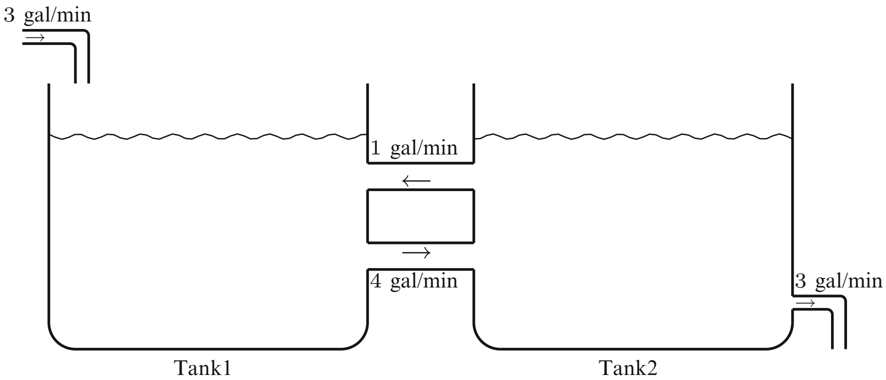
Chapter 9
Linear Systems of Differential Equations

9.1 Introduction

In previous chapters, we have discussed ordinary differential equations in a single unknown function, $y(t)$. These are adequate to model real-world systems as they evolve in time, provided that only one state, that is, the number $y(t)$, is needed to describe the system. For instance, we might be interested in the temperature of an object, the concentration of a pollutant in a lake, or the displacement of a weight attached to a spring. In each of these cases, the system we wish to describe is adequately represented by a single function of time. However, a single ordinary differential equation is inadequate for describing the evolution over time of a system with interdependent subsystems, each with its own state. Consider such an example.

Example 1. Two tanks are interconnected as illustrated below.



Assume that Tank 1 contains 10 gallons of brine in which 2 pounds of salt are initially dissolved and Tank 2 initially contains 10 gallons of pure water. Moreover, the mixtures are pumped between the two tanks, 4 gal/min from Tank 1 to Tank

W.A. Adkins and M.G. Davidson, *Ordinary Differential Equations*, Undergraduate Texts in Mathematics, DOI 10.1007/978-1-4614-3618-8_9,

2 and 1 gal/min going from Tank 2 back to Tank 1. Assume that a brine mixture containing 1 lb salt/gal enters Tank 1 at a rate of 3 gal/min, and the well-stirred mixture is removed from Tank 2 at the same rate of 3 gal/min. Let $y_1(t)$ be the amount of salt in Tank 1 at time t and let $y_2(t)$ be the amount of salt in Tank 2 at time t. Determine how y_1 and y_2 and their derivatives are related.

► **Solution.** The underlying principle is the same as that of the single tank mixing problem. Namely, we apply the balance equation

$$y'(t) = \text{input rate} - \text{output rate}$$

to the amount of salt in *each* tank. If $y_1(t)$ denotes the amount of salt at time t in Tank 1, then the ***concentration*** of salt at time t in Tank 1 is $c_1(t) = (y_1(t)/10)$ lb/gal. Similarly, the concentration of salt in Tank 2 at time t is $c_2(t) = (y_2(t)/10)$ lb/gal. The input and output rates are determined by the product of the concentration and the flow rate of the fluid at time t. The relevant rates of change are summarized in the following table.

From	To	Rate
Outside	Tank 1	$(1\text{ lb/gal}) \cdot (3\text{ gal/min}) = 3$ lb/min
Tank 1	Tank2	$\frac{y_1(t)}{10}\text{lb/gal} \cdot 4\text{ gal/min} = \frac{4y_1(t)}{10}$ lb/min
Tank 2	Tank 1	$\frac{y_2(t)}{10}\text{lb/gal} \cdot 1\text{ gal/min} = \frac{y_2(t)}{10}$ lb/min
Tank 2	Outside	$\frac{y_2(t)}{10}\text{lb/gal} \cdot 3\text{ gal/min} = \frac{3y_2(t)}{10}$ lb/min

The data for the balance equations can then be read from the following table:

Tank	Input rate	Output rate
1	$3 + \frac{y_2(t)}{10}$	$\frac{4y_1(t)}{10}$
2	$\frac{4y_1(t)}{10}$	$\frac{4y_2(t)}{10}$

Putting these data in the balance equations then gives

$$\begin{aligned} y_1'(t) &= -\frac{4}{10}y_1(t) + \frac{1}{10}y_2(t) + 3, \\ y_2'(t) &= \frac{4}{10}y_1(t) - \frac{4}{10}y_2(t). \end{aligned}$$

These equations thus describe the relationship between y_1 and y_2 and their derivatives. We observe also that the statement of the problem includes initial conditions, namely, $y_1(0) = 2$ and $y_2(0) = 0$. ◄

These two equations together comprise an example of a (constant coefficient) linear system of ordinary differential equations. Notice that two states are involved: the amount of salt in each tank, $y_1(t)$ and $y_2(t)$. Notice also that y_1' depends not only on y_1 but also y_2 and likewise for y_2'. When such occurs, we say that y_1 and y_2 are coupled. In order to find one, we need the other. This chapter is devoted to developing theory and solution methods for such equations. Before we discuss such methods, let us lay down the salient definitions, notation, and basic facts.

9.2 Linear Systems of Differential Equations

A system of equations of the form

$$\begin{aligned} y_1'(t) &= a_{11}(t)y_1(t) + \cdots + a_{1n}(t)y_n(t) + f_1(t) \\ y_2'(t) &= a_{21}(t)y_1(t) + \cdots + a_{2n}(t)y_n(t) + f_2(t) \\ &\vdots \\ y_n'(t) &= a_{n1}(t)y_1(t) + \cdots + a_{nn}(t)y_n(t) + f_n(t), \end{aligned} \tag{1}$$

where $a_{ij}(t)$ and $f_i(t)$ are functions defined on some common interval, is called a ***first order linear system of ordinary differential equations*** or just ***linear differential system***, for short. If

$$\boldsymbol{y}(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{bmatrix}, \quad A(t) = \begin{bmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{bmatrix}, \text{ and } \quad \boldsymbol{f}(t) = \begin{bmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{bmatrix}$$

then (1) can be written more succinctly in matrix form as

$$\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t) + \boldsymbol{f}(t). \tag{2}$$

If $A(t) = A$ is a matrix of constants, the linear differential system is said to be ***constant coefficient***. A linear differential system is ***homogeneous*** if $\boldsymbol{f} = 0$; otherwise it is ***nonhomogeneous***. The homogeneous linear differential system obtained from (2) (equivalently (1)) by setting $\boldsymbol{f} = \boldsymbol{0}$ (equivalently setting each $f_i(t) = 0$), namely,

$$\boldsymbol{y}' = A(t)\boldsymbol{y}, \tag{3}$$

is known as the ***associated homogeneous equation*** for the system (2).

As was the case for a single differential equation, it is conventional to suppress the independent variable t in the unknown functions $y_i(t)$ and their derivatives $y_i'(t)$. Thus, (1) and (2) would normally be written as

$$\begin{aligned} y_1' &= a_{11}(t)y_1 + \cdots + a_{1n}(t)y_n + f_1(t) \\ y_2' &= a_{21}(t)y_1 + \cdots + a_{2n}(t)y_n + f_2(t) \\ &\vdots \\ y_n' &= a_{n1}(t)y_1 + \cdots + a_{nn}(t)y_n + f_n(t), \end{aligned} \tag{4}$$

and

$$\boldsymbol{y}' = A(t)\boldsymbol{y} + \boldsymbol{f}(t), \tag{5}$$

respectively.

Example 1. For each of the following systems, determine which is a linear differential system. If it is, write it in matrix form and determine whether it is homogeneous and whether it is constant coefficient:

1.
$$\begin{aligned} y_1' &= (1-t)y_1 + \mathrm{e}^t y_2 + (\sin t)y_3 + 1 \\ y_2' &= 3y_1 + t y_2 + (\cos t)y_3 + t\mathrm{e}^{2t} \\ y_3' &= y_3 \end{aligned}$$

2.
$$\begin{aligned} y_1' &= a y_1 - b y_1 y_2 \\ y_2' &= -c y_1 + d y_1 y_2 \end{aligned}$$

3.
$$\begin{aligned} y_1' &= -\frac{4}{10} y_1 + \frac{1}{10} y_2 + 3 \\ y_2' &= \frac{4}{10} y_1 - \frac{4}{10} y_2 \end{aligned}$$

► **Solution.** 1. This is a linear differential system with

$$A(t) = \begin{bmatrix} 1-t & \mathrm{e}^t & \sin t \\ 3 & t & \cos t \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{f}(t) = \begin{bmatrix} 1 \\ t\mathrm{e}^{2t} \\ 0 \end{bmatrix}.$$

Since $\boldsymbol{f} \neq 0$, this system is nonhomogeneous, and since $A(t)$ is not a constant function, this system is not a constant coefficient linear differential system.

2. This system is not a *linear* differential system because of the presence of the products $y_1 y_2$.
3. This is a linear differential system with

$$A(t) = A = \begin{bmatrix} -\frac{4}{10} & \frac{1}{10} \\ \frac{4}{10} & -\frac{4}{10} \end{bmatrix} \quad \text{and} \quad \boldsymbol{f} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}.$$

It is constant coefficient but nonhomogeneous. This is the linear differential system we introduced in Example 1 of Sect. 9.1. ◄

If all the entry functions of $A(t)$ and $\boldsymbol{f}(t)$ are defined on a common interval, I, then a vector function $\boldsymbol{y}$, defined on I, that satisfies (2) (or, equivalently, (1)) is a ***solution***. A solution of the associated homogeneous equation $\boldsymbol{y}' = A(t)\boldsymbol{y}$ of (2) is referred to as a ***homogeneous solution*** to (2). Note that a homogeneous solution to (2) is not a solution of the given equation $\boldsymbol{y}' = A(t)\boldsymbol{y} + \boldsymbol{f}(t)$, but rather a solution of the related equation $\boldsymbol{y}' = A(t)\boldsymbol{y}$.

Example 2. Consider the following first order system of ordinary differential equations:

$$\begin{aligned} y_1' &= 3y_1 - y_2 \\ y_2' &= 4y_1 - 2y_2. \end{aligned} \tag{6}$$

Let

$$\boldsymbol{y}(t)=\begin{bmatrix}\mathrm{e}^{2t}\\ \mathrm{e}^{2t}\end{bmatrix} \quad \text{and} \quad \boldsymbol{z}(t)=\begin{bmatrix}\mathrm{e}^{-t}\\ 4\mathrm{e}^{-t}\end{bmatrix}.$$

Show that $\boldsymbol{y}(t)$, $\boldsymbol{z}(t)$, and $\boldsymbol{w}(t)=c_1\boldsymbol{y}(t)+c_2\boldsymbol{z}(t)$, where c_1 and c_2 are scalars, are solutions to (6).

► **Solution.** Let $A=\begin{bmatrix}3 & -1\\ 4 & -2\end{bmatrix}$. Then (6) can be written

$$\boldsymbol{y}'=A\boldsymbol{y}.$$

This system is a constant coefficient homogeneous linear differential system. For $\boldsymbol{y}(t)=\begin{bmatrix}\mathrm{e}^{2t}\\ \mathrm{e}^{2t}\end{bmatrix}$, we have on the one hand

$$\boldsymbol{y}'(t)=\begin{bmatrix}2\mathrm{e}^{2t}\\ 2\mathrm{e}^{2t}\end{bmatrix}$$

and on the other hand,

$$A\boldsymbol{y}(t)=\begin{bmatrix}3 & -1\\ 4 & -2\end{bmatrix}\begin{bmatrix}\mathrm{e}^{2t}\\ \mathrm{e}^{2t}\end{bmatrix}=\begin{bmatrix}3\mathrm{e}^{2t}-\mathrm{e}^{2t}\\ 4\mathrm{e}^{2t}-2\mathrm{e}^{2t}\end{bmatrix}=\begin{bmatrix}2\mathrm{e}^{2t}\\ 2\mathrm{e}^{2t}\end{bmatrix}.$$

It follows that $\boldsymbol{y}'=A\boldsymbol{y}$, and hence, $\boldsymbol{y}$ is a solution.

For $\boldsymbol{z}(t)=\begin{bmatrix}\mathrm{e}^{-t}\\ 4\mathrm{e}^{-t}\end{bmatrix}$, we have on the one hand

$$\boldsymbol{z}'(t)=\begin{bmatrix}-\mathrm{e}^{-t}\\ -4\mathrm{e}^{-t}\end{bmatrix}$$

and on the other hand,

$$A\boldsymbol{z}(t)=\begin{bmatrix}3 & -1\\ 4 & -2\end{bmatrix}\begin{bmatrix}\mathrm{e}^{-t}\\ 4\mathrm{e}^{-t}\end{bmatrix}=\begin{bmatrix}3\mathrm{e}^{-t}-4\mathrm{e}^{-t}\\ 4\mathrm{e}^{-t}-8\mathrm{e}^{-t}\end{bmatrix}=\begin{bmatrix}-\mathrm{e}^{-t}\\ -4\mathrm{e}^{-t}\end{bmatrix}.$$

Again, it follows that $\boldsymbol{z}$ is a solution.

Suppose c_1 and c_2 are *any* constants and $\boldsymbol{w}(t)=c_1\boldsymbol{y}(t)+c_2\boldsymbol{z}(t)$. Since differentiation is linear, we have $\boldsymbol{w}'(t)=c_1\boldsymbol{y}'(t)+c_2\boldsymbol{z}'(t)$. Since matrix multiplication is linear, we have

$$A\boldsymbol{w}(t)=c_1A\boldsymbol{y}(t)+c_2A\boldsymbol{z}(t)=c_1\boldsymbol{y}'(t)+c_2\boldsymbol{z}'(t)=\boldsymbol{w}'(t).$$

It follows that $\boldsymbol{w}(t)$ is a solution. ◄

Example 3. Consider the following first order system of ordinary differential equations:

$$\begin{aligned} y_1' &= 3y_1 - y_2 + 2t \\ y_2' &= 4y_1 - 2y_2 + 2. \end{aligned} \tag{7}$$

1. Verify that $\boldsymbol{y}_{\mathrm{p}}(t) = \begin{bmatrix} -2t+1 \\ -4t+5 \end{bmatrix}$ is a solution of (7).
2. Verify that $\boldsymbol{z}_{\mathrm{p}}(t) = 2\boldsymbol{y}_{\mathrm{p}}(t) = \begin{bmatrix} -4t+2 \\ -8t+10 \end{bmatrix}$ is *not* a solution to (7).
3. Verify that $\boldsymbol{y}_{\mathrm{g}}(t) = \boldsymbol{w}(t) + \boldsymbol{y}_{\mathrm{p}}(t)$ is a solution of (7), where $\boldsymbol{w}(t)$ is the general solution of (6) from the previous example.

► **Solution.** We begin by writing (7) in matrix form as:

$$\boldsymbol{y}' = A\boldsymbol{y} + \boldsymbol{f},$$

where $A = \begin{bmatrix} 3 & -1 \\ 4 & -2 \end{bmatrix}$ and $\boldsymbol{f} = \begin{bmatrix} 2t \\ 2 \end{bmatrix}$. Note that the associated homogeneous linear differential system, that is, the equation $\boldsymbol{y}' = A\boldsymbol{y}$ obtained by setting $\boldsymbol{f} = \boldsymbol{0}$, is the system from the previous example.

1. On the one hand,

$$\boldsymbol{y}_{\mathrm{p}}'(t) = \begin{bmatrix} -2 \\ -4 \end{bmatrix},$$

and on the other hand,

$$\begin{aligned} A\boldsymbol{y}_{\mathrm{p}} + \boldsymbol{f} = \begin{bmatrix} 3 & -1 \\ 4 & -2 \end{bmatrix}\begin{bmatrix} -2t+1 \\ -4t+5 \end{bmatrix} + \begin{bmatrix} 2t \\ 2 \end{bmatrix} &= \begin{bmatrix} 3(-2t+1) - (-4t+5) + 2t \\ 4(-2t+1) - 2(-4t+5) + 2 \end{bmatrix} \\ &= \begin{bmatrix} -2 \\ -4 \end{bmatrix}. \end{aligned}$$

Since $\boldsymbol{y}_{\mathrm{p}}' = A\boldsymbol{y}_{\mathrm{p}} + \boldsymbol{f}$, it follows that $\boldsymbol{y}_{\mathrm{p}}$ is a solution to (7).

2. On the one hand,

$$\boldsymbol{z}_{\mathrm{p}}' = \begin{bmatrix} -4 \\ -8 \end{bmatrix},$$

and on the other hand,

$$A\boldsymbol{z}_{\mathrm{p}} + \boldsymbol{f} = \begin{bmatrix} 3 & -1 \\ 4 & -2 \end{bmatrix}\begin{bmatrix} -4t+2 \\ -8t+10 \end{bmatrix} + \begin{bmatrix} 2t \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 3(-4t+2) - (-8t+10) + 2t \\ 4(-4t+2) - 2(-8t+10) + 2 \end{bmatrix}$$

$$= \begin{bmatrix} -2t-4 \\ -10 \end{bmatrix}.$$

Since $z_p' \neq Az_p + \boldsymbol{f}$, z_p is *not* a solution.

3. Since $\boldsymbol{y}' = A\boldsymbol{y}$ is the homogeneous linear differential system associated to (7), we know from the previous example that $\boldsymbol{w}$ is a solution to the homogeneous equation $\boldsymbol{y}' = A\boldsymbol{y}$. Since differentiation is linear, we have

$$\begin{aligned} \boldsymbol{y}_g' &= \boldsymbol{w}' + \boldsymbol{y}_p' \\ &= A\boldsymbol{w} + A\boldsymbol{y}_p + \boldsymbol{f} \\ &= A(\boldsymbol{w} + \boldsymbol{y}_p) + \boldsymbol{f} \\ &= A\boldsymbol{y}_g + \boldsymbol{f}. \end{aligned}$$

It follows that $\boldsymbol{y}_g$ is a solution to (7). ◄

These two examples illustrate the power of linearity, a concept that we have repeatedly encountered, and are suggestive of the following general statement.

Theorem 4. *Consider the linear differential system*

$$\boldsymbol{y}' = A(t)\boldsymbol{y} + \boldsymbol{f}(t). \tag{8}$$

1. *If $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are solutions to the associated homogeneous linear differential system*

$$\boldsymbol{y}' = A(t)\boldsymbol{y}, \tag{9}$$

and c_1 and c_2 are constants, then $c_1\boldsymbol{y}_1 + c_2\boldsymbol{y}_2$ is also a solution to (9).

2. *If $\boldsymbol{y}_p$ is a fixed particular solution to (8) and $\boldsymbol{y}_h$ is any homogeneous solution (i.e., any solution to (9)), then*

$$\boldsymbol{y}_h + \boldsymbol{y}_p$$

is also a solution to (8) and all solutions to (8) are of this form.

Proof. Let $\boldsymbol{L} = \boldsymbol{D} - A(t)$ be the operator on vector-valued functions given by

$$\boldsymbol{L}\boldsymbol{y} = (\boldsymbol{D} - A(t))\boldsymbol{y} = \boldsymbol{y}' - A(t)\boldsymbol{y}.$$

Then $\boldsymbol{L}$ is linear since differentiation and matrix multiplication are linear. The rest of the proof follows the same line of argument given in the proof of Theorem 6 of Sect. 3.1. □

Linear differential systems also satisfy the superposition principle, an analogue to Theorem 8 of Sect. 3.4.

