = \mathbf{A}\mathbf{x}$ satisfies $\mathbf{x}(t) \rightarrow \mathbf{0}$ entrywise as $t \rightarrow \infty$;
- (ii) every eigenvalue λ of $\mathbf{A}$ has $\text{Re } \lambda < 0$.

Hint: compute the entries of $e^{t\mathbf{J}_k(\lambda)}$ and use (A5). For the failing direction, an eigenvector gives a solution you can write down.

Solution.

- (ii) $\Rightarrow$ (i). By Theorem 10.9.6 every solution is $\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}_0$, so it suffices to show $e^{t\mathbf{A}} \rightarrow \mathbf{0}$ entrywise. Write $\mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1}$ with $\mathbf{J} = \bigoplus_i \mathbf{J}_{k_i}(\lambda_i)$ (Theorem 10.3.1). Each entry of

$e^{t\mathbf{A}} = \mathbf{P}e^{t\mathbf{J}}\mathbf{P}^{-1}$ is a fixed linear combination of entries of $e^{t\mathbf{J}}$, so by (A3) it is enough that $e^{t\mathbf{J}} \rightarrow \mathbf{O}$. By [Corollary 10.9.5](#), the nonzero entries of $e^{t\mathbf{J}_k(\lambda)}$ are

$$\frac{t^j e^{\lambda t}}{j!} \quad (0 \leq j \leq k-1).$$

Writing $\lambda = c + i\omega$ with $c = \operatorname{Re} \lambda < 0$, we get $|t^j e^{\lambda t}| = t^j e^{ct} |e^{i\omega t}| = t^j e^{ct}$ for $t \geq 0$, since $|e^{i\omega t}| = |\cos \omega t + i \sin \omega t| = 1$ by (A1). By (A5) this tends to 0, and a complex sequence tends to 0 exactly when its moduli do. Hence $e^{t\mathbf{J}} \rightarrow \mathbf{O}$.

- (iii) $\Rightarrow$ (ii). Suppose some eigenvalue λ has $\operatorname{Re} \lambda = c \geq 0$, and let $\mathbf{v} \neq \mathbf{0}$ be an eigenvector. Then $\mathbf{x}(t) = e^{\lambda t} \mathbf{v}$ is a solution: it is differentiable with $\mathbf{x}'(t) = \lambda e^{\lambda t} \mathbf{v} = e^{\lambda t} \mathbf{A} \mathbf{v} = \mathbf{A} \mathbf{x}(t)$, and $\mathbf{x}(0) = \mathbf{v}$. Pick a coordinate with $v_r \neq 0$; then

$$|x_r(t)| = |e^{\lambda t}| |v_r| = e^{ct} |v_r| \geq |v_r| > 0$$

for every $t \geq 0$, because $e^{ct} \geq 1$ when $c \geq 0$. So $x_r(t)$ does not tend to 0, and (i) fails. This proves the equivalence.