Theorem 5. *Suppose $\boldsymbol{y}_{p_1}$ is a solution to $\boldsymbol{y}' = A(t)\boldsymbol{y} + \boldsymbol{f}_1(t)$ and $\boldsymbol{y}_{p_2}$ is a solution to $\boldsymbol{y}' = A(t)\boldsymbol{y} + \boldsymbol{f}_2(t)$. Then $\boldsymbol{y} = \boldsymbol{y}_{p_1}(t) + \boldsymbol{y}_{p_2}(t)$ is a solution to $\boldsymbol{y}' = A(t)\boldsymbol{y} + \boldsymbol{f}_1(t) + \boldsymbol{f}_2(t)$.*

Proof. Let $\boldsymbol{L} = \boldsymbol{D} - A(t)$ be as above. Then linearity implies

$$\boldsymbol{L}(\boldsymbol{y}_{p_1} + \boldsymbol{y}_{p_2}) = \boldsymbol{L}\boldsymbol{y}_{p_1} + \boldsymbol{L}\boldsymbol{y}_{p_2} = \boldsymbol{f}_1(t) + \boldsymbol{f}_2(t).$$ □

Initial Value Problems

For a linear differential system $\boldsymbol{y}' = A(t)\boldsymbol{y} + \boldsymbol{f}(t)$, assume the entries of $A(t)$ and $\boldsymbol{f}(t)$ are defined on a common interval I. Let $t_0 \in I$. When we associate to a linear differential system an initial value $\boldsymbol{y}(t_0) = \boldsymbol{y}_0$, we call the resulting problem an ***initial value problem***. The mixing problem we discussed in the introduction to this chapter is an example of an initial value problem:

$$\boldsymbol{y}' = A\boldsymbol{y} + \boldsymbol{f}, \quad \boldsymbol{y}(0) = \boldsymbol{y}_0,$$

where $A = \begin{bmatrix} -\frac{4}{10} & \frac{1}{10} \\ \frac{4}{10} & -\frac{4}{10} \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$, and $\boldsymbol{y}(0) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

Linear Differential Equations and Systems

For each linear *ordinary* differential equation $\boldsymbol{L}y = f$ with initial conditions, we can construct a corresponding linear system with initial condition. The solution of one will imply the solution of the other. The following example will illustrate the procedure.

Example 6. Construct a first order linear differential system with initial value from the second order differential equation

$$y'' + ty' + y = \sin t \quad y(0) = 1,\ y'(0) = 2.$$

► **Solution.** If $y(t)$ is a solution to the second order equation, form a vector function $\boldsymbol{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$ by setting $y_1(t) = y(t)$ and $y_2(t) = y'(t)$. Then

$$\begin{aligned} y_1'(t) &= y'(t) = y_2(t), \\ y_2'(t) &= y''(t) = -y(t) - ty'(t) + \sin t = -y_1(t) - ty_2(t) + \sin t. \end{aligned}$$

The second equation is obtained by solving $y'' + ty' + y = \sin t$ for y'' and substituting y_1 for y and y_2 for y'. In vector form, this equation becomes

$$\boldsymbol{y}' = A(t)\boldsymbol{y} + \boldsymbol{f},$$

where $A(t) = \begin{bmatrix} 0 & 1 \\ -1 & -t \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} 0 \\ \sin t \end{bmatrix}$, and $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. ◄

The solution to the first order differential system implies a solution to the original second order differential equation and vice versa. Specifically, if $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ is a solution of the system, then the first entry of $\boldsymbol{y}$, namely, y_1, is a solution of the second order equation, and conversely, as illustrated in the above example, if y is a solution of the second order equation, then $\boldsymbol{y} = \begin{bmatrix} y \\ y' \end{bmatrix}$ is a solution of the linear system.

Linear differential equations of order n are transformed into linear systems in a similar way.

Extension of Basic Definitions and Operations to Matrix-Valued Functions

It is convenient to state most of our results on linear systems of ordinary differential equations in the language of matrices. To this end, we extend several definitions familiar for real-valued functions to matrix-(or vector-)valued functions. Let $v(t)$ be an $n \times m$ matrix-valued function with entries $v_{i,j}(t)$, for $i = 1 \ldots n$ and $j = 1 \ldots m$.

1. $v(t)$ is ***defined*** on an interval I of $\mathbb{R}$ if each $v_{ij}(t)$ is defined on I.
2. $v(t)$ is ***continuous*** on an interval I of $\mathbb{R}$ if each $v_{ij}(t)$ is continuous on I. For instance, the matrix

$$v(t) = \begin{bmatrix} \dfrac{1}{t+2} & \cos 2t \\ e^{-2t} & \dfrac{1}{(2t-3)^2} \end{bmatrix}$$

   is continuous on each of the intervals $I_1 = (-\infty, -2)$, $I_2 = (-2, 3/2)$ and $I_3 = (3/2, \infty)$, but it is not continuous on the interval $I_4 = (0, 2)$.
3. $v(t)$ is ***differentiable*** on an interval I of $\mathbb{R}$ if each $v_{ij}(t)$ is differentiable on I. Moreover, $v'(t) = [a'_{ij}(t)]$. That is, the matrix $v(t)$ is differentiated by differentiating each entry of the matrix. For instance, for the matrix

$$v(t) = \begin{bmatrix} e^t & \sin t & t^2 + 1 \\ \ln t & \cos t & \sinh t \end{bmatrix},$$

we have

$$v'(t) = \begin{bmatrix} e^t & \cos t & 2t \\ 1/t & -\sin t & \cosh t \end{bmatrix}.$$

4. An ***antiderivative*** of $v(t)$ is a matrix-valued function $V(t)$ (necessarily of the same size) such that $V'(t) = v(t)$. Since the derivative is calculated entry by entry, so likewise is the antiderivative. Thus, if

$$v(t) = \begin{bmatrix} e^{4t} & \sin t \\ 2t & \ln t \\ \cos 2t & 5 \end{bmatrix},$$

then an antiderivative is

$$\begin{aligned} V(t) = \int v(t)\mathrm{d}t &= \begin{bmatrix} \frac{1}{4}e^{4t} + c_{11} & -\cos t + c_{12} \\ t^2 + c_{21} & t\ln t - t + c_{22} \\ \frac{1}{2}\sin 2t + c_{31} & 5t + c_{32} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{4}e^{4t} & -\cos t \\ t^2 & t\ln t - t \\ \frac{1}{2}\sin 2t & 5t \end{bmatrix} + C, \end{aligned}$$

where C is the matrix of constants $[c_{ij}]$. Thus, if $v(t)$ is defined on an interval I and $V_1(t)$ and $V_2(t)$ are two antiderivatives of $v(t)$, then they differ by a constant matrix.

5. The integral of $v(t)$ on the interval $[a, b]$ is computed by computing the integral of each entry of the matrix over $[a, b]$, that is, $\int_a^b v(t)\,\mathrm{d}t = \left[\int_a^b v_{ij}(t)\,\mathrm{d}t\right]$. For the matrix $v(t)$ of item 2 above, this gives

$$\int_0^1 v(t)\,\mathrm{d}t = \begin{bmatrix} \displaystyle\int_0^1 \frac{1}{t+2}\,\mathrm{d}t & \displaystyle\int_0^1 \cos 2t\,\mathrm{d}t \\ \displaystyle\int_0^1 e^{-2t}\,\mathrm{d}t & \displaystyle\int_0^1 \frac{1}{(2t-3)^2}\,\mathrm{d}t \end{bmatrix} = \begin{bmatrix} \ln\frac{3}{2} & \frac{1}{2}\sin 2 \\ \frac{1}{2}(1-e^{-2}) & \frac{1}{3} \end{bmatrix}.$$

6. If each entry $v_{ij}(t)$ of $v(t)$ is of exponential type (recall the definition on page 111), we can take the ***Laplace transform of*** $v(t)$, by taking the Laplace transform of each entry. That is, $\mathcal{L}(v(t))(s) = \left[\mathcal{L}(v_{ij}(t))(s)\right]$. For example, if $v(t) = \begin{bmatrix} te^{-2t} & \cos 2t \\ e^{3t}\sin t & (2t-3)^2 \end{bmatrix}$, this gives

$$\mathcal{L}(v(t))(s) = \begin{bmatrix} \mathcal{L}\left(te^{-2t}\right)(s) & \mathcal{L}(\cos 2t)(s) \\ \mathcal{L}(e^{3t}\sin t)(s) & \mathcal{L}\left((2t-3)^2\right)(s) \end{bmatrix} = \begin{bmatrix} \frac{1}{(s+2)^2} & \frac{2}{s^2+4} \\ \frac{1}{(s-3)^2+1} & \frac{9s^2-12s+8}{s^3} \end{bmatrix}.$$

7. We define the ***inverse Laplace transform*** entry by entry as well. For example, if

$$V(s) = \begin{bmatrix} \frac{1}{s-1} & \frac{1}{(s-1)^2} \\ \frac{s}{s^2+4} & \frac{2}{(s-1)^2+4} \end{bmatrix},$$

then

$$\mathcal{L}^{-1}\{V(s)\}(t) = \begin{bmatrix} e^t & te^t \\ \cos 2t & e^t \sin 2t \end{bmatrix}.$$

8. Finally, we extend convolution to matrix products. Suppose $v(t)$ and $w(t)$ are matrix-valued functions such that the usual matrix product $v(t)w(t)$ is defined. We define the convolution $v(t) * w(t)$ as follows:

$$(v(t) * w(t))_{i,j} = \sum_k v_{i,k} * w_{k,j}(t).$$

Thus, in the matrix product, we replace each product of terms by the convolution product. For example, if

$$v(t) = \begin{bmatrix} e^t & e^{2t} \\ -e^{2t} & 2e^t \end{bmatrix} \quad \text{and} \quad w(t) = \begin{bmatrix} 3e^t \\ -e^t \end{bmatrix},$$

then

$$\begin{aligned} v * w(t) &= \begin{bmatrix} e^t & e^{2t} \\ -e^{2t} & 2e^t \end{bmatrix} * \begin{bmatrix} 3e^t \\ -e^t \end{bmatrix} \\ &= \begin{bmatrix} 3e^t * e^t - e^{2t} * e^t \\ -3e^{2t} * e^t - 2e^t * e^t \end{bmatrix} \\ &= \begin{bmatrix} 3te^t - (e^{2t} - e^t) \\ -3(e^{2t} - e^t) - 2te^t \end{bmatrix} \\ &= te^t \begin{bmatrix} 3 \\ -2 \end{bmatrix} + (e^{2t} - e^t) \begin{bmatrix} -1 \\ -3 \end{bmatrix}. \end{aligned}$$

Alternately, if we write

$$v(t) = e^t \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} + e^{2t} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{and} \quad w(t) = e^t \begin{bmatrix} 3 \\ -1 \end{bmatrix},$$

then the preceding calculations can be performed as follows:

$$v * w(t) = \left(e^t \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} + e^{2t} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right) * e^t \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

$$= e^t * e^t \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} + e^{2t} * e^t \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

$$= te^t \begin{bmatrix} 3 \\ -2 \end{bmatrix} + (e^{2t} - e^t) \begin{bmatrix} -1 \\ -3 \end{bmatrix}.$$

The following theorem extends basic operations of calculus and Laplace transforms to matrix-valued functions.

Theorem 7. *Assume $v(t)$ and $w(t)$ are matrix-valued functions, $g(t)$ is a real-valued function, and A is a matrix of constants.*

1. *Suppose v is differentiable and the product $Av(t)$ is defined. Then $Av(t)$ is differentiable and*
$$(Av(t))' = Av'(t).$$
2. *Suppose v is differentiable and the product $v(t)A$ is defined. Then $v(t)A$ is differentiable and*
$$(v(t)A)' = v'(t)A.$$
3. *Suppose the product, $v(t)w(t)$, is defined and v and w are both differentiable. Then*
$$(v(t)w(t))' = v'(t)w(t) + v(t)w'(t).$$
4. *Suppose the composition $v(g(t))$ is defined. Then $(v(g(t)))' = v'(g(t))g'(t)$.*
5. *Suppose v is integrable over the interval $[a,b]$ and the product $Av(t)$ is defined. Then $Av(t)$ is integrable over the interval $[a,b]$ and*
$$\int_a^b Av(t)\,dt = A\int_a^b v(t)\,dt.$$
6. *Suppose v is integrable over the interval $[a,b]$ and the product $v(t)A$ is defined. The $v(t)A$ is integrable over the interval $[a,b]$ and*
$$\int_a^b v(t)A\,dt = \left(\int_a^b v(t)\,dt\right)A.$$
7. *Suppose $v(t)$ is defined on $[0,\infty)$, has a Laplace transform, and the product $Av(t)$ is defined. Then $Av(t)$ has a Laplace transform and*
$$\mathcal{L}\{Av(t)\}(s) = A\mathcal{L}\{v(t)\}(s).$$
8. *Suppose $v(t)$ is defined on $[0,\infty)$, has a Laplace transform, and the product $v(t)A$ is defined. Then $v(t)A$ has a Laplace transform and*
$$\mathcal{L}\{v(t)A\}(s) = (\mathcal{L}\{v(t)\}(s))A.$$

9. *Suppose $v(t)$ is defined on $[0, \infty)$ and $v'(t)$ exists and has a Laplace transform. Then*

$$\mathcal{L}\left\{v'(t)\right\}(s) = s\mathcal{L}\left\{v(t)\right\}(s) - v(0).$$

10. *The convolution theorem extends as well: Suppose $v(t)$ and $w(t)$ have Laplace transforms and $v * w(t)$ is defined. Then*

$$\mathcal{L}\left\{v * w(t)\right\}(s) = \mathcal{L}\left\{v(t)\right\}(s) \cdot \mathcal{L}\left\{w(t)\right\}(s).$$

Remark 8. Where matrix multiplication is involved in these formulas it is important to preserve the order of the multiplication. It is particularly worth emphasizing this dependency on the order of multiplication in the product rule for the derivative of the product of matrix-valued functions (formula (3) above). Also note that formula (4) is just the chain rule in the context of matrix-valued functions.

Exercises

1-6. For each of the following systems of differential equations, determine if it is linear. For each of those which are linear, write it in matrix form; determine if the equation is (1) homogeneous or nonhomogeneous and (2) constant coefficient. Do *not* try to solve the equations.

1. $\begin{aligned} y_1' &= y_2 \\ y_2' &= y_1 y_2 \end{aligned}$

2. $\begin{aligned} y_1' &= y_1 + y_2 + t^2 \\ y_2' &= -y_1 + y_2 + 1 \end{aligned}$

3. $\begin{aligned} y_1' &= (\sin t) y_1 \\ y_2' &= y_1 + (\cos t) y_2 \end{aligned}$

4. $\begin{aligned} y_1' &= t \sin y_1 - y_2 \\ y_2' &= y_1 + t \cos y_2 \end{aligned}$

5. $\begin{aligned} y_1' &= y_1 \\ y_2' &= 2y_1 + y_4 \\ y_3' &= y_4 \\ y_4' &= y_2 + 2y_3 \end{aligned}$

6. $\begin{aligned} y_1' &= \frac{1}{2} y_1 - y_2 + 5 \\ y_2' &= -y_1 + \frac{1}{2} y_2 - 5 \end{aligned}$

7–10. Verify that the given vector function $\boldsymbol{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$ is a solution to the given linear differential system with the given initial value.

7.
$$\boldsymbol{y}' = \begin{bmatrix} 5 & -2 \\ 4 & -1 \end{bmatrix} \boldsymbol{y}; \quad \boldsymbol{y}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \boldsymbol{y}(t) = \begin{bmatrix} e^t - e^{3t} \\ 2e^t - e^{3t} \end{bmatrix}.$$

8.
$$\boldsymbol{y}' = \begin{bmatrix} 3 & -1 \\ 4 & -1 \end{bmatrix} \boldsymbol{y}; \quad \boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \boldsymbol{y}(t) = \begin{bmatrix} e^t + 2te^t \\ 4te^t \end{bmatrix}.$$

9.
$$\boldsymbol{y}' = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix} \boldsymbol{y} + \begin{bmatrix} e^t \\ e^t \end{bmatrix}; \quad \boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \qquad \boldsymbol{y}(t) = \begin{bmatrix} e^{-t} + te^t \\ 3e^{-t} + te^t \end{bmatrix}.$$

10.

$$\boldsymbol{y} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{y} + \begin{bmatrix} t \\ -t \end{bmatrix}, \quad \boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \boldsymbol{y}(t) = \begin{bmatrix} 1 - t + 2\sin t \\ -1 - t + 2\cos t \end{bmatrix}.$$

11–15. Rewrite each of the following initial value problems for an ordinary differential equation as an initial value problem for a first order system of ordinary differential equations.

11. $y'' + 5y' + 6y = e^{2t}, \quad y(0) = 1, y'(0) = -2.$
12. $y'' + k^2 y = 0, \quad y(0) = -1, y'(0) = 0$
13. $y'' - k^2 y = A\cos \omega t, \quad y(0) = 0, y'(0) = 0$
14. $ay'' + by' + cy = 0, \quad y(0) = \alpha, y'(0) = \beta$
15. $t^2 y'' + 2ty' + y = 0, \quad y(1) = -2, y'(1) = 3$

16–21. Compute the derivative of each of the following matrix functions.

16. $A(t) = \begin{bmatrix} \cos 2t & \sin 2t \\ -\sin 2t & \cos 2t \end{bmatrix}$

17. $A(t) = \begin{bmatrix} e^{-3t} & t \\ t^2 & e^{2t} \end{bmatrix}$

18. $A(t) = \begin{bmatrix} e^{-t} & te^{-t} & t^2 e^{-t} \\ 0 & e^{-t} & te^{-t} \\ 0 & 0 & e^{-t} \end{bmatrix}$

19. $\boldsymbol{y}(t) = \begin{bmatrix} t \\ t^2 \\ \ln t \end{bmatrix}$

20. $A(t) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

21. $\boldsymbol{v}(t) = \begin{bmatrix} e^{-2t} & \ln(t^2 + 1) & \cos 3t \end{bmatrix}$

22–25. For each of the following matrix functions, compute the requested integral.

22. Compute $\int_0^{\frac{\pi}{2}} A(t)\, dt$ if $A(t) = \begin{bmatrix} \cos 2t & \sin 2t \\ -\sin 2t & \cos 2t \end{bmatrix}$.

23. Compute $\int_0^1 A(t)\, dt$ if $A(t) = \dfrac{1}{2} \begin{bmatrix} e^{2t} + e^{-2t} & e^{2t} - e^{-2t} \\ e^{-2t} - e^{2t} & e^{2t} + e^{-2t} \end{bmatrix}$.

24. Compute $\int_1^2 \boldsymbol{y}(t)\,\mathrm{d}t$ if $\boldsymbol{y}(t) = \begin{bmatrix} t \\ t^2 \\ \ln t \end{bmatrix}$.

25. Compute $\int_1^5 A(t)\,\mathrm{d}t$ if $A(t) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

26. On which of the following intervals is the matrix function $A(t) = \begin{bmatrix} t & (t+1)^{-1} \\ (t-1)^{-2} & t+6 \end{bmatrix}$ continuous?

(a) $I_1 = (-1, 1)$ (b) $I_2 = (0, \infty)$ (c) $I_3 = (-1, \infty)$
(d) $I_4 = (-\infty, -1)$ (e) $I_5 = (2, 6)$

27–32. Compute the Laplace transform of each of the following matrix functions.

27. $A(t) = \begin{bmatrix} 1 & t \\ t^2 & \mathrm{e}^{2t} \end{bmatrix}$

28. $A(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$

29. $A(t) = \begin{bmatrix} t^3 & t\sin t & t\mathrm{e}^{-t} \\ t^2 - t & \mathrm{e}^{3t}\cos 2t & 3 \end{bmatrix}$

30. $A(t) = \begin{bmatrix} t \\ t^2 \\ t^3 \end{bmatrix}$

31. $A(t) = \mathrm{e}^{t}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \mathrm{e}^{-t}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$

32. $A(t) = \begin{bmatrix} 1 & \sin t & 1 - \cos t \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{bmatrix}$

33–36. Compute the inverse Laplace transform of each matrix function:

33. $\begin{bmatrix} \dfrac{1}{s} & \dfrac{2}{s^2} & \dfrac{6}{s^3} \end{bmatrix}$

34. $\begin{bmatrix} \dfrac{1}{s} & \dfrac{1}{s^2} \\ \dfrac{s}{s^2-1} & \dfrac{s}{s^2+1} \end{bmatrix}$

35. $$\begin{bmatrix} \dfrac{2s}{s^2-1} & \dfrac{2}{s^2-1} \\ \dfrac{2}{s^2-1} & \dfrac{2s}{s^2-1} \end{bmatrix}$$

36. $$\begin{bmatrix} \dfrac{1}{s-1} & \dfrac{1}{s^2-2s+1} \\ \dfrac{4}{s^3+2s^2-3s} & \dfrac{1}{s^2+1} \\ \dfrac{3s}{s^2+9} & \dfrac{1}{s-3} \end{bmatrix}$$

9.3 The Matrix Exponential and Its Laplace Transform

One of the most basic Laplace transforms that we learned early on was that of the exponential function:

$$\mathcal{L}\left\{e^{at}\right\} = \frac{1}{s-a}. \tag{1}$$

This basic formula has proved to be a powerful tool for solving constant coefficient linear differential equations of order n. Our goal here is to extend (1) to the case where the constant a is replaced by an $n \times n$ matrix A. The resulting extension will prove to be an equally powerful tool for solving linear systems of differential equations with constant coefficients.

Let A be an $n \times n$ matrix of scalars. Formally, we define the ***matrix exponential***, e^{At}, by the formula

$$e^{At} = I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \cdots \tag{2}$$

When A is a scalar, this definition is the usual power series expansion of the exponential function. Equation (2) is an infinite sum of $n \times n$ matrices: The first term I is the $n \times n$ identity matrix, the second term is the $n \times n$ matrix At, the third term is the $n \times n$ matrix $\frac{A^2t^2}{2!}$, and so forth. To compute the sum, one must compute each (i, j) entry and add the corresponding terms. Thus, the (i, j) entry of e^{At} is

$$(e^{At})_{i,j} = (I)_{i,j} + t(A)_{i,j} + \frac{t^2}{2!}(A^2)_{i,j} + \cdots, \tag{3}$$

which is a power series centered at the origin. To determine this sum, one must be able to calculate the (i, j) entry of all the powers of A. This is easy enough for $I = A^0$ and $A = A^1$. For the (i, j) entry of A^2, we get the ith row of A times the jth column of A. For the higher powers, A^3, A^4, etc., the computations become more complicated and the resulting power series is difficult to identify, unless A is very simple. However, one can see that each entry is some power series in the variable t and thus defines a function (if it converges). In Appendix A.4, we show that the series in (3) converges absolutely for all $t \in \mathbb{R}$ and for all matrices A so that the matrix exponential is a well-defined matrix function defined on $\mathbb{R}$. However, knowing that the series converges is a far cry from knowing the sum of the series.

In the following examples, A is simple enough to allow the computation of e^{At}.

Example 1. Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. Compute e^{At}.

► **Solution.** In this case, the powers of the matrix A are easy to compute. In fact

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \cdots, \quad A^n = \begin{bmatrix} 2^n & 0 & 0 \\ 0 & 3^n & 0 \\ 0 & 0 & (-1)^n \end{bmatrix},$$

so that

$$\begin{aligned} e^{At} &= I + At + \frac{1}{2}A^2t^2 + \frac{1}{3!}A^3t^3 + \cdots \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix} t + \frac{1}{2!}\begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1 \end{bmatrix} t^2 + \frac{1}{3!}\begin{bmatrix} 8 & 0 & 0 \\ 0 & 27 & 0 \\ 0 & 0 & -1 \end{bmatrix} t^3 + \cdots \\ &= \begin{bmatrix} 1 + 2t + \frac{4t^2}{2} + \cdots & 0 & 0 \\ 0 & 1 + 3t + \frac{9t^2}{2} + \cdots & 0 \\ 0 & 0 & 1 - t + \frac{t^2}{2} + \cdots \end{bmatrix} \\ &= \begin{bmatrix} e^{2t} & 0 & 0 \\ 0 & e^{3t} & 0 \\ 0 & 0 & e^{-t} \end{bmatrix}. \end{aligned}$$

◄

Example 2. Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Compute e^{At}.

► **Solution.** In this case, $A^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{0}$ and $A^n = \mathbf{0}$ for all $n \geq 2$. Hence,

$$\begin{aligned} e^{At} &= I + At + \frac{1}{2}A^2t^2 + \frac{1}{3!}A^3t^3 + \cdots \\ &= I + At \\ &= \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Note that in this case, the individual entries of e^{At} are not exponential functions. ◄

Example 3. Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Compute e^{At}.

► **Solution.** The first few powers of A are $A^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, $A^3 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $A^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$, $A^5 = A$, $A^6 = A^2$, etc. That is, the powers repeat with period 4. Thus,

$$
\begin{aligned}
e^{At} &= I + At + \frac{1}{2}A^2t^2 + \frac{1}{3!}A^3t^3 + \cdots \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} -t^2 & 0 \\ 0 & -t^2 \end{bmatrix} + \frac{1}{3!}\begin{bmatrix} 0 & -t^3 \\ t^3 & 0 \end{bmatrix} + \frac{1}{4!}\begin{bmatrix} t^4 & 0 \\ 0 & t^4 \end{bmatrix} + \cdots \\
&= \begin{bmatrix} 1 - \frac{1}{2}t^2 + \frac{1}{4!}t^4 + \cdots & t - \frac{1}{3!}t^3 + \frac{1}{5!}t^5 + \cdots \\ -t + \frac{1}{3!}t^3 - \frac{1}{5!}t^5 + \cdots & 1 - \frac{1}{2}t^2 + \frac{1}{4!}t^4 + \cdots \end{bmatrix} \\
&= \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}.
\end{aligned}
$$

(cf. (3) and (4) of Sect. 7.1). In this example also, the individual entries of e^{At} are not themselves exponential functions. ◄

Do not let Examples 1–3 fool you. Unless A is very special, it is difficult to directly determine the entries of A^n and use this to compute e^{At}. In the following subsection, we will compute the Laplace transform of the matrix exponential function. The resulting inversion formula provides an effective method for explicitly computing e^{At}.

The Laplace Transform of the Matrix Exponential

Let A be an $n \times n$ matrix of scalars. As discussed above, each entry of e^{At} converges absolutely on $\mathbb{R}$. From a standard theorem in calculus, we have that e^{At} is differentiable and the derivative can be computed termwise. Thus,

$$
\begin{aligned}
\frac{d}{dt}e^{At} &= \frac{d}{dt}\left(I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \ldots\right) \\
&= A + \frac{A^2t}{1!} + \frac{A^3t^2}{2!} + \frac{A^4t^3}{3!} + \ldots \\
&= A\left(I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \ldots\right) \\
&= Ae^{At}.
\end{aligned}
$$

By factoring A out on the right-hand side in the second line, we also get

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{At} = \mathrm{e}^{At}A.$$

Appendix A.4 shows that each entry is of exponential type and thus has a Laplace transform. Now apply the derivative formula

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{At} = A\mathrm{e}^{At},$$

and the first derivative principle for the Laplace transform of matrix-valued functions (Theorem 7 of Sect. 9.2) applied to $v(t) = \mathrm{e}^{At}$, to get

$$A\mathcal{L}\left\{\mathrm{e}^{At}\right\} = \mathcal{L}\left\{A\mathrm{e}^{At}\right\} = \mathcal{L}\left\{\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{At}\right\} = s\mathcal{L}\left\{\mathrm{e}^{At}\right\} - I,$$

where we have used $v(0) = \mathrm{e}^{At}|_{t=0} = I$. Combining terms gives

$$(sI - A)\mathcal{L}\left\{\mathrm{e}^{At}\right\} = I$$

and thus,

$$\mathcal{L}\left\{\mathrm{e}^{At}\right\} = (sI - A)^{-1}.$$

This is the extension of (1) mentioned above. We summarize this discussion in the following theorem.

Theorem 4. *Let A be an $n \times n$ matrix. Then e^{At} is a well-defined matrix-valued function and*

$$\mathcal{L}\left\{\mathrm{e}^{At}\right\} = (sI - A)^{-1}. \tag{4}$$

The Laplace inversion formula is given by

$$\mathrm{e}^{At} = \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\}. \tag{5}$$

The matrix $(sI - A)^{-1}$ is called the ***resolvent matrix*** of A. It is a function of s, defined for all s for which the inverse exists. Let $c_A(s) = \det(sI - A)$. It is not hard to see that $c_A(s)$ is a polynomial of degree n. We call $c_A(s)$ the ***characteristic polynomial of*** A. As a polynomial of degree n, it has at most n roots. The roots are called ***characteristic values*** or ***eigenvalues of*** A. Thus, if s is larger than the absolute value of all the eigenvalues of A then $sI - A$ is invertible and the resolvent matrix is defined. By the adjoint formula for matrix inversion, Corollary 11 of Sect. 8.4 each entry of $(sI - A)^{-1}$ is the quotient of a cofactor of $sI - A$ and the characteristic polynomial $c_A(s)$, hence, a proper rational function. Thus, e^{At} is a matrix of exponential polynomials. The Laplace inversion formula given in Theorem 4 now provides a method to compute explicitly the matrix exponential without appealing to the power series expansion given by (2). It will frequently involve partial fraction decompositions of each entry of the resolvent matrix. Consider the following examples.

Example 5. Let $A = \begin{bmatrix} 3 & 4 \\ -2 & -3 \end{bmatrix}$. Compute the resolvent matrix $(sI - A)^{-1}$ and use the Laplace inversion formula to compute e^{At}.

► **Solution.** The characteristic polynomial is

$$\begin{aligned} c_A(s) &= \det(sI - A) \\ &= \det \begin{bmatrix} s-3 & -4 \\ 2 & s+3 \end{bmatrix} \\ &= (s-3)(s+3) + 8 \\ &= s^2 - 1 = (s-1)(s+1). \end{aligned}$$

The adjoint formula for the inverse thus gives

$$\begin{aligned} (sI - A)^{-1} &= \frac{1}{(s-1)(s+1)} \begin{bmatrix} s+3 & 4 \\ -2 & s-3 \end{bmatrix} \\ &= \begin{bmatrix} \dfrac{s+3}{(s-1)(s+1)} & \dfrac{4}{(s-1)(s+1)} \\ \dfrac{-2}{(s-1)(s+1)} & \dfrac{s-3}{(s-1)(s+1)} \end{bmatrix} \\ &= \frac{1}{s-1} \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix} + \frac{1}{s+1} \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}, \end{aligned}$$

where the third line is obtained by computing partial fractions of each entry in the second line.

To compute the matrix exponential, we use the Laplace inversion formula from Theorem 4 to get

$$\begin{aligned} e^{At} &= \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\} \\ &= e^{t} \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix} + e^{-t} \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 2e^{t} - e^{-t} & 2e^{t} - 2e^{-t} \\ -e^{t} + e^{-t} & -e^{t} + 2e^{-t} \end{bmatrix}. \end{aligned}$$ ◄

As a second example, we reconsider Example 3.

Example 6. Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Compute the resolvent matrix $(sI - A)^{-1}$ and use the Laplace inversion formula to compute e^{At}.

► **Solution.** The characteristic polynomial is

$$\begin{aligned} c_A(s) &= \det(sI - A) \\ &= \det \begin{bmatrix} s & -1 \\ 1 & s \end{bmatrix} \\ &= s^2 + 1. \end{aligned}$$

The adjoint formula for the inverse thus gives

$$\begin{aligned} (sI - A)^{-1} &= \frac{1}{s^2+1} \begin{bmatrix} s & 1 \\ -1 & s \end{bmatrix} \\ &= \begin{bmatrix} \dfrac{s}{s^2+1} & \dfrac{1}{s^2+1} \\ \dfrac{-1}{s^2+1} & \dfrac{s}{s^2+1} \end{bmatrix}. \end{aligned}$$

Using the inversion formula, we get

$$\begin{aligned} e^{At} &= \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\} \\ &= \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}, \end{aligned}$$ ◄

Theorem 4 thus gives an effective method for computing the matrix exponential. There are many other techniques. In the next section, we discuss a useful alternative that circumvents the need to compute partial fraction decompositions.

Exercises

1–7. Use the power series definition of the matrix exponential to compute e^{At} for the given matrix A.

1. $A = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$
2. $A = \begin{bmatrix} 3 & -1 \\ 9 & -3 \end{bmatrix}$
3. $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
4. $A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$
5. $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
6. $A = \begin{bmatrix} 0 & 2 & 0 \\ 1 & 0 & -1 \\ 0 & 2 & 0 \end{bmatrix}$
7. $A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

8–13. For each matrix A given below

(i) Compute the resolvent matrix $(sI - A)^{-1}$.
(ii) Compute the matrix exponential $e^{At} = \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\}$.

8. $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$
9. $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$
10. $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
11. $A = \begin{bmatrix} 3 & 5 \\ -1 & -1 \end{bmatrix}$
12. $A = \begin{bmatrix} 4 & -10 \\ 1 & -2 \end{bmatrix}$
13. $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

14. Suppose $A = \begin{bmatrix} M & 0 \\ 0 & N \end{bmatrix}$ where M is an $r \times r$ matrix and N is an $s \times s$ matrix. Show that $e^{At} = \begin{bmatrix} e^{Mt} & 0 \\ 0 & e^{Nt} \end{bmatrix}$.

15–16. Use Exercise 14 to compute the matrix exponential for each A.

15. $A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

16. $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -2 & 2 \end{bmatrix}$

9.4 Fulmer's Method for Computing e^{At}

The matrix exponential is fundamental to much of what we do in this chapter. It is therefore useful to have efficient techniques for calculating it. Here we will present a small variation on a technique[1] due to Fulmer[2] for computing the matrix exponential, e^{At}. It is based on the knowledge of what *types* of functions are included in the individual entries of e^{At}. This knowledge is derived from our understanding of the Laplace transform formula

$$e^{At} = \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\}.$$

Assume that A is an $n \times n$ constant matrix. Let $c_A(s) = \det(sI - A)$ be the characteristic polynomial of A. The characteristic polynomial has degree n, and by the adjoint formula for matrix inversion, $c_A(s)$ is in the denominator of each term of the inverse of $sI - A$. Therefore, each entry in $(sI - A)^{-1}$ belongs to $\mathcal{R}_{c_A}$, and hence, each entry of the matrix exponential, $e^{At} = \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\}$, is in $\mathcal{E}_{c_A}$. Recall from Sects. 2.6 and 2.7 that if $q(s)$ is a polynomial, then $\mathcal{R}_q$ is the set of proper rational functions that can be written with denominator $q(s)$, and $\mathcal{E}_q$ is the set of all exponential polynomials $f(t)$ with $\mathcal{L}\{f(t)\} \in \mathcal{R}_q$. If $\mathcal{B}_{c_A} = \{\phi_1, \phi_2, \ldots, \phi_n\}$ is the standard basis of $\mathcal{E}_{c_A}$, then it follows that

$$e^{At} = M_1\phi_1 + \cdots + M_n\phi_n,$$

where M_i is an $n \times n$ matrix for each index $i = 1, \ldots, n$. Fulmer's method is a procedure to determine the coefficient matrices $M_1, \ldots, M_n$.

Before considering the general procedure and its justification, we illustrate Fulmer's method with a simple example. If

$$A = \begin{bmatrix} 3 & 5 \\ -1 & -1 \end{bmatrix},$$

then the characteristic polynomial is

$$c_A(s) = \det(sI - A) = s^2 - 2s + 2 = (s-1)^2 + 1$$

[1]There are many other techniques. For example, see the articles "Nineteen Dubious Ways to Compute the Exponential of a Matrix" in *Siam Review*, Vol 20, no. 4, pp 801-836, October 1978 and "Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later" in *Siam Review*, Vol 45, no. 1, pp 3-49, 2003.

[2]Edward P. Fulmer, Computation of the Matrix Exponential, *American Mathematical Monthly*, **82** (1975) 156–159.

and the standard basis is $\mathcal{B}_{c_A} = \{e^t \cos t,\ e^t \sin t\}$. It follows that

$$e^{At} = M_1 e^t \cos t + M_2 e^t \sin t. \tag{1}$$

Differentiate (1) to get

$$Ae^{At} = M_1(e^t \cos t - e^t \sin t) + M_2(e^t \sin t + e^t \cos t) \tag{2}$$

and evaluate (1) and (2) at $t = 0$ to get the system

$$\begin{aligned} I &= M_1 \\ A &= M_1 + M_2. \end{aligned} \tag{3}$$

It is immediate that $M_1 = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $M_2 = A - M_1 = A - I = \begin{bmatrix} 2 & 5 \\ -1 & -2 \end{bmatrix}$. Substituting these matrices into (1) gives

$$\begin{aligned} e^{At} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} e^t \cos t + \begin{bmatrix} 2 & 5 \\ -1 & -2 \end{bmatrix} e^t \sin t \\ &= \begin{bmatrix} e^t \sin t + 2e^t \cos t & 5e^t \sin t \\ -e^t \sin t & e^t \sin t - 2e^t \cos t \end{bmatrix}. \end{aligned}$$

Compare the results here to those obtained in Exercise 11 in Sect. 9.3.

The General Case. Let A be an $n \times n$ matrix and $c_A(s)$ its characteristic polynomial. Suppose $\mathcal{B}_{c_A} = \{\phi_1, \ldots, \phi_n\}$. Reasoning as we did above, there are matrices $M_1, \ldots, M_n$ so that

$$e^{At} = M_1\phi_1(t) + \ldots + M_n\phi_n(t). \tag{4}$$

We need to find these matrices. By taking $n - 1$ derivatives, we obtain a system of linear equations (with matrix coefficients)

$$\begin{aligned} e^{At} &= M_1\phi_1(t) + \cdots + M_n\phi_n(t) \\ Ae^{At} &= M_1\phi_1'(t) + \cdots + M_n\phi_n'(t) \\ &\ \vdots \\ A^{n-1}e^{At} &= M_1\phi_1^{(n-1)}(t) + \cdots + M_n\phi_n^{(n-1)}(t). \end{aligned}$$

Now we evaluate this system at $t = 0$ to obtain

$$\begin{aligned} I &= M_1\phi_1(0) + \cdots + M_n\phi_n(0) \\ A &= M_1\phi_1'(0) + \cdots + M_n\phi_n'(0) \\ &\vdots \\ A^{n-1} &= M_1\phi_1^{(n-1)}(0) + \cdots + M_n\phi_n^{(n-1)}(0). \end{aligned} \tag{5}$$

At this point, we want to argue that it is always possible to solve (5) by showing that the coefficient matrix is invertible. However, in the examples and exercises, it is usually most efficient to solve (5) by elimination of variables. Let

$$W = \begin{bmatrix} \phi_1(0) & \dots & \phi_n(0) \\ \vdots & \ddots & \vdots \\ \phi_1^{(n-1)}(0) & \dots & \phi_n^{(n-1)}(0) \end{bmatrix}.$$

Then W is the Wronskian of $\phi_1, \dots, \phi_n$ at $t = 0$. Since $\phi_1, \dots, \phi_n$ are solutions to the linear homogeneous constant coefficient differential equation $c_A(\boldsymbol{D})(y) = 0$ (by Theorem 2 of Sect. 4.2) and since they are linearly independent ($\mathcal{B}_{c_A}$ is a *basis* of $\mathcal{E}_{c_A}$), Abel's formula, Theorem 6 of Sect. 4.2, applies to show the determinant of W is nonzero so W is invertible. The above system of equations can now be written:

$$\begin{bmatrix} I \\ A \\ \vdots \\ A^{n-1} \end{bmatrix} = W \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_n \end{bmatrix}.$$

Therefore,

$$\begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_n \end{bmatrix} = W^{-1} \begin{bmatrix} I \\ A \\ \vdots \\ A^{n-1} \end{bmatrix}.$$

Having solved for $M_1, \dots, M_n$, we obtain e^{At}.

Remark 1. Note that this last equation implies that each matrix M_i is a polynomial in the matrix A since W^{-1} is a constant matrix. Specifically, $M_i = p_i(A)$ where

$$p_i(s) = \mathrm{Row}_i(W^{-1}) \begin{bmatrix} 1 \\ s \\ \vdots \\ s^{n-1} \end{bmatrix}.$$

The following algorithm outlines Fulmer's method.

Algorithm 2 (Fulmer's Method).

Computation of e^{At} Where A Is a Given $n \times n$ Constant Matrix.

1. Compute the characteristic polynomial $c_A(s) = \det(sI - A)$.
2. Determine the standard basis $\mathcal{B}_{c_A} = \{\phi_1, \ldots, \phi_n\}$ of $\mathcal{E}_{c_A}$.
3. We then have
$$e^{At} = M_1\phi_1(t) + \cdots + M_n\phi_n(t) \tag{6}$$
where M_i $i = 1, \ldots, n$ are $n \times n$ matrices.
4. Take the derivative of (6) $n - 1$ times and evaluate each resulting equation at $t = 0$ to get a system of matrix equations.
5. Solve the matrix equations for $M_1, \ldots, M_n$.

Example 3. Find the matrix exponential, e^{At}, if $A = \begin{bmatrix} 2 & 1 \\ -4 & 6 \end{bmatrix}$.

► **Solution.** The characteristic polynomial is $c_A(s) = (s-2)(s-6) + 4 = s^2 - 8s + 16 = (s-4)^2$. Hence, $\mathcal{B}_{c_A} = \left\{e^{4t}, te^{4t}\right\}$ and it follows that

$$e^{At} = M_1e^{4t} + M_2te^{4t}.$$

Differentiating and evaluating at $t = 0$ gives

$$\begin{aligned} I &= M_1 \\ A &= 4M_1 + M_2. \end{aligned}$$

It follows that

$$M_1 = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } M_2 = A - 4I = \begin{bmatrix} -2 & 1 \\ -4 & 2 \end{bmatrix}.$$

Thus,

$$\begin{aligned} e^{At} &= M_1e^{4t} + M_2te^{4t} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} e^{4t} + \begin{bmatrix} -2 & 1 \\ -4 & 2 \end{bmatrix} te^{4t} \\ &= \begin{bmatrix} e^{4t} - 2te^{4t} & te^{4t} \\ -4te^{4t} & e^{4t} + 2te^{4t} \end{bmatrix}. \end{aligned}$$ ◄

As a final example, consider the following 3×3 matrix.

Example 4. Find the matrix exponential, e^{At}, if $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

► **Solution.** The characteristic polynomial is

$$\begin{aligned} c_A(s) &= \det \begin{bmatrix} s & 0 & -1 \\ 0 & s-2 & 0 \\ -1 & 0 & s \end{bmatrix} \\ &= (s-2)\det \begin{bmatrix} s & -1 \\ -1 & s \end{bmatrix} \\ &= (s-2)(s^2-1) = (s-2)(s-1)(s+1). \end{aligned}$$

It follows that $\mathcal{B}_{c_A} = \left\{e^{2t}, e^{t}, e^{-t}\right\}$ and

$$e^{At} = M_1 e^{2t} + M_2 e^{t} + M_3 e^{-t}.$$

Differentiating twice gives

$$\begin{aligned} Ae^{At} &= 2M_1 e^{2t} + M_2 e^{t} - M_3 e^{-t} \\ A^2 e^{At} &= 4M_1 e^{2t} + M_2 e^{t} + M_3 e^{-t} \end{aligned}$$

and evaluating at $t = 0$ gives

$$\begin{aligned} I &= M_1 + M_2 + M_3 \\ A &= 2M_1 + M_2 - M_3 \\ A^2 &= 4M_1 + M_2 + M_3. \end{aligned}$$

It is an easy exercise to solve for M_1, M_2, and M_3. We get

$$\begin{aligned} M_1 &= \frac{A^2 - I}{3} = \frac{(A-I)(A+I)}{3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ M_2 &= -\frac{A^2 - A - 2I}{2} = -\frac{(A-2I)(A+I)}{2} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}, \\ M_3 &= \frac{A^2 - 3A + 2I}{6} = \frac{(A-2I)(A-I)}{6} = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}. \end{aligned}$$

It follows now that

$$\begin{aligned} e^{At} &= e^{2t} M_1 + e^{t} M_2 + e^{-t} M_3 \\ &= \begin{bmatrix} \cosh t & 0 & \sinh t \\ 0 & e^{2t} & 0 \\ \sinh t & 0 & \cosh t \end{bmatrix}. \end{aligned}$$

◄

Exercises

1–19. Use Fulmer's method to compute the matrix exponential e^{At}.

1. $A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$

2. $A = \begin{bmatrix} -1 & 2 \\ -3 & 4 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & 1 \\ -4 & -2 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$

5. $A = \begin{bmatrix} 4 & -10 \\ 1 & -2 \end{bmatrix}$

6. $A = \begin{bmatrix} 4 & -1 \\ 1 & 2 \end{bmatrix}$

7. $A = \begin{bmatrix} -9 & 11 \\ -7 & 9 \end{bmatrix}$

8. $A = \begin{bmatrix} -5 & -8 \\ 4 & 3 \end{bmatrix}$

9. $A = \begin{bmatrix} 26 & 39 \\ -15 & -22 \end{bmatrix}$

10. $A = \begin{bmatrix} 3 & -3 \\ 2 & -2 \end{bmatrix}$

11. $A = \begin{bmatrix} -3 & 1 \\ -1 & -1 \end{bmatrix}$

12. $A = \begin{bmatrix} 6 & -4 \\ 2 & 0 \end{bmatrix}$

13. $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 2 & 0 & -2 \end{bmatrix}$, where $c_A(s) = s(s-1)(s+1)$

14. $A = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 1 & 1 \\ 2 & -1 & 0 \end{bmatrix}$, where $c_A(s) = (s-1)^2(s-2)$

15. $A = \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 1 & 1 & -1 \\ 0 & \frac{1}{2} & 1 \end{bmatrix}$, where $c_A(s) = (s-1)(s^2-2s+2)$

16. $A = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$, where $c_A(s) = (s-2)^3$

17. $A = \begin{bmatrix} -1 & -2 & 1 \\ 4 & 0 & -2 \\ -2 & -2 & 2 \end{bmatrix}$, where $c_A(s) = (s-1)(s^2+4)$

18. $A = \begin{bmatrix} 4 & 1 & -2 & 0 \\ 0 & 4 & 0 & -2 \\ 4 & 0 & -2 & 1 \\ 0 & 4 & 0 & -2 \end{bmatrix}$, where $c_A(s) = s^2(s-2)^2$

19. $A = \begin{bmatrix} -1 & 1 & 1 & 0 \\ -1 & -1 & 0 & 1 \\ -1 & 0 & 1 & 1 \\ 0 & -1 & -1 & 1 \end{bmatrix}$, where $c_A(s) = (s^2+1)^2$

20–22. Suppose A is a 2×2 real matrix with characteristic polynomial $c_A(s) = \det(sI - A) = s^2 + bs + c$. In these exercises, you are asked to derive a general formula for the matrix exponential e^{At}. We distinguish three cases.

20. **Distinct Real Roots:** Suppose $c_A(s) = (s - r_1)(s - r_2)$ with r_1 and r_2 distinct real numbers. Show that

$$e^{At} = \frac{A - r_2 I}{r_1 - r_2} e^{r_1 t} + \frac{A - r_1 I}{r_2 - r_1} e^{r_2 t}.$$

21. **Repeated Root:** Suppose $c_A(s) = (s - r)^2$. Show that

$$e^{At} = (I + (A - rI)t)\, e^{rt}.$$

22. **Complex Roots:** Suppose $c_A(s) = (s - \alpha)^2 + \beta^2$ where $\beta \neq 0$. Show that

$$e^{At} = I e^{\alpha t} \cos \beta t + \frac{(A - \alpha I)}{\beta} e^{\alpha t} \sin \beta t.$$

9.5 Constant Coefficient Linear Systems

In this section, we turn our attention to solving a first order constant coefficient linear differential system

$$\boldsymbol{y}' = A\boldsymbol{y} + \boldsymbol{f}, \qquad \boldsymbol{y}(t_0) = \boldsymbol{y}_0. \tag{1}$$

The solution method parallels that of Sect. 1.4. Specifically, the matrix exponential e^{-At} serves as an integrating factor to simplify the equivalent system

$$\boldsymbol{y}' - A\boldsymbol{y} = \boldsymbol{f}. \tag{2}$$

The result is the existence and uniqueness theorem for such systems. As a corollary, we obtain the existence and uniqueness theorems for ordinary constant coefficient linear differential equations, as stated in Sects. 3.1 and 4.1.

We begin with a lemma that lists the necessary properties of the matrix exponential to implement the solution method.

Lemma 1. *Let A be an $n \times n$ matrix. The following statements then hold:*

1. $\mathrm{e}^{At}\big|_{t=0} = I$.
2. $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{At} = A\mathrm{e}^{At} = \mathrm{e}^{At}A$ *for all* $t \in \mathbb{R}$.
3. $\mathrm{e}^{A(t+a)} = \mathrm{e}^{At}\mathrm{e}^{Aa} = \mathrm{e}^{Aa}\mathrm{e}^{At}$ *for all* $t, a \in \mathbb{R}$.
4. e^{At} *is an invertible matrix with inverse* $\left(\mathrm{e}^{At}\right)^{-1} = \mathrm{e}^{-At}$ *for all* $t \in \mathbb{R}$.

Proof. Items 1. and 2. were proved in Sect. 9.3. Fix $a \in \mathbb{R}$ and let

$$\Phi(t) = \mathrm{e}^{-At}\mathrm{e}^{A(t+a)}.$$

Then

$$\begin{aligned}\Phi'(t) &= -A\mathrm{e}^{-At}\mathrm{e}^{A(t+a)} + \mathrm{e}^{-At}A\mathrm{e}^{A(t+a)} \\ &= -A\mathrm{e}^{-At}\mathrm{e}^{A(t+a)} + A\mathrm{e}^{-At}\mathrm{e}^{A(t+a)} = 0,\end{aligned}$$

which follows from the product rule and part 2. It follows that Φ is a constant matrix and since $\Phi(0) = \mathrm{e}^{-A0}\mathrm{e}^{Aa} = \mathrm{e}^{Aa}$ by part 1, we have

$$\mathrm{e}^{-At}\mathrm{e}^{A(t+a)} = \mathrm{e}^{Aa}, \tag{3}$$

for all $t, a \in \mathbb{R}$. Now let $a = 0$ then $\mathrm{e}^{-At}\mathrm{e}^{At} = I$. From this, it follows that e^{At} is invertible and $\left(\mathrm{e}^{At}\right)^{-1} = \mathrm{e}^{-At}$. This proves item 4. Further, from (3), we have

$$\mathrm{e}^{A(t+a)} = (\mathrm{e}^{-At})^{-1}\mathrm{e}^{Aa} = \mathrm{e}^{At}\mathrm{e}^{Aa}.$$

This proves item 3. □

To solve (1), multiply (2) by e^{-At} to get

$$e^{-At}\boldsymbol{y}'(t) - e^{-At}A\boldsymbol{y}(t) = e^{-At}\boldsymbol{f}(t). \tag{4}$$

By the product rule and Lemma 1, part 2, we have

$$(e^{-At}\boldsymbol{y})'(t) = -e^{-At}A\boldsymbol{y}(t) + e^{-At}\boldsymbol{y}'(t).$$

We can thus rewrite (4) as

$$(e^{-At}\boldsymbol{y})'(t) = e^{-At}\boldsymbol{f}(t).$$

Now change the variable from t to u and integrate both sides from t_0 to t to get

$$e^{-At}\boldsymbol{y}(t) - e^{-At_0}\boldsymbol{y}(t_0) = \int_{t_0}^{t} (e^{-Au}\boldsymbol{y})'(u)\,du = \int_{t_0}^{t} e^{-Au}\boldsymbol{f}(u)\,du.$$

Now add $e^{-At_0}\boldsymbol{y}(t_0)$ to both sides and multiply by the inverse of e^{-At}, which is e^{At} by Lemma 1, part 4, to get

$$\boldsymbol{y}(t) = e^{At}e^{-At_0}\boldsymbol{y}(t_0) + e^{At}\int_{t_0}^{t} e^{-Au}\boldsymbol{f}(u)\,du.$$

Now use Lemma 1, part 3, to simplify. We get

$$\boldsymbol{y}(t) = e^{A(t-t_0)}\boldsymbol{y}_0 + \int_{t_0}^{t} e^{A(t-u)}\,\boldsymbol{f}(u)\,du. \tag{5}$$

This argument shows that if there is a solution, it must take this form. However, it is a straightforward calculation to verify that (5) is a solution to (1). We thereby obtain

Theorem 2 (Existence and Uniqueness Theorem). *Let A be an $n \times n$ constant matrix and $\boldsymbol{f}(t)$ an $\mathbb{R}^n$-valued continuous function defined on an interval I. Let $t_0 \in I$ and $\boldsymbol{y}_0 \in \mathbb{R}^n$. Then the unique solution to the initial value problem*

$$\boldsymbol{y}'(t) = A\boldsymbol{y}(t) + \boldsymbol{f}(t), \quad \boldsymbol{y}(t_0) = \boldsymbol{y}_0, \tag{6}$$

is the function $\boldsymbol{y}(t)$ defined for $t \in I$ by

Solution to a First Order Differential System

$$\boldsymbol{y}(t) = e^{A(t-t_0)}\boldsymbol{y}_0 + \int_{t_0}^{t} e^{A(t-u)}\boldsymbol{f}(u)\,du. \tag{7}$$

Let us break up this general solution into its two important parts. First, when $\boldsymbol{f} = 0$, (6) reduces to $\boldsymbol{y}'(t) = A\boldsymbol{y}(t)$, the ***associated homogeneous equation***. Its solution is the ***homogeneous solution*** given simply by

$$\boldsymbol{y}_{\rm h} = \mathrm{e}^{A(t-t_0)}\boldsymbol{y}_0.$$

Let $\boldsymbol{y}_0 = \boldsymbol{e}_i$, the column vector with 1 in the ith position and 0's elsewhere. Define $\boldsymbol{y}_i = \mathrm{e}^{A(t-t_0)}\boldsymbol{e}_i$. Then $\boldsymbol{y}_i$ is the ith column of $\mathrm{e}^{A(t-t_0)}$. This means that each column of $\mathrm{e}^{A(t-t_0)}$ is a solution to the associated homogeneous equation. Furthermore, if

$$\boldsymbol{y}_0 = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

then

$$\boldsymbol{y}_{\rm h} = \mathrm{e}^{A(t-t_0)}\boldsymbol{y}_0 = a_1\boldsymbol{y}_1 + a_2\boldsymbol{y}_2 + \cdots + a_n\boldsymbol{y}_n$$

is a linear combination of the columns of e^{At} and all homogeneous solutions are of this form.

The other piece of the general solution is the ***particular solution***:

$$\boldsymbol{y}_{\rm p}(t) = \int_{t_0}^{t} \mathrm{e}^{A(t-u)}\boldsymbol{f}(u)\,\mathrm{d}u. \tag{8}$$

We then get the familiar formula

$$\boldsymbol{y} = \boldsymbol{y}_{\rm h} + \boldsymbol{y}_{\rm p}.$$

The ***general solution*** to $\boldsymbol{y}' = A\boldsymbol{y} + \boldsymbol{f}$ is thus obtained by adding all possible homogeneous solutions to one fixed particular solution.

When $t_0 = 0$, the particular solution $\boldsymbol{y}_{\rm p}$ of (8) becomes the convolution product of e^{At} and $\boldsymbol{f}(t)$. We record this important special case as a corollary.

Corollary 3. *Let A be an $n \times n$ constant matrix and $\boldsymbol{f}(t)$ an $\mathbb{R}^n$-valued continuous function defined on an interval I containing the origin. Let $\boldsymbol{y}_0 \in \mathbb{R}^n$. The unique solution to*

$$\boldsymbol{y}'(t) = A\boldsymbol{y}(t) + \boldsymbol{f}(t), \quad \boldsymbol{y}(0) = y_0,$$

is the function defined for $t \in I$ by

$$\boldsymbol{y}(t) = \mathrm{e}^{At}\boldsymbol{y}_0 + \mathrm{e}^{At} * \boldsymbol{f}(t). \tag{9}$$

When each entry of $\boldsymbol{f}$ is of exponential type, then the convolution theorem, Theorem 1 of Sect. 2.8, can be used to compute $y_{\rm p} = \mathrm{e}^{At} * \boldsymbol{f}(t)$. Let us consider a few examples.

Example 4. Solve the following linear system of differential equations:

$$\begin{aligned} y_1' &= -y_1 + 2y_2 + e^t \\ y_2' &= -3y_1 + 4y_2 - 2e^t, \end{aligned} \tag{10}$$

with initial conditions $y_1(0) = 1$ and $y_2(0) = 1$.

► **Solution.** We begin by writing the given system in matrix form:

$$\boldsymbol{y}' = A\boldsymbol{y} + \boldsymbol{f}, \quad \boldsymbol{y}(0) = \boldsymbol{y}_0 \tag{11}$$

where

$$A = \begin{bmatrix} -1 & 2 \\ -3 & 4 \end{bmatrix}, \quad \boldsymbol{f}(t) = \begin{bmatrix} e^t \\ -2e^t \end{bmatrix} = e^t \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{y}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The characteristic polynomial is

$$c_A(s) = \det(sI - A) = \det \begin{bmatrix} s+1 & -2 \\ 3 & s-4 \end{bmatrix} = s^2 - 3s + 2 = (s-1)(s-2). \tag{12}$$

Therefore,

$$\begin{aligned} (sI - A)^{-1} &= \begin{bmatrix} \dfrac{s-4}{(s-1)(s-2)} & \dfrac{2}{(s-1)(s-2)} \\ \dfrac{-3}{(s-1)(s-2)} & \dfrac{s+1}{(s-1)(s-2)} \end{bmatrix} \\ &= \frac{1}{s-1}\begin{bmatrix} 3 & -2 \\ 3 & -2 \end{bmatrix} + \frac{1}{s-2}\begin{bmatrix} -2 & 2 \\ -3 & 3 \end{bmatrix}. \end{aligned}$$

It now follows that

$$e^{At} = \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\} = e^t \begin{bmatrix} 3 & -2 \\ 3 & -2 \end{bmatrix} + e^{2t}\begin{bmatrix} -2 & 2 \\ -3 & 3 \end{bmatrix}.$$

By Corollary 3, the homogeneous part of the solution is given by

$$\begin{aligned} \boldsymbol{y}_h &= e^{At}\boldsymbol{y}_0 \\ &= \left(e^t \begin{bmatrix} 3 & -2 \\ 3 & -2 \end{bmatrix} + e^{2t}\begin{bmatrix} -2 & 2 \\ -3 & 3 \end{bmatrix}\right)\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \end{aligned}$$

To compute $\boldsymbol{y}_{\mathrm{p}}$ we will use two simple convolution formulas: $e^t * e^t = te^t$ and $e^{2t} * e^t = e^{2t} - e^t$. By Corollary 3, we have

$$\begin{aligned}
\boldsymbol{y}_{\mathrm{p}}(t) &= e^{At} * \boldsymbol{f}(t) \\
&= \left(e^t \begin{bmatrix} 3 & -2 \\ 3 & -2 \end{bmatrix} + e^{2t} \begin{bmatrix} -2 & 2 \\ -3 & 3 \end{bmatrix} \right) * e^t \begin{bmatrix} 1 \\ -2 \end{bmatrix} \\
&= e^t * e^t \begin{bmatrix} 3 & -2 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} + e^{2t} * e^t \begin{bmatrix} -2 & 2 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} \\
&= te^t \begin{bmatrix} 7 \\ 7 \end{bmatrix} + (e^{2t} - e^t) \begin{bmatrix} -6 \\ -9 \end{bmatrix} \\
&= \begin{bmatrix} 7te^t + 6e^t - 6e^{2t} \\ 7te^t + 9e^t - 9e^{2t} \end{bmatrix}.
\end{aligned}$$

Now, adding the homogeneous and particular solutions together leads to the solution:

$$\begin{aligned}
\boldsymbol{y}(t) = \boldsymbol{y}_{\mathrm{h}}(t) + \boldsymbol{y}_{\mathrm{p}}(t) &= e^{At}\boldsymbol{y}(0) + e^{At} * \boldsymbol{f}(t) \\
&= \begin{bmatrix} e^t \\ e^t \end{bmatrix} + \begin{bmatrix} 7te^t + 6e^t - 6e^{2t} \\ 7te^t + 9e^t - 9e^{2t} \end{bmatrix} \\
&= \begin{bmatrix} 7te^t + 7e^t - 6e^{2t} \\ 7te^t + 10e^t - 9e^{2t} \end{bmatrix}.
\end{aligned} \tag{13}$$

◀

Example 5. Find the general solution to the following system of differential equations:

$$\begin{aligned}
y_1' &= y_2 + t \\
y_2' &= -y_1 - t
\end{aligned}$$

► **Solution.** If

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \boldsymbol{f}(t) = \begin{bmatrix} t \\ -t \end{bmatrix},$$

the given system can be expressed as

$$\boldsymbol{y}' = A\boldsymbol{y} + \boldsymbol{f}.$$

By Example 6 of Sect. 9.3, we have $e^{At} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$. If $\boldsymbol{y}_0 = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ then

$$\boldsymbol{y}_{\mathrm{h}} = e^{At}\boldsymbol{y}_0 = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = a_1 \begin{bmatrix} \cos t \\ -\sin t \end{bmatrix} + a_2 \begin{bmatrix} \sin t \\ \cos t \end{bmatrix}.$$

Further,

$$\boldsymbol{y}_{\mathrm{p}} = \mathrm{e}^{At} * f(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} * \begin{bmatrix} t \\ -t \end{bmatrix} = \begin{bmatrix} (\cos t) * t - (\sin t) * t \\ -(\sin t) * t - (\cos t) * t \end{bmatrix}.$$

Table 2.11 gives formulas for the convolutions $(\cos t) * t$ and $(\sin t) * t$. However, we will use the convolution principle to make these computations. First,

$$\mathcal{L}\{(\cos t) * t\} = \frac{s}{(s^2+1)}\frac{1}{s^2} = \frac{1}{(s^2+1)s} = \frac{1}{s} - \frac{s}{s^2+1},$$

$$\text{and} \quad \mathcal{L}\{(\sin t) * t\} = \frac{1}{(s^2+1)s^2} = \frac{1}{s^2} - \frac{1}{s^2+1},$$

so that $(\cos t) * t = 1 - \cos t$ and $(\sin t) * t = t - \sin t$. Therefore,

$$\boldsymbol{y}_{\mathrm{p}} = \begin{bmatrix} 1 - \cos t - (t - \sin t) \\ -(t - \sin t) - (1 - \cos t) \end{bmatrix}.$$

By Corollary 3, the general solution is thus

$$\boldsymbol{y}(t) = a_1 \begin{bmatrix} \cos t \\ -\sin t \end{bmatrix} + a_2 \begin{bmatrix} \sin t \\ \cos t \end{bmatrix} + \begin{bmatrix} 1 - t - \cos t + \sin t \\ -1 - t + \sin t + \cos t \end{bmatrix},$$

which we can rewrite more succinctly as

$$\begin{aligned} \boldsymbol{y}(t) &= (a_1 - 1)\begin{bmatrix} \cos t \\ -\sin t \end{bmatrix} + (a_2 + 1)\begin{bmatrix} \sin t \\ \cos t \end{bmatrix} + \begin{bmatrix} 1 - t \\ -1 - t \end{bmatrix} \\ &= \alpha_1 \begin{bmatrix} \cos t \\ -\sin t \end{bmatrix} + \alpha_2 \begin{bmatrix} \sin t \\ \cos t \end{bmatrix} + \begin{bmatrix} 1 - t \\ -1 - t \end{bmatrix} \end{aligned}$$

after relabeling the coefficients. ◀

Example 6. Solve the following system of equations:

$$\begin{aligned} y_1' &= 2y_1 + y_2 + 1/t \qquad & y_1(1) &= 2 \\ y_2' &= -4y_1 - 2y_2 + 2/t \qquad & y_2(1) &= 1 \end{aligned}$$

on the interval $(0, \infty)$.

► **Solution.** We can write the given system as $\boldsymbol{y}' = A\boldsymbol{y} + \boldsymbol{f}$, where

$$A = \begin{bmatrix} 2 & 1 \\ -4 & -2 \end{bmatrix}, \quad \boldsymbol{f}(t) = \begin{bmatrix} 1/t \\ 2/t \end{bmatrix}, \quad \text{and} \quad \boldsymbol{y}(1) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Observe that $\boldsymbol{f}$ is continuous on $(0, \infty)$. The characteristic polynomial is

$$c_A(s) = \det \begin{bmatrix} s-2 & -1 \\ 4 & s+2 \end{bmatrix} = s^2 \tag{14}$$

and $\mathcal{B}_{c_A} = \{1, t\}$. Therefore, $e^{At} = M_1 + M_2 t$ and its derivative is $Ae^{At} = M_2$. Evaluating at $t = 0$ gives $M_1 = I$ and $M_2 = A$. Fulmer's method now gives

$$e^{At} = I + At = \begin{bmatrix} 1+2t & t \\ -4t & 1-2t \end{bmatrix}.$$

It follows that

$$\boldsymbol{y}_{\mathrm{h}}(t) = e^{A(t-1)}\boldsymbol{y}(1) = \begin{bmatrix} 1+2(t-1) & t-1 \\ -4(t-1) & 1-2(t-1) \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2+5(t-1) \\ 1-10(t-1) \end{bmatrix}.$$

The following calculation gives the particular solution:

$$\begin{aligned}
\boldsymbol{y}_{\mathrm{p}}(t) &= \int_1^t e^{A(t-u)} \boldsymbol{f}(u)\, du \\
&= e^{At} \int_1^t e^{-Au} \boldsymbol{f}(u)\, du \\
&= e^{At} \int_1^t \begin{bmatrix} 1-2u & -u \\ 4u & 1+2u \end{bmatrix} \begin{bmatrix} 1/u \\ 2/u \end{bmatrix} du \\
&= e^{At} \int_1^t \begin{bmatrix} \frac{1}{u} - 4 \\ \frac{2}{u} + 8 \end{bmatrix} du \\
&= \begin{bmatrix} 1+2t & t \\ -4t & 1-2t \end{bmatrix} \begin{bmatrix} \ln t - 4(t-1) \\ 2\ln t + 8(t-1) \end{bmatrix} \\
&= \begin{bmatrix} (1+4t)\ln t - 4t + 4 \\ (2-8t)\ln t + 8t - 8 \end{bmatrix}.
\end{aligned}$$

Where in line 2 we used Lemma 1 to write $e^{A(t-u)} = e^{At}e^{-Au}$. We now get

$$\begin{aligned}
\boldsymbol{y}(t) &= \boldsymbol{y}_{\mathrm{h}}(t) + \boldsymbol{y}_{\mathrm{p}}(t) \\
&= \begin{bmatrix} 2+5(t-1) \\ 1-10(t-1) \end{bmatrix} + \begin{bmatrix} (1+4t)\ln t - 4t + 4 \\ (2-8t)\ln t + 8t - 8 \end{bmatrix} \\
&= \begin{bmatrix} 1+t+(1+4t)\ln t \\ 3-2t+(2-8t)\ln t \end{bmatrix},
\end{aligned}$$

valid for all $t > 0$. ◀

Example 7. Solve the mixing problem introduced at the beginning of this chapter in Example 1 of Sect. 9.1. Namely,

$$\begin{aligned} y_1'(t) &= \frac{-4}{10} y_1(t) + \frac{1}{10} y_2(t) + 3 \\ y_2'(t) &= \frac{4}{10} y_1(t) - \frac{4}{10} y_2(t), \end{aligned}$$

with initial condition $y_1(0) = 2$ and $y_2(0) = 0$. Also, determine the amount of salt in each tank at time $t = 10$.

► **Solution.** In matrix form, this system can be written as

$$\boldsymbol{y}'(t) = A\boldsymbol{y}(t) + \boldsymbol{f}(t),$$

where

$$A = \begin{bmatrix} -4/10 & 1/10 \\ 4/10 & -4/10 \end{bmatrix}, \quad \boldsymbol{f}(t) = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{y}_0 = \boldsymbol{y}(0) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

We let the reader verify that

$$\mathrm{e}^{At} = \mathrm{e}^{-2t/10} \begin{bmatrix} 1/2 & 1/4 \\ 1 & 1/2 \end{bmatrix} + \mathrm{e}^{-6t/10} \begin{bmatrix} 1/2 & -1/4 \\ -1 & 1/2 \end{bmatrix}.$$

The homogeneous solution is

$$\begin{aligned} \boldsymbol{y}_{\mathrm{h}}(t) = \mathrm{e}^{At}\boldsymbol{y}_0 &= \left(\mathrm{e}^{-2t/10} \begin{bmatrix} 1/2 & 1/4 \\ 1 & 1/2 \end{bmatrix} + \mathrm{e}^{-6t/10} \begin{bmatrix} 1/2 & -1/4 \\ -1 & 1/2 \end{bmatrix} \right) \begin{bmatrix} 2 \\ 0 \end{bmatrix} \\ &= \mathrm{e}^{-2t/10} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \mathrm{e}^{-6t/10} \begin{bmatrix} 1 \\ -2 \end{bmatrix}. \end{aligned}$$

We use the fact that $\mathrm{e}^{-2t/10} * 1 = 5 - 5\mathrm{e}^{-2t/10}$ and $\mathrm{e}^{-6t/10} * 1 = \frac{1}{3}(5 - 5\mathrm{e}^{-6t/10})$ to get

$$\begin{aligned} \boldsymbol{y}_{\mathrm{p}} &= \mathrm{e}^{At} * \begin{bmatrix} 3 \\ 0 \end{bmatrix} = 3\mathrm{e}^{At} * \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= 3\left(\mathrm{e}^{-2t/10} * 1\right) \begin{bmatrix} 1/2 & 1/4 \\ 1 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\left(\mathrm{e}^{-6t/10} * 1\right) \begin{bmatrix} 1/2 & -1/4 \\ -1 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{aligned}$$

$$= 3\left(5 - 5\mathrm{e}^{-2t/10}\right)\begin{bmatrix}1/2\\1\end{bmatrix} + \left(5 - 5\mathrm{e}^{-6t/10}\right)\begin{bmatrix}1/2\\-1\end{bmatrix}$$

$$= \begin{bmatrix}10\\10\end{bmatrix} - \mathrm{e}^{-2t/10}\begin{bmatrix}15/2\\15\end{bmatrix} - \mathrm{e}^{-6t/10}\begin{bmatrix}5/2\\-5\end{bmatrix}.$$

We now obtain the solution:

$$\begin{aligned}\boldsymbol{y}(t) &= \boldsymbol{y}_{\mathrm{h}}(t) + \boldsymbol{y}_{\mathrm{p}}(t)\\ &= \begin{bmatrix}10\\10\end{bmatrix} + \mathrm{e}^{-2t/10}\begin{bmatrix}-13/2\\-13\end{bmatrix} + \mathrm{e}^{-6t/10}\begin{bmatrix}-3/2\\3\end{bmatrix}.\end{aligned}$$

At time $t = 10$, we have

$$\boldsymbol{y}(10) = \begin{bmatrix}10 - (13/2)\mathrm{e}^{-2} - (3/2)\mathrm{e}^{-6}\\ 10 - 13\mathrm{e}^{-2} + 3\mathrm{e}^{-6}\end{bmatrix} = \begin{bmatrix}9.117\\8.248\end{bmatrix}.$$

At $t = 10$ minutes, Tank 1 contains 9.117 pounds of salt and Tank 2 contains 8.248 pounds of salt. ◀

We now summarize in the following algorithm the procedure for computing the solution to a constant coefficient first order system.

Algorithm 8. Given a constant coefficient first order system,

$$\boldsymbol{y}' = A\boldsymbol{y} + \boldsymbol{f}, \quad \boldsymbol{y}(t_0) = \boldsymbol{y}_0$$

we proceed as follows to determine the solution set.

Solution Method for a Constant Coefficient First Order System

1. Determine e^{At}: This may be done by the inverse Laplace transform formula $\mathrm{e}^{At} = \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\}$ or by Fulmer's method.
2. Determine the homogeneous part $\boldsymbol{y}_{\mathrm{h}}(t) = \mathrm{e}^{A(t-t_0)}\boldsymbol{y}(t_0)$.
3. Determine the particular solution $\boldsymbol{y}_{\mathrm{p}}(t) = \int_{t_0}^{t} \mathrm{e}^{A(t-u)}\boldsymbol{f}(u)\,\mathrm{d}u$. It is sometimes useful to use $\mathrm{e}^{A(t-u)} = \mathrm{e}^{At}\mathrm{e}^{-Au}$.
4. The general solution is $\boldsymbol{y}_{\mathrm{g}} = \boldsymbol{y}_{\mathrm{h}} + \boldsymbol{y}_{\mathrm{p}}$.

Eigenvectors and Eigenvalues

When the initial condition $\boldsymbol{y}_0 = \boldsymbol{v}$ is an eigenvector for A, then the solution to $\boldsymbol{y}' = A\boldsymbol{y}$, $\boldsymbol{y}(0) = \boldsymbol{v}$, takes on a very simple form.

Lemma 9. *Suppose A is an $n \times n$ matrix and v is an eigenvector with eigenvalue λ. Then the solution to*

$$\boldsymbol{y}' = A\boldsymbol{y}, \quad \boldsymbol{y}(0) = \boldsymbol{v}$$

is

$$\boldsymbol{y} = e^{\lambda t}\boldsymbol{v}.$$

In other words,

$$e^{At}\boldsymbol{v} = e^{\lambda t}\boldsymbol{v}.$$

Proof. Let $\boldsymbol{y}(t) = e^{\lambda t}\boldsymbol{v}$. Then $\boldsymbol{y}'(t) = \lambda e^{\lambda t}\boldsymbol{v}$ and $A\boldsymbol{y}(t) = e^{\lambda t}A\boldsymbol{v} = \lambda e^{\lambda t}\boldsymbol{v}$. Therefore, $\boldsymbol{y}'(t) = A\boldsymbol{y}(t)$. By the uniqueness and existence theorem, we have

$$e^{At}\boldsymbol{v} = e^{\lambda t}\boldsymbol{v}.$$ □

Example 10. Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. Find the solution to

$$\boldsymbol{y}'(t) = A\boldsymbol{y}(t), \quad \boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

► **Solution.** We observe that $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of A with eigenvalue 5:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Thus,

$$\boldsymbol{y}(t) = e^{5t}\begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

is the unique solution to $\boldsymbol{y}' = A\boldsymbol{y}$, $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. ◄

More generally, we have

Theorem 11. *Suppose $\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k$, where $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ are eigenvectors of A with corresponding eigenvalues $\lambda_1, \ldots, \lambda_k$. Then the solution to*

$$\boldsymbol{y}' = A\boldsymbol{y}, \quad \boldsymbol{y}(0) = \boldsymbol{v}$$

is

$$\boldsymbol{y} = c_1 \mathrm{e}^{\lambda_1 t} \boldsymbol{v}_1 + \cdots + c_2 \mathrm{e}^{\lambda_k t} \boldsymbol{v}_k.$$

Proof. By Theorem 2, the solution is $\boldsymbol{y}(t) = \mathrm{e}^{At}\boldsymbol{v}$. By linearity and Lemma 9, we get

$$\begin{aligned}\boldsymbol{y}(t) = \mathrm{e}^{At}\boldsymbol{v} &= c_1 \mathrm{e}^{At}\boldsymbol{v}_1 + \cdots + c_k \mathrm{e}^{At}\boldsymbol{v}_k \\ &= c_1 \mathrm{e}^{\lambda_1 t}\boldsymbol{v}_1 + \cdots + c_k \mathrm{e}^{\lambda_k t}\boldsymbol{v}_k.\end{aligned}$$ □

Existence and Uniqueness Theorems

We conclude this section with the existence and uniqueness theorems referred to earlier in the text, namely, Theorem 10 of Sect. 3.1 and Theorem 5 of Sect. 4.1.

For convenience of expression, we will say that an $\mathbb{R}^n$-valued function $\boldsymbol{f}$ is an ***exponential polynomial*** if each component f_i of $\boldsymbol{f}$ is an exponential polynomial. Similarly, we say that an $\mathbb{R}^n$-valued function $\boldsymbol{f}$ is of ***exponential type*** if each component f_i of $\boldsymbol{f}$ is of exponential type.

Corollary 12. *Suppose $\boldsymbol{f}$ is an exponential polynomial. Then the solution to*

$$\boldsymbol{y}'(t) = A\boldsymbol{y}(t) + \boldsymbol{f}(t), \quad \boldsymbol{y}(t_0) = \boldsymbol{y}_0$$

is an exponential polynomial defined on $\mathbb{R}$.

Proof. The formula for the solution is given by (7) in Theorem 2. Each entry in e^{At} is an exponential polynomial. Therefore, each entry of $\mathrm{e}^{A(t-t_0)} = \mathrm{e}^{At}\mathrm{e}^{-At_0}$ is a linear combination of entries of e^{At}, hence an exponential polynomial. It follows that $\mathrm{e}^{A(t-t_0)}\boldsymbol{y}_0$ is an exponential polynomial. The function $u \to \mathrm{e}^{A(t-u)}\boldsymbol{f}(u)$ is a translation and product of exponential polynomials. Thus, by Exercises 34 and 35 of Sect. 2.7 it is in $\mathcal{E}$, and by Exercise 37 of Sect. 2.7 we have $\int_{t_0}^{t} \mathrm{e}^{A(t-u)}\boldsymbol{f}(u)\,\mathrm{d}u$ is in $\mathcal{E}$. Thus, each piece in (7) is an exponential polynomial so the solution $\boldsymbol{y}(t)$ is an exponential polynomial. □

Corollary 13. *Suppose $\boldsymbol{f}$ is of exponential type. Then the solution to*

$$\boldsymbol{y}'(t) = A\boldsymbol{y}(t) + \boldsymbol{f}(t), \quad \boldsymbol{y}(t_0) = y_0$$

is of exponential type.

Proof. By Proposition 1 of Sect. 2.2, Exercise 37 of Sect. 2.2, and Lemma 4 of Sect. 2.2, we find that sums, products, and integrals of functions that are of exponential type are again of exponential type. Reasoning as above, we obtain the result. □

Theorem 14 (The Existence and Uniqueness Theorem for Constant Coefficient Linear Differential Equations). *Suppose $f(t)$ is a continuous real-valued function on an interval I. Let $t_0 \in I$. Then there is a unique real-valued function y*

defined on I satisfying

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots a_1 y' + a_0 y = f(t), \tag{15}$$

with initial conditions $y(t_0) = y_0$, $y'(t_0) = y_1$, $\ldots y^{(n-1)} = y_{n-1}$. If $f(t)$ is of exponential type, so is the solution $y(t)$ and its derivatives $y^{(i)}(t)$, for $i = 1, \ldots, n-1$. Furthermore, if $f(t)$ is in $\mathcal{E}$, then $y(t)$ is also in $\mathcal{E}$.

Proof. We may assume $a_n = 1$ by dividing by a_n, if necessary. Let $y_1 = y$, $y_2 = y'$, $\ldots$, $y_n = y^{(n-1)}$, and let $\boldsymbol{y}$ be the column vector with entries $y_1, \ldots, y_n$. Then $y_1' = y' = y_2, \ldots, y_{n-1}' = y^{(n-1)} = y_n$, and

$$\begin{aligned} y_n' = y^{(n)} &= -a_0 y - a_1 y' - a_2 y'' - \cdots - a_{n-1} y^{(n-1)} + f \\ &= -a_0 y_1 - a_1 y_2 - a_2 y_3 - \cdots - a_{n-1} y_n + f. \end{aligned}$$

It is simple to check that y is a solution to (15) if and only if $\boldsymbol{y}$ is a solution to (1), where

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix}, \quad \boldsymbol{f} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f \end{bmatrix}, \quad \text{and} \quad \boldsymbol{y}_0 = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-2} \\ y_{n-1} \end{bmatrix}.$$

By Theorem 2, there is a unique solution $\boldsymbol{y}$. The first entry, y, in $\boldsymbol{y}$ is the unique solution to (15). If f is an exponential polynomial, then $\boldsymbol{f}$ is likewise, and Corollary 12 implies that $\boldsymbol{y}$ is an exponential polynomial. Hence, $y_1 = y$, $y_2 = y'$, $\ldots$, $y_n = y^{(n-1)}$ are all exponential polynomials. If f is of exponential type, then so is $\boldsymbol{f}$, and Corollary 13 implies $\boldsymbol{y}$ is of exponential type. This, in turn, implies y, $y', \ldots, y^{n-1}$ are each of exponential type. $\square$

Exercises

1–9. Solve the homogeneous systems $\boldsymbol{y}' = A\boldsymbol{y}$, $\boldsymbol{y}(0) = \boldsymbol{y}_0$, for the given A and $\boldsymbol{y}_0$.

1. $A = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$; $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$
2. $A = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$; $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$
3. $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$; $\boldsymbol{y}(0) = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$
4. $A = \begin{bmatrix} -1 & 2 \\ -2 & -1 \end{bmatrix}$; $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
5. $A = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}$; $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$
6. $A = \begin{bmatrix} 2 & -5 \\ 1 & -2 \end{bmatrix}$; $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$
7. $A = \begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix}$; $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
8. $A = \begin{bmatrix} -1 & 0 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$
9. $A = \begin{bmatrix} 0 & 4 & 0 \\ -1 & 0 & 0 \\ 1 & 4 & -1 \end{bmatrix}$; $\boldsymbol{y}(0) = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$

10–17. Use Corollary 3 to solve $\boldsymbol{y}' = A\boldsymbol{y} + \boldsymbol{f}$ for the given matrix A, forcing function $\boldsymbol{f}$, and initial condition $\boldsymbol{y}(0)$.

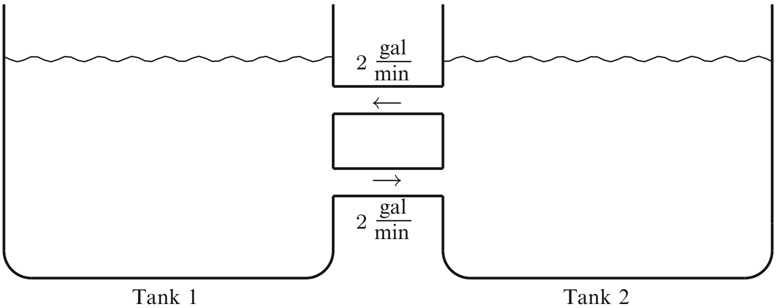
10. $A = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} e^t \\ e^t \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$
11. $A = \begin{bmatrix} -1 & 2 \\ -2 & -1 \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
12. $A = \begin{bmatrix} 2 & -5 \\ 1 & -2 \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} 2\cos t \\ \cos t \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$
13. $A = \begin{bmatrix} -1 & -4 \\ 1 & -1 \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$
14. $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} e^t \\ -e^t \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
15. $A = \begin{bmatrix} 5 & 2 \\ -8 & -3 \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} t \\ -2t \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

16. $A = \begin{bmatrix} -2 & 2 & 1 \\ 0 & -1 & 0 \\ 2 & -2 & -1 \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} e^{-2t} \\ 0 \\ -e^{-2t} \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$

17. $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & -1 \\ -2 & 1 & 3 \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} e^{2t} \\ e^{2t} \\ -e^{2t} \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

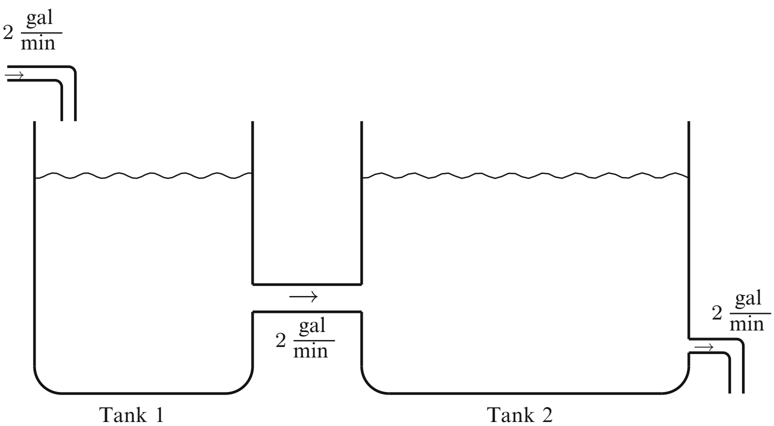
18–21. Solve each mixing problem.

18. Two tanks are interconnected as illustrated below.



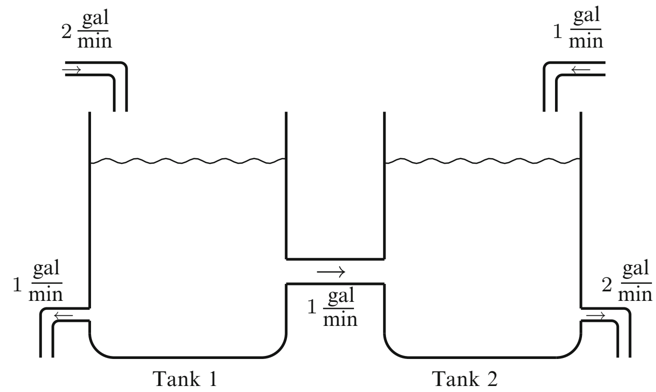
Assume that Tank 1 contains 1 gallon of brine in which 4 pounds of salt are initially dissolved and Tank 2 initially contains 1 gallon of pure water. Moreover, at time $t = 0$, the mixtures are pumped between the two tanks, each at a rate of 2 gal/min. Assume the mixtures are well stirred. Let $y_1(t)$ be the amount of salt in Tank 1 at time t and let $y_2(t)$ be the amount of salt in Tank 2 at time t. Determine y_1, y_2. Find the amount of salt in each tank after 30 seconds.

19. Two tanks are interconnected as illustrated below.



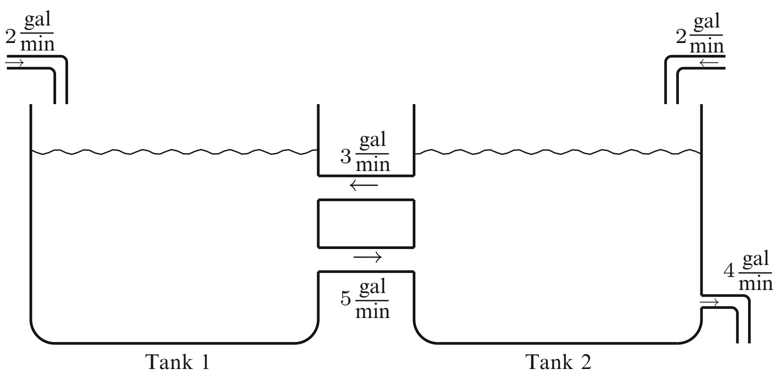
Assume that Tank 1 contains 1 gallon of brine in which 4 pounds of salt is initially dissolved and Tank 2 contains 2 gallons of pure water. Moreover, the mixture from Tank 1 is pumped into Tank 2 at a rate of 2 gal/min. Assume that a brine mixture containing 1 lb salt/gal enters Tank 1 at a rate of 2 gal/min and the well-stirred mixtures in both tanks are removed from Tank 2 at the same rate. Let $y_1(t)$ be the amount of salt in Tank 1 at time t and let $y_2(t)$ be the amount of salt in Tank 2 at time t. Determine y_1 and y_2. Determine when the concentration of salt in Tank 2 is $1/2$ lbs/gal.

20. Two tanks are interconnected as illustrated below.



Assume that Tank 1 contains 1 gallon of pure water and Tank 2 contains 1 gallon of brine in which 4 pounds of salt is initially dissolved. Moreover, the mixture from Tank 1 is pumped into Tank 2 at a rate of 1 gal/min. Assume that a brine mixture containing 4 lb salt/gal enters Tank 1 at a rate of 2 gal/min and pure water enters Tank 2 at a rate of 1 gal/min. Assume the tanks are well stirred. Brine is removed from Tank 1 at the rate 1 gal/min and from Tank 2 at a rate of 2 gal/min. Let $y_1(t)$ be the amount of salt in Tank 1 at time t and let $y_2(t)$ be the amount of salt in Tank 2 at time t. Determine y_1 and y_2. Determine when the amount of salt in Tank 2 is at a minimum. What is the minimum?

21. Two tanks are interconnected as illustrated below.



Assume initially that Tank 1 and Tank 2 each contains 1 gallon of pure water. Moreover, the mixture from Tank 1 is pumped into Tank 2 at a rate of 5 gal/min and the mixture from Tank 2 is pumped into Tank 1 at a rate of 3 gal/min. Assume that a brine mixture containing 1 lb salt/gal enters both tanks at a rate of 2 gal/min. Assume the tanks are well stirred. Brine is removed from Tank 2 at the rate 4 gal/min. Let $y_1(t)$ be the amount of salt in Tank 1 at time t and let $y_2(t)$ be the amount of salt in Tank 2 at time t. Determine y_1 and y_2.

9.6 The Phase Plane

This section addresses some of the qualitative features of homogeneous solutions to constant coefficient systems of differential equations. We restrict our attention to the case $n = 2$ and write the dependent variables in the form

$$z(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}.$$

Thus, if $z(0) = z_0$, then the initial value problem that we consider is

$$z'(t) = Az(t), \quad z(0) = z_0, \tag{1}$$

where A is a 2×2 matrix. We will think of each $z(t)$ as a point $(x(t), y(t))$ in the Euclidean plane, usually referred to as the ***phase plane*** in this context.[3] The set of points $\{(x(t), y(t)) : t \in \mathbb{R}\}$ traces out a ***path*** or ***orbit***, and to each path, we can associate a direction: the one determined by t increasing. Such directed paths are called ***trajectories***. The ***phase portrait*** shows trajectories for various initial values in the phase plane. The shape of the paths, the direction of the trajectories, and equilibrium solutions are some of the qualitative features in which we are interested. As we will see, the eigenvalues of A play a decisive role in determining many of important characteristics of the phase portrait.

Affine Equivalence

Our study of the phase portrait for $z' = Az$ can be simplified by considering an affine equivalent system $\boldsymbol{w}' = B\boldsymbol{w}$, where B is a 2×2 matrix that has a particularly simple form. Let P be a 2×2 invertible matrix. The change in variables

$$z = P\boldsymbol{w} \tag{2}$$

is called an ***affine transformation***. [4]More specifically, if

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}, \text{ and } \boldsymbol{w} = \begin{bmatrix} u \\ v \end{bmatrix},$$

[3]See the article "The Tangled Tale of Phase Space" by David D. Nolte (published in *Physics Today*, April 2010) for an account of the history of 'phase space'.

[4]More generally, an affine transformation is a transformation of the form $z = P\boldsymbol{w} + \boldsymbol{w}_0$, where P is an invertible matrix and $\boldsymbol{w}_0$ is a fixed ***translation*** vector. For our purposes, it will suffice to assume $\boldsymbol{w}_0 = 0$.

then the equation $\boldsymbol{z} = P\boldsymbol{w}$ becomes

$$x = p_{11}u + p_{12}v,$$
$$y = p_{21}u + p_{22}v.$$

Since P is invertible, we also have $\boldsymbol{w} = P^{-1}\boldsymbol{z}$. Thus, we are able to go from one set of variables to the other. Since P is a constant matrix, we have the following

$$\boldsymbol{w}' = P^{-1}\boldsymbol{z}' = P^{-1}A\boldsymbol{z} = P^{-1}AP\boldsymbol{w}. \tag{3}$$

If we set $B = P^{-1}AP$, then (3) becomes

$$\boldsymbol{w}' = B\boldsymbol{w} \tag{4}$$

and the associated initial condition becomes $\boldsymbol{w}_0 = \boldsymbol{w}(0) = P^{-1}\boldsymbol{z}_0$. Once $\boldsymbol{w}$ is determined, we are able to recover $\boldsymbol{z}$ by the equation $\boldsymbol{z} = P\boldsymbol{w}$, and vice versa, once $\boldsymbol{z}$ is determined, we are able to recover $\boldsymbol{w}$ by the equation $\boldsymbol{w} = P^{-1}\boldsymbol{z}$. The idea is to find an affine transformation P in such a way that $B = P^{-1}AP$ is particularly simple, for example, diagonal, something "close to diagonal", or something that is distinctively simple[5] (see "Jordan Canonical Forms" below for a description of exactly what we mean). Two matrices A and B are called ***similar*** if there is an invertible matrix P such that $B = P^{-1}AP$.

Affine transformations are important for us because certain shapes in the (u, v) phase plane are preserved in the (x, y) phase plane. Specifically, if $\boldsymbol{z} = P\boldsymbol{w}$ is an affine transformation, then

1. A line in the (u, v) phase plane is transformed to a line in the (x, y) phase plane. If the line in the (u, v) plane goes through the origin, then so does the transformed line.
2. An ellipse in the (u, v) phase plane is transformed to an ellipse in the (x, y) phase plane. (In particular, a circle is transformed to an ellipse.)
3. A spiral in the (u, v) phase plane is transformed to a spiral in the (x, y) phase plane.
4. A power curve[6] in the (u, v) phase plane is transformed to a power curve in the (x, y) phase plane, for example, parabolas and hyperbolas.
5. A tangent line L to a curve C in the (u, v) phase plane is transformed to the tangent line $P(L)$ to the curve $P(C)$ in the (x, y) phase plane.

[5] Many texts take this approach to solve constant coefficient systems. Our development of the Laplace transform allows us to get at the solution rather immediately. However, to understand some of the qualitative features, we make use of the notion of affine equivalence.

[6] By a power curve we mean the graph of a relation $Ax + By = (Cx + Dy)^p$, where p is a real number and all constants and variables are suitably restricted so the power is well defined.

6. A curve C that lies in a region R in the (u, v) phase plane is transformed to a curve $P(C)$ that lies in the region $P(R)$ in the (x, y) phase plane.

You will be guided through a proof of these statements in the exercises. In view of the discussion above, we say that the phase portraits of $z' = Az$ and $z' = Bz$ are ***affine equivalent*** if A and B are similar.

To illustrate the value of affine equivalence, consider the following example.

Example 1. Discuss the phase portrait for the linear differential system

$$z'(t) = Az(t),$$

where

$$A = \begin{bmatrix} 4 & 1 \\ 2 & 5 \end{bmatrix}.$$

► **Solution.** The characteristic polynomial of A is $c_A(s) = s^2 - 9s + 18 = (s-3)(s-6)$. The eigenvalues of A are thus 3 and 6. By Fulmer's method, we have $e^{At} = Me^{3t} + Ne^{6t}$ from which we get

$$\begin{aligned} I &= M + N \\ A &= 3M + 6N. \end{aligned}$$

From these equations, it follows that

$$\begin{aligned} M &= \frac{1}{3}(6I - A) = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -2 & 1 \end{bmatrix} \\ N &= \frac{1}{3}(A - 3I) = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}. \end{aligned}$$

Hence,

$$e^{At} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -2 & 1 \end{bmatrix} e^{3t} + \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} e^{6t}.$$

A short calculation gives the solution to $z'(t) = Az(t)$ with initial value $z(0) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ as

$$\begin{aligned} x(t) &= \frac{1}{3}(2c_1 - c_2)e^{3t} + \frac{1}{3}(c_1 + c_2)e^{6t} \\ y(t) &= -\frac{1}{3}(2c_1 - c_2)e^{3t} + \frac{2}{3}(c_1 + c_2)e^{6t} \end{aligned} \tag{5}$$

The orbits $\{(x(t), y(t)) : t \in \mathbb{R}\}$ are difficult to describe in general except for a few carefully chosen initial values. Notice that

1. If $c_1 = 0$ and $c_2 = 0$, then $x(t) = y(t) = 0$ so that the origin $(0, 0)$ is a trajectory.
2. If $c_2 = -c_1 \neq 0$, then $x(t) = c_1 e^{3t}$ and $y(t) = -c_1 e^{3t}$. This means that the trajectory is of the form
$$(x(t), y(t)) = (c_1 e^{3t}, -c_1 e^{3t}) = c_1 e^{3t}(1, -1).$$
The function e^{3t} is positive and increasing as a function of t. Thus, if c_1 is positive, then the trajectory is a half-line in the fourth quadrant consisting of all positive multiples of the vector $(1, -1)$ and pointing away from the origin. This is the trajectory marked **A** in Fig. 9.1 in the (x, y) phase plane. If c_1 is negative, then the trajectory is a half-line in the second quadrant consisting of all positive multiples of the vector $(-1, 1)$ and pointing away from the origin.
3. If $c_2 = 2c_1 \neq 0$, then $x(t) = c_1 e^{6t}$ and $y(t) = 2c_1 e^{6t}$. This means that the trajectory is of the form
$$(x(t), y(t)) = (c_1 e^{6t}, 2c_1 e^{6t}) = c_1 e^{6t}(1, 2).$$
Again e^{6t} is positive and increasing as a function of t. Thus, if c_1 is positive, the trajectory is a half-line in the first quadrant consisting of all positive multiples of the vector $(1, 2)$ and pointing away from the origin. This is the trajectory marked **D** in Fig. 9.1 in the (x, y) phase plane. If c_1 is negative, then the trajectory is a half-line in the third quadrant consisting of all positive multiples of the vector $(-1, -2)$ and pointing away from the origin.

For initial values other than the ones listed above, it is rather tedious to directly describe the trajectories. Notice though how a change in coordinates simplifies matters significantly. In (5), we can eliminate e^{6t} by subtracting y(t) from twice $x(t)$ and we can eliminate e^{3t} by adding $x(t)$ and $y(t)$. We then get

$$\begin{aligned} 2x(t) - y(t) &= (2c_1 - c_2)\, e^{3t} = k_1 e^{3t}, \\ x(t) + y(t) &= (c_1 + c_2)\, e^{6t} = k_2 e^{6t}, \end{aligned} \tag{6}$$

where $k_1 = 2c_1 - c_2$ and $k_2 = c_1 + c_2$. Now let

$$\begin{aligned} u &= 2x - y \\ v &= x + y \end{aligned} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

The matrix $\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$ is invertible and we can thus solve for $\begin{bmatrix} x \\ y \end{bmatrix}$ to get

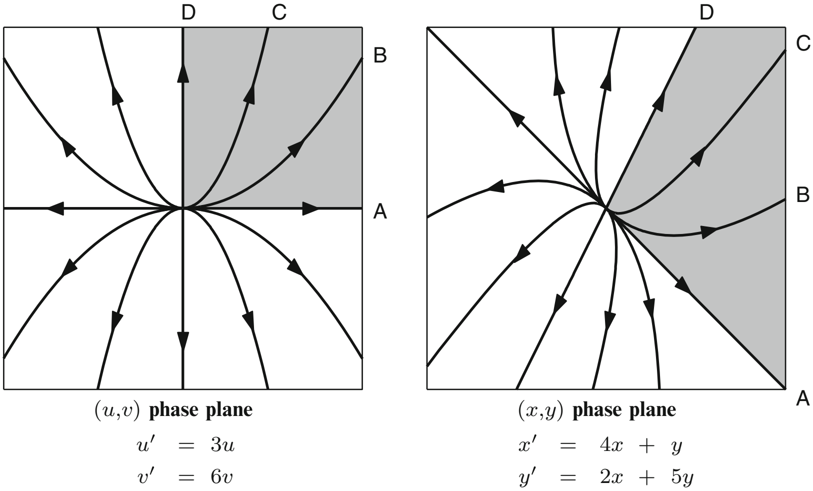


Fig. 9.1 Affine equivalent phase portraits

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}.$$

Let $P = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$. This is the affine transformation that implements the change in variables that we need. A simple calculation gives

$$B = P^{-1}AP = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 4 & 1 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 6 \end{bmatrix}, \tag{7}$$

a diagonal matrix consisting of the eigenvalues 3 and 6.

We make an important observation about P here: Since A has two distinct eigenvalues, there are two linearly independent eigenvectors. Notice that the first column of P is an eigenvector with eigenvalue 3 and the second column is an eigenvector with eigenvalue 6. The importance of this will be made clear when we talk about Jordan canonical forms below.

If $\boldsymbol{w} = \begin{bmatrix} u \\ v \end{bmatrix}$, then

$$\boldsymbol{w}' = B\boldsymbol{w},$$

and the initial condition is given by

$$\boldsymbol{w}(0) = P^{-1}\boldsymbol{z}(0) = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_1 \end{bmatrix} = \begin{bmatrix} 2c_1 - c_2 \\ c_1 + c_2 \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}.$$

The equations for $\boldsymbol{w}' = B\boldsymbol{w}$ are simply

$$u' = 3u,$$
$$v' = 6v.$$

The solutions can be computed directly and are

$$u(t) = k_1 e^{3t} \quad \text{and} \quad v(t) = k_2 e^{6t},$$

which are also consistent with (6). In these variables, it is a simple matter to compute the phase portrait. Let us first take care of some special cases. Notice that

1. If $k_1 = 0$ and $k_2 = 0$, then $u(t) = v(t) = 0$ so that the origin $(0, 0)$ is a trajectory.
2. If $k_1 \neq 0$ and $k_2 = 0$, then the trajectory is of the form

   $$(u(t), v(t)) = k_1 e^{3t}(1, 0).$$

   As before, we observe that e^{3t} is positive and increasing. Thus, if k_1 is positive, the trajectory is the positive u-axis pointing away from the origin. This is the trajectory marked **A** in Fig. 9.1 in the (u, v) phase plane. If k_1 is negative, then the trajectory is the negative u-axis pointing away from the origin.
3. If $k_1 = 0$ and $k_2 \neq 0$ then the trajectory is of the form

   $$(u(t), v(t)) = k_2 e^{6t}(0, 1).$$

   Again e^{6t} is positive and increasing. Thus if k_2 is positive, then the trajectory is the positive y-axis pointing away from the origin. This is the trajectory marked **D** in Fig. 9.1 in the (u, v) phase plane. If k_2 is negative, then the trajectory is the negative y-axis pointing away from the origin.

Now assume $k_1 \neq 0$ and $k_2 \neq 0$. Since e^{3t} and e^{6t} take on all positive real numbers and are increasing, the trajectories $(u(t), v(t)) = (k_1 e^{3t}, k_2 e^{6t})$ are located in the quadrant determined by the initial value (k_1, k_2) and are pointed in the direction away from the origin. To see what kinds of curves arise, let us determine how $u(t)$ and $v(t)$ are related. For notation's sake, we drop the "t" in $u(t)$ and $v(t)$. Observe that $k_2 u^2 = k_1^2 k_2 e^{6t} = k_1^2 v$, and hence,

$$v = \frac{k_2}{k_1^2} u^2.$$

Hence, a trajectory is that portion of a parabola that lies in the quadrant determined by the initial value (k_1, k_2). In Fig. 9.1, the two trajectories marked **B** and **C** are those trajectories that go through $(4, 3)$ and $(1, 1)$, respectively. The other unmarked trajectories are obtained from initial values, $(\pm 1, \pm 1)$ and $(\pm 4, \pm 3)$.

Now to determine the trajectories in the (x, y) phase plane for $z' = Az$, we utilize the affine transformation P. Since

$$P\begin{bmatrix} u \\ v \end{bmatrix} = \frac{u}{3}\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \frac{v}{3}\begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

it follows that the region determined by the first quadrant in the (u, v) phase plane transforms to the region consisting of all sums of nonnegative multiples of $(1, -1)$ and nonnegative multiples of $(1, 2)$. We have shaded those regions in Fig. 9.1. An affine transformation such as P transforms parabolas to parabolas and preserves tangent lines. The parabola marked **B** in the (u, v) plane lies in the first quadrant and is tangent to the trajectory marked **A** at the origin. Therefore, the transformed trajectory in the (x, y) phase plane must (1) lie in the shaded region, (2) be a parabola, and (3) be tangent to trajectory **A** at the origin. It is similarly marked **B**. Now consider the region between trajectory **B** and **D** in the (u, v) phase plane. Trajectory **C** lies in this region and must therefore transform to a trajectory which (1) lies in the region between trajectory **B** and **D** in the (x, y) phase plane, (2) is a parabola, and (3) is tangent to trajectory **A** at the origin. We have marked it correspondingly **C**. The analysis of the other trajectories in the (u, v) phase plane is similar; they are each transformed to parabolically shaped trajectories in the (x, y) phase plane. ◄

Jordon Canonical Forms

In the example above, we showed by (7) that A is similar to a diagonal matrix. In general, this cannot always be done. However, in the theorem below, we show that A is similar to one of four forms, called the ***Jordan Canonical Forms***.

Theorem 2. *Let A be a real 2×2 matrix. Then there is an invertible matrix P so that $P^{-1}AP$ is one of the following matrices:*

1. $J_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, $\begin{matrix} \lambda_1, \lambda_2 \in \mathbb{R} \\ \lambda_1 \neq \lambda_2 \end{matrix}$,
2. $J_2 = \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix}$, $\lambda \in \mathbb{R}$,
3. $J_3 = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$, $\lambda \in \mathbb{R}$,
4. $J_4 = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$, $\alpha \in \mathbb{R}$, $\beta > 0$.

Furthermore, the affine transformation P may be determined correspondingly as follows:

1. *If A has two distinct real eigenvalues, λ_1 and λ_2, then the first column of P is an eigenvector for λ_1 and the second column of P is an eigenvector for λ_2.*

2. *If A has only one real eigenvalue λ with eigenspace of dimension* 1*, then the first column of P may be chosen to be any vector v that is not an eigenvector and the second column of P is $(A - \lambda I)v$.*
3. *If A has only one real eigenvalue with eigenspace of dimension* 2 *then A is J_3. Hence, P may be chosen to be the identity.*
4. *If A has a complex eigenvalue, then one of them is of the form $\alpha - i\beta$ with $\beta > 0$. If w is a corresponding eigenvector, then the first column of P is the real part of w and the second column of P, is the imaginary part of w.*

Remark 3. Any one of the four matrices, $J_1, \ldots, J_4$ is called a ***Jordan matrix***. Note that the affine transformation P is not unique.

Proof. We consider the eigenvalues of A. There are four possibilities.

1. Suppose A has two distinct real eigenvalues λ_1 and λ_2. Let v_1 be an eigenvector with eigenvalue λ_1 and v_2 an eigenvector with eigenvalue λ_2. Let $P = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$, the matrix 2×2 matrix with v_1 the first column and v_2 the second column. Then

$$AP = \begin{bmatrix} Av_1 & Av_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = PJ_1.$$

   Now multiply both sides on the left by P^{-1} to get $P^{-1}AP = J_1$.
2. Suppose A has only a single real eigenvalue λ. Then the characteristic polynomial is $c_A(s) = (s-\lambda)^2$. Let E_λ be the eigenspace for λ and suppose further that $E_\lambda \neq \mathbb{R}^2$. Let v_1 be a vector outside of E_λ. Then $(A - \lambda I)v_1 \neq 0$. However, by the Cayley-Hamilton theorem (see Appendix A.5), $(A - \lambda I)(A - \lambda I)v_1 = (A - \lambda I)^2 v_1 = 0$. It follows that $(A - \lambda I)v_1$ is an eigenvector. Let $v_2 = (A - \lambda I)v_1$. Then $Av_1 = ((A - \lambda I) + \lambda I)v_1 = v_2 + \lambda v_1$ and $Av_2 = \lambda v_2$. Let $P = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$ be the matrix with v_1 in the first column and v_2 in the second column. Then

$$AP = \begin{bmatrix} Av_1 & Av_2 \end{bmatrix} = \begin{bmatrix} v_2 + \lambda v_1 & \lambda v_2 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix} = PJ_2.$$

   Now multiply both sides on the left by P^{-1} to get $P^{-1}AP = J_2$.
3. Suppose A has only a single real eigenvalue λ and the eigenspace $E_\lambda = \mathbb{R}^2$. Then $Av = \lambda v$ for all $v \in \mathbb{R}^2$. This means A must already be $J_3 = \lambda I$.
4. Suppose A does not have a real eigenvalue. Since A is real, the two complex eigenvalues are of the form $\alpha + i\beta$ and $\alpha - i\beta$, with $\beta > 0$. Let w be an eigenvector in the complex plane $\mathbb{C}^2$ with eigenvalue $\alpha - i\beta$. Let v_1 be the real part of w and v_2 the imaginary part of w. Then $w = v_1 + iv_2$ and since

$$A(v_1 + iv_2) = (\alpha - i\beta)(v_1 + iv_2) = (\alpha v_1 + \beta v_2) + i(-\beta v_1 + \alpha v_2)$$

   we get

$$\begin{aligned} Av_1 &= \alpha v_1 + \beta v_2, \\ Av_2 &= -\beta v_1 + \alpha v_2. \end{aligned}$$

Let $P = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$. Then

$$AP = \begin{bmatrix} Av_1 & Av_2 \end{bmatrix} = \begin{bmatrix} \alpha v_1 + \beta v_2 & -\beta v_1 + \alpha v_2 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} = PJ_4.$$

Now multiply both sides on the left by P^{-1} to get $P^{-1}AP = J_4$. ☐

Example 4. For each of the following matrices, determine an affine transformation P so the $P^{-1}AP$ is a Jordan matrix:

1. $A = \begin{bmatrix} 6 & -2 \\ -3 & 7 \end{bmatrix}$ 2. $A = \begin{bmatrix} -5 & 2 \\ -2 & -1 \end{bmatrix}$ 3. $A = \begin{bmatrix} -5 & -8 \\ 4 & 3 \end{bmatrix}$

► **Solution.** 1. The characteristic polynomial is $c_A(s) = (s-6)(s-7) - 6 = s^2 - 13s + 36 = (s-4)(s-9)$. There are two distinct eigenvalues, 4 and 9. It is an easy calculation to see that $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue 4 and $v_2 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ is an eigenvector with eigenvalue 9. Let

$$P = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}.$$

Then an easy calculation gives

$$P^{-1}AP = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}.$$

2. The characteristic polynomial is $c_A(s) = (s+3)^2$. Thus, $\lambda = -3$ is the only eigenvalue. Since $A - -3I = \begin{bmatrix} -2 & 2 \\ -2 & 2 \end{bmatrix}$, it is easy to see that all eigenvectors are multiples of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Let $v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Then v_1 is not an eigenvector. Let

$$v_2 = (A - \lambda I)v = \begin{bmatrix} -2 & 2 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -4 \\ -4 \end{bmatrix}.$$

Let

$$P = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 1 & -4 \\ -1 & -4 \end{bmatrix}.$$

Then an easy calculation gives

$$P^{-1}AP = \begin{bmatrix} -3 & 0 \\ 1 & -3 \end{bmatrix}.$$

3. The characteristic polynomial is $c_A(s) = s^2 + 2s + 17 = (s+1)^2 + 16$ so the eigenvalues are $-1 \pm 4i$. We compute an eigenvector with eigenvalue $-1 - 4i$. To do this, we solve

$$((-1-4i)I - A)\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

for a and b. This is equivalent to

$$\begin{aligned} (4-4i)a + 8b &= 0, \\ -4a + (-4-4i)b &= 0. \end{aligned}$$

If we choose $b = 1$, then $a = \dfrac{-8}{4-4i} = \dfrac{-8(4+4i)}{(4-4i)(4+4i)} = -1-i$. Therefore,

$$v = \begin{bmatrix} -1-i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + i\begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

is an eigenvector for A with eigenvalue $-1-4i$. Let $v_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ be the real and imaginary parts of v. Let

$$P = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}.$$

Then

$$P^{-1}AP = \begin{bmatrix} -1 & -4 \\ 4 & -1 \end{bmatrix}.$$ ◄

Notice in the following example the direct use of affine equivalence.

Example 5. Discuss the phase portrait for the linear differential system

$$z'(t) = Az(t),$$

where

$$A = \begin{bmatrix} -5 & -8 \\ 4 & 3 \end{bmatrix}.$$

► **Solution.** The characteristic polynomial is

$$c_A(s) = s^2 + 2s + 17 = (s+1)^2 + 16.$$

It is straightforward to determine that

$$e^{At} = e^{-t}\begin{bmatrix} \cos 4t - \sin 4t & 2\sin 4t \\ \sin 4t & \cos 4t + \sin 4t \end{bmatrix}.$$

For a given initial condition $z(0) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, we have

$$\begin{aligned} x(t) &= e^{-t}(c_1 \cos 4t + (2c_2 - c_1)\sin 4t) \\ y(t) &= e^{-t}((c_1 + c_2)\sin 4t + c_2 \cos 4t). \end{aligned}$$

The phase plane portrait for this system is very difficult to directly deduce without the help of an affine transformation. In Example 4, part 3, we determined

$$P = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}$$

and

$$B = P^{-1}AP = \begin{bmatrix} -1 & -4 \\ 4 & -1 \end{bmatrix}.$$

In the new variables $z = Pw$, we get $w' = Bw$, and the initial condition is given by

$$w(0) = P^{-1}z(0) = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_2 \\ -c_1 - c_2 \end{bmatrix}.$$

Let $\begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} c_2 \\ -c_1 - c_2 \end{bmatrix}$. A straightforward calculation gives

$$e^{Bt} = e^{-t}\begin{bmatrix} \cos 4t & -\sin 4t \\ \sin 4t & \cos 4t \end{bmatrix}$$

and

$$\begin{bmatrix} u \\ v \end{bmatrix} = e^{Bt}\begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = e^{-t}\begin{bmatrix} k_1 \cos 4t - k_2 \sin 4t \\ k_1 \sin 4t + k_2 \cos 4t \end{bmatrix}.$$

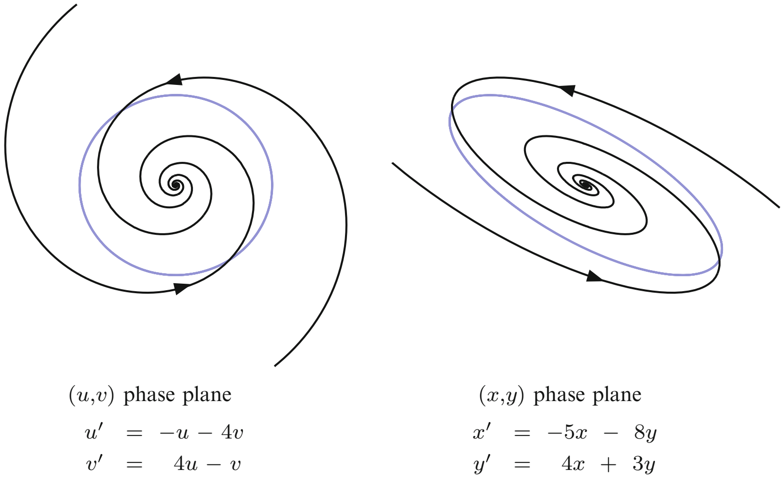


Fig. 9.2 Spiral phase portraits

If ϕ is the angle made by the vector (k_1, k_2) and the positive u-axis, then

$$\cos\phi = \frac{k_1}{\sqrt{k_1^2 + k_2^2}} \quad \text{and} \quad \sin\phi = \frac{k_2}{\sqrt{k_1^2 + k_2^2}}.$$

We can then express u and v as

$$u(t) = \sqrt{k_1^2 + k_2^2}\mathrm{e}^{-t}(\cos\phi\cos 4t - \sin\phi\sin 4t) = \sqrt{k_1^2 + k_2^2}\mathrm{e}^{-t}(\cos(4t + \phi)),$$
$$v(t) = \sqrt{k_1^2 + k_2^2}\mathrm{e}^{-t}(\cos\phi\sin 4t + \sin\phi\cos 4t) = \sqrt{k_1^2 + k_2^2}\mathrm{e}^{-t}(\sin(4t + \phi)),$$

Now observe that

$$u^2(t) + v^2(t) = (k_1^2 + k_2^2)\mathrm{e}^{-2t}.$$

If $u^2(t) + v^2(t)$ were constant, then the trajectories would be circles. However, the presence of the factor e^{-2t} shrinks the distance to the origin as t increases. The result is that the trajectory is a spiral pointing toward the origin. We show two such trajectories in Fig. 9.2 in the (u, v) phase plane.

Notice that the trajectories in the (u, v) phase plane rotate one onto another. By this we mean if we rotate a fixed trajectory, you will get another trajectory. In fact, all trajectories can be obtained by rotating a fixed one. Specifically, if we rotate a trajectory by an angle θ. Then the matrix that implements that rotation is given by

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

It is a nice property about rotation matrices that they commute. In other words, if θ_1 and θ_2 are two angles, then

$$R(\theta_1)R(\theta_2) = R(\theta_2)R(\theta_1).$$

Now observe that $\mathrm{e}^{Bt} = \mathrm{e}^{-t}R(4t)$. When we apply $R(\theta)$ to $\boldsymbol{w}$, we get

$$R(\theta)\boldsymbol{w}(t) = \mathrm{e}^{-t}R(\theta)R(4t)\boldsymbol{k} = \mathrm{e}^{-t}R(4t)\,(R(\theta)\boldsymbol{k}) = \mathrm{e}^{Bt}(R(\theta)\boldsymbol{k}). \tag{8}$$

Thus, $R(\theta)\boldsymbol{w}(t)$ is the solution to $\boldsymbol{w}' = B\boldsymbol{w}$ with just a different initial condition, namely, $R(\theta)\boldsymbol{k}$.

Now to see what is going on in (x, y) phase plane, we use the affine map P. In the (u, v) phase plane, we have drawn a gray circle centered at the origin (it is not a trajectory). Suppose its radius is k. Recall that an affine transformation maps circles to ellipses. Specifically, since $\boldsymbol{w} = P^{-1}\boldsymbol{z}$, we have $u = y$ and $v = -x - y$. From this we get

$$k^2 = u^2 + v^2 = y^2 + x^2 + 2xy + y^2 = x^2 + 2xy + 2y^2.$$

This equation defines an ellipse. That ellipse is drawn in gray in the (x, y) phase plane. The trajectories in the (x, y) phase plane are still spirals that point toward the origin but elongate in the direction of the semimajor axis of the ellipse. ◀

Critical Points

A solution for which the associated path

$$\{(x(t), y(t)), t \in \mathbb{R}\}$$

is just a point is called an ***equilibrium solution*** or ***critical point***. This means then that $x(t) = c_1$ and $y(t) = c_2$ for all $t \in \mathbb{R}$ and occurs if and only if

$$A\boldsymbol{c} = \boldsymbol{0}, \quad \boldsymbol{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \tag{9}$$

If A is nonsingular, that is, $\det A \neq 0$, then (9) implies $\boldsymbol{c} = \boldsymbol{0}$. In the phase portrait, the origin is an orbit and it is the only orbit consisting of a single point. On the other hand, if A is singular, then the solutions to (9) consists of the whole plane in the case $A = 0$ and a line through the origin in the case $A \neq 0$. If $A = 0$, then the phase portrait is trivial; each point in the plane is an orbit. If $A \neq 0$ (but singular), then each point on the line is an orbit and off that line there will be nontrivial orbits. We will assume A is nonsingular for the remainder of this section and develop the case where A is nonzero but singular in the exercises.

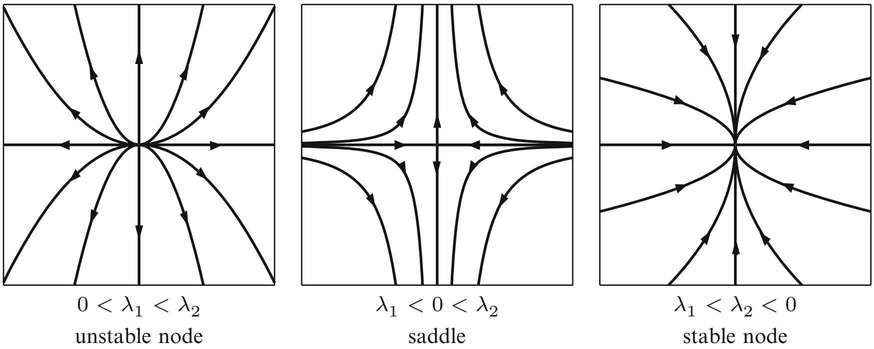


Fig. 9.3 Phase portrait for the canonical system J_1

The Canonical Phase Portraits

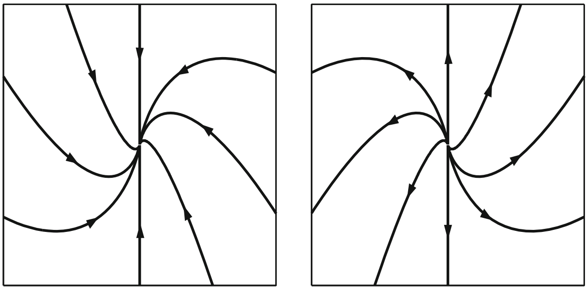
In view of Theorem 2, all phase portraits are affine equivalent to just a few simple types. These types are referred to as the ***canonical phase portraits***. There are four of them (if we do not take into account directions associated with the paths) corresponding to each of the four Jordan canonical forms. Let us consider each case.

$\mathbf{J_1}$: In J_1, we will order the eigenvalues λ_1 and λ_2 to satisfy $\lambda_1 < \lambda_2$. Since $\det A \neq 0$, neither λ_1 nor λ_2 are zero. The solutions to $z' = J_1 z$ are $x(t) = c_1 e^{\lambda_1 t}$ and $y(t) = c_2 e^{\lambda_2 t}$. Therefore, the trajectories lie on the power curve defined by $|y(t)| = K\,|x(t)|^{\frac{\lambda_2}{\lambda_1}}$ (a similar calculation was done in Example 1). The shape of the trajectories are determined by $p = \lambda_2/\lambda_1$. Refer to Fig. 9.3 for the phase portrait for each of the following three subcases:

1. Suppose $0 < \lambda_1 < \lambda_2$. Then $x(t)$ and $y(t)$ become infinite as t gets large. The trajectories lie on the curve $|y| = K\,|x|^p$, $p > 1$. All of the trajectories point away from the origin; the origin is said to be an ***unstable node***.
2. Suppose $\lambda_1 < 0 < \lambda_2$. Then $x(t)$ approaches zero while $y(t)$ becomes infinite as t gets large. The trajectories lie on the curve $|y| = K/\,|x|^q$, where $q = -p > 0$. The origin is said to be a ***saddle***.
3. Suppose $\lambda_1 < \lambda_2 < 0$. Then $x(t)$ and $y(t)$ approach zero as t gets large. The trajectories lie on the curve $|y| = K\,|x|^p$, $0 < p < 1$. In this case, all the trajectories point toward the origin; the origin is said to be a ***stable node***.

$\mathbf{J_2}$: It is straightforward to see that the solutions to $z' = J_2 z$ are $x(t) = c_1 e^{\lambda t}$ and $y(t) = (c_1 t + c_2)e^{\lambda t}$. Observe that if $c_1 = 0$, then $x(t) = 0$ and $y(t) = c_2 e^{\lambda t}$. If $c_2 > 0$, then the positive y-axis is a trajectory, and if $c_2 < 0$, then the negative y-axis is a trajectory. Now assume $c_1 \neq 0$. We can solve for y in terms of x as

Fig. 9.4 Phase portrait for the canonical system J_2



follows. In the equation, $x(t) = c_1 e^{\lambda t}$, we get $t = \frac{1}{\lambda} \ln\left(\frac{x}{c_1}\right)$. Substituting into $y(t)$ gives

$$y = \frac{x}{\lambda} \ln\left(\frac{x}{c_1}\right) + \frac{c_2 x}{c_1}. \tag{10}$$

Note that if $c_1 > 0$, then $x > 0$ and the trajectory lies in the right half plane, and if $c_1 < 0$, then $x < 0$ and the trajectory lies in the left half plane. An easy exercise shows that the graph of Equation (10) has a vertical tangent at the origin; the origin is called an ***improper node***. Now refer to Fig. 9.4 for the phase portrait for each of the following two subcases:

1. Suppose $\lambda < 0$. Then $x(t)$ and $y(t)$ approach zero as t gets large. Thus, all trajectories point toward the origin. In this case, the origin is ***stable***. If $c_1 < 0$, then the trajectory is concave upward and has a single local minimum. If $c_1 > 0$, then the trajectory is concave downward and has a single local maximum.
2. Suppose $\lambda > 0$. Then $x(t)$ and $y(t)$ approach infinity as t gets large. Thus, all trajectories point away from the origin. In this case, the origin is ***unstable***. If $c_1 < 0$, then the trajectory is concave downward and has a single local maximum. If $c_1 > 0$, then the trajectory is concave upward and has a single local minimum.

$\mathbf{J_3}$: It is straightforward to see that the solutions to $z' = J_3 z$ are $x(t) = c_1 e^{\lambda t}$ and $y(t) = c_2 e^{\lambda t}$. Thus, the orbits are of the form $(x(t), y(t)) = e^{\lambda t}(c_1, c_2)$ and hence are rays from the origin through the initial condition (c_1, c_2). The origin is called a ***star node***. Now refer to Fig. 9.5 for the phase portrait for each of the following two subcases:

1. If $\lambda < 0$, then $x(t)$ and $y(t)$ approach zero as x gets large. All trajectories point toward the origin. In this case, the origin is a ***stable star node***.
2. If $\lambda > 0$, then $x(t)$ and $y(t)$ approach infinity as x gets large. All trajectories point away from the origin. In this case, the origin is an ***unstable star node***.

$\mathbf{J_4}$: By a calculation similar to the one done in Example 5, it is easy to see that the solutions to $z' = J_3 z$ are $x(t) = e^{\alpha t}(c_1 \cos \beta t - c_2 \sin \beta t) = |\boldsymbol{c}|\, e^{\alpha t} \cos(\beta t + \phi)$

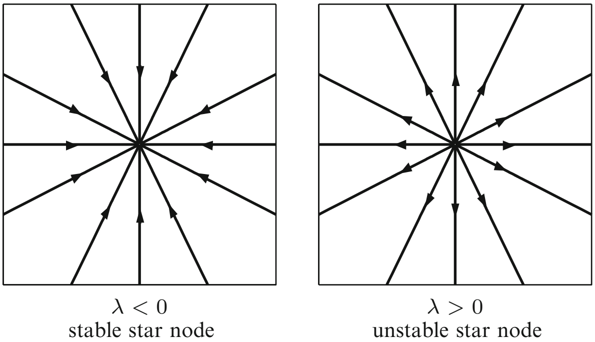


Fig. 9.5 Phase portrait for the canonical system J_3

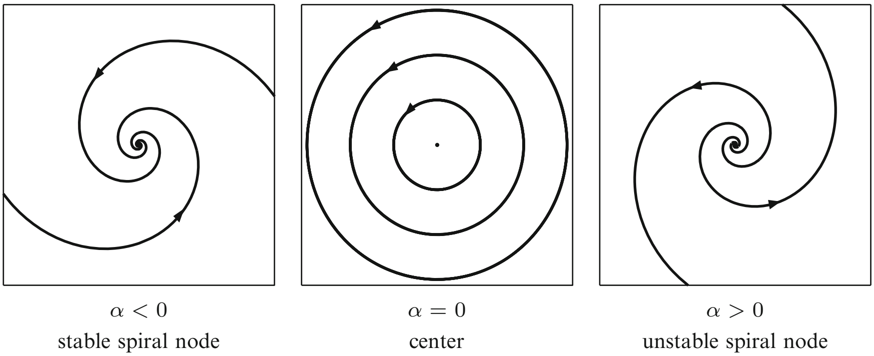


Fig. 9.6 Phase portrait for the canonical system J_4

and $y(t) = \mathrm{e}^{\alpha t}(c_1 \sin \beta t + c_2 \cos \beta t) = |\boldsymbol{c}|\, \mathrm{e}^{\alpha t} \sin(\beta t + \phi)$, where $|\boldsymbol{c}| = \sqrt{c_1^2 + c_2^2}$ and ϕ is the angle made by the vector (c_1, c_2) and the x-axis. From this it follows that

$$x^2 + y^2 = |\boldsymbol{c}|^2\, \mathrm{e}^{2\alpha t}$$

and the trajectories are spirals if $\alpha \neq 0$. Now refer to Fig. 9.6 for the phase portrait for each of the following three subcases:

1. If $\alpha < 0$, then $x(t)$ and $y(t)$ approach zero as t gets large. Thus, the trajectories point toward the origin; the origin is called a ***stable spiral node***.
2. If $\alpha = 0$, then $x^2 + y^2 = |\boldsymbol{c}|^2$ and the trajectories are circles with center at the origin; the origin is simply called a ***center***.
3. If $\alpha > 0$, then $x(t)$ and $y(t)$ approach infinity as t gets large. Thus, the trajectories point away from the origin; the origin is called a ***unstable spiral node***.

Classification of Critical Points

Now let A be any nonsingular matrix. The critical point $(0, 0)$ is classified as a ***node, center, or saddle*** according to whether the critical point in its affine equivalent is a node, center, or saddle. In like manner, we extend the adjectives proper, improper, stable, unstable, spiral, and star. Thus, in Example 1, the origin is an unstable node, and in Example 5, the origin is stable spiral node. Below we summarize the classification of the critical points in terms of the eigenvalues of A.

Classification of critical points

Jordan form	Eigenvalues of A	Critical point
J_1 :	$\lambda_1 \neq \lambda_2$	
	$\lambda_1 < \lambda_2 < 0$	Stable node
	$\lambda_1 < 0 < \lambda_2$	Saddle
	$0 < \lambda_1 < \lambda_2$	Unstable node
J_2 :	$\lambda \neq 0$	
	$\lambda < 0$	Stable improper node
	$\lambda > 0$	Unstable improper node
J_3 :	$\lambda \neq 0$	
	$\lambda < 0$	Stable star node
	$\lambda > 0$	Unstable star node
J_4 :	$\alpha \pm i\beta,\ \beta > 0$	
	$\alpha < 0$	Stable spiral node
	$\alpha = 0$	Center
	$\alpha > 0$	Unstable spiral node

Example 6. Classify the critical points for the system $z' = Az$ where

1. $A = \begin{bmatrix} 6 & -2 \\ -3 & 7 \end{bmatrix}$ 2. $A = \begin{bmatrix} -5 & 2 \\ -2 & -1 \end{bmatrix}$ 3. $A = \begin{bmatrix} -5 & -8 \\ 4 & 3 \end{bmatrix}$

► **Solution.** 1. In Example 4 part (1), we found that A is of type J_1 with positive eigenvalues, $\lambda_1 = 4$ and $\lambda_2 = 9$. The origin is an unstable node.

2. In Example 4 part (2), we found that A is of type J_2 with single eigenvalue, $\lambda = -3$. Since it is negative, the origin is an improper stable node.
3. In Example 4 part (3), we found that A is of type J_4 with eigenvalue, $\lambda = -1 \pm 4i$. Since the real part is negative, the origin is a stable star node. ◄

Exercises

1–10. For each of the following matrices A determine an affine transformation P and a Jordan matrix J so that $J = P^{-1}AP$. Then classify the critical point of A.

1. $A = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}$
2. $A = \begin{bmatrix} -3 & 2 \\ -2 & 1 \end{bmatrix}$
3. $A = \begin{bmatrix} -5 & -2 \\ 5 & 1 \end{bmatrix}$
4. $A = \begin{bmatrix} 2 & -2 \\ 4 & -2 \end{bmatrix}$
5. $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$
6. $A = \begin{bmatrix} 4 & 1 \\ -1 & 2 \end{bmatrix}$
7. $A = \begin{bmatrix} 5 & 3 \\ -1 & 1 \end{bmatrix}$
8. $A = \begin{bmatrix} 1 & -2 \\ 4 & -5 \end{bmatrix}$
9. $A = \begin{bmatrix} 3 & 1 \\ -8 & -1 \end{bmatrix}$
10. $A = \begin{bmatrix} -3 & 0 \\ 0 & -3 \end{bmatrix}$

11–14. In the following exercises we examine how an affine transformation preserves basic kinds of shapes. Let P be an affine transformation.

11. The general equation of a line is $Du + Ev + F = 0$, with D and E not both zero. Show that the change of variable $\boldsymbol{z} = P\boldsymbol{w}$ transforms a line L in the (u, v) plane to a line $P(L)$ in the (x, y) plane. If the line goes through the origin show that the transformed line also goes through the origin.
12. The general equation of a conic section is given by

$$Au^2 + Buv + Cv^2 + Du + Ev + F = 0,$$

where A, B, and C are not all zero. Let $\Delta = B^2 - 4AC$ be the ***discriminant***. If

1. $\Delta < 0$ the graph is an ellipse.
2. $\Delta = 0$ the graph is a parabola.
3. $\Delta > 0$ the graph is a hyperbola.

Show that the change of variable $\boldsymbol{z} = P\boldsymbol{w}$ transforms an ellipse to an ellipse, a parabola to a parabola, and a hyperbola to a hyperbola.

13. Suppose C is a power curve, i.e., the graph of a relation $Au + Bv = (Cu + Dv)^p$, where p is a real number and the variables are suitably restricted so the power is well defined. Show that $P(C)$ is again a power curve.
14. Suppose C is a differentiable curve with tangent line L at the point (u_0, v_0) in the (u, v) plane. Show that $P(L)$ is a tangent line to the curve $P(C)$ at the point $P(u_0, v_0)$.

15–19. In this set of exercises we consider the phase portraits when A is non zero but singular. Thus assume $\det A = 0$ and $A \neq 0$.

15. Show that A is similar to one of the following matrices:

$$J_1 = \begin{bmatrix} 0 & 0 \\ 0 & \lambda \end{bmatrix}, \lambda \neq 0 \quad \text{or} \quad J_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

(Hint: Since 0 is an eigenvalue consider two cases: the second eigenvalue is nonzero or it is 0. Then mimic what was done for the cases J_1 and J_2 in Theorem 2.)

16. Construct the Phase Portrait for $J_1 = \begin{bmatrix} 0 & 0 \\ 0 & \lambda \end{bmatrix}$.
17. Construct the Phase Portrait for $J_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$.
18. Suppose $\det A = 0$ and A is similar to J_1. Let λ be the nonzero eigenvalue. Show that the phase portrait for A consists of equilibrium points on the zero eigenspace and half lines parallel to the eigenvector for the nonzero eigenvalue λ with one end at an equilibrium point. If $\lambda > 0$ then the half lines point away from the equilibrium point and if $\lambda < 0$ they point toward the equilibrium point.
19. Suppose $\det A = 0$ and A is similar to J_2. Show that the phase portrait for A consists of equilibrium points on the zero eigenspace and lines parallel to the eigenspace.

20–23. In this set of problems we consider some the properties mentioned in the text about the canonical phase portrait J_2. Let (c_1, c_2) be a point in the plane but not on the y axis and let $\lambda \neq 0$. Let $y = \dfrac{x}{\lambda} \ln\left(\dfrac{x}{c_1}\right) + \dfrac{c_2 x}{c_1}$.

20. If $c_1 > 0$ show $\lim_{x \to 0^+} y = 0$ and if $c_1 < 0$ $\lim_{x \to 0^-} y = 0$.
21. Show that y has a vertical tangent at the origin.
22. Show that y has a single critical point.
23. Assume $c_1 > 0$ and hence $x > 0$. Show y is concave upward on $(0, \infty)$ if $\lambda > 0$ and y is concave downward on $(0, \infty)$ if $\lambda < 0$.

9.7 General Linear Systems

In this section, we consider the broader class of linear differential systems where the coefficient matrix $A(t)$ in

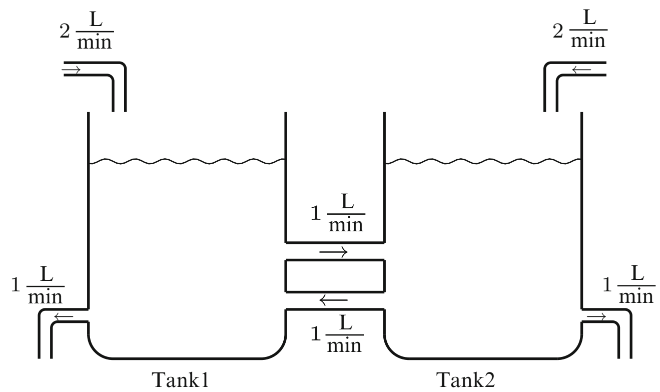
$$\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t) + \boldsymbol{f}(t) \tag{1}$$

is a function of t and not necessarily a constant matrix. Under rather mild conditions on $A(t)$ and the ***forcing function*** $\boldsymbol{f}(t)$, there is an existence and uniqueness theorem. The linearity of (1) implies that the structure of the solution set is very similar to that of the constant coefficient case. However, it can be quite difficult to find solution methods unless rather strong conditions are imposed on $A(t)$.

An Example: The Mixing Problem

The following example shows that a simple variation of the mixing problem produces a linear differential system with nonconstant coefficient matrix.

Example 1. Two tanks are interconnected as illustrated below.



Assume that Tank 1 initially contains 1 liter of brine in which 8 grams of salt are dissolved and Tank 2 initially contains 1 liter of pure water. The mixtures are then pumped between the two tanks, 1 L/min from Tank 1 to Tank 2 and 1 L/min from Tank 2 back to Tank 1. Assume that a brine mixture containing 6 grams salt/L enters Tank 1 at a rate of 2 L/min, and the well-stirred mixture is removed from Tank 1 at the rate of 1 L/min. Assume that a brine mixture containing 10 grams salt/L enters Tank 2 at a rate of 2 L/min and the well-stirred mixture is removed from Tank 2 at the rate of 1 L/min. Let $y_1(t)$ be the amount of salt in Tank 1 at time t and let $y_2(t)$ be the amount of salt in Tank 2 at time t. Determine a linear differential system that describes how y_1, y_2, and their derivatives are related.

► **Solution.** In this example, the amount of brine solution is not constant. The net increase of brine in each tank is one liter per minute. Thus, if $v_1(t)$ and $v_2(t)$ represent the amount of brine in Tank 1 and Tank 2, respectively, then $v_1(t) = v_2(t) = 1 + t$. Further, the concentration of salt in each tank is given by $\frac{y_1(t)}{1+t}$ and $\frac{y_2(t)}{1+t}$. As usual, y_1' and y_2' are the differences between the input rate of salt and the output rate of salt, and each rate is the product of the flow rate and the concentration. The relevant rates of change are summarized in the following table.

From	To	Rate		
Outside	Tank 1	(6 g/L)·(2 L/min)	=	12 g/min
Tank 1	Outside	$\left(\frac{y_1(t)}{1+t}\text{ g/L}\right)\cdot(1\text{L/min})$	=	$\frac{y_1(t)}{1+t}$ g/min
Tank 1	Tank2	$\left(\frac{y_1(t)}{1+t}\text{ g/L}\right)\cdot(1\text{ L/min})$	=	$\frac{y_1(t)}{1+t}$ g/min
Tank 2	Tank 1	$\left(\frac{y_2(t)}{1+t}\text{g/L}\right)\cdot(1\text{ L/min})$	=	$\frac{y_2(t)}{1+t}$ g/min
Tank 2	Outside	$\left(\frac{y_2(t)}{10}\text{g/L}\right)\cdot 1\text{ L/min}$	=	$\frac{y_2(t)}{1+t}$ g/min
Outside	Tank 2	(10 g/L) · (2 L/min)	=	20 g/min

The input and output rates are given as follows:

Tank	Input rate	Output rate
1	$12 + \frac{y_2(t)}{1+t}$	$\frac{2y_1(t)}{1+t}$
2	$20 + \frac{y_1(t)}{1+t}$	$\frac{2y_2(t)}{1+t}$

We thus obtain

$$\begin{aligned} y_1'(t) &= \frac{-2}{1+t}y_1(t) + \frac{1}{1+t}y_2(t) + 12, \\ y_2'(t) &= \frac{1}{1+t}y_1(t) - \frac{2}{1+t}y_2(t) + 20. \end{aligned}$$

The initial conditions are $y_1(0) = 8$ and $y_2(0) = 0$. We may now write the linear differential system in the form

$$\boldsymbol{y}' = A(t)\boldsymbol{y}(t) + \boldsymbol{f}(t), \quad \boldsymbol{y}(0) = \boldsymbol{y}_0,$$

where

$$A(t) = \begin{bmatrix} \dfrac{-2}{1+t} & \dfrac{1}{1+t} \\ \dfrac{1}{1+t} & \dfrac{-2}{1+t} \end{bmatrix}, \quad \boldsymbol{f}(t) = \begin{bmatrix} 12 \\ 20 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{y}(0) = \begin{bmatrix} 8 \\ 0 \end{bmatrix}.$$

We will show how to solve this nonconstant linear differential system later in this section. ◀

The Existence and Uniqueness Theorem

The following theorem is the fundamental result for linear systems. It guarantees that solutions exist, and if we can find a solution to a initial value problem by any means whatsoever, then we know that we have found the only possible solution.

Theorem 2 (Existence and Uniqueness). [7]*Suppose that the $n \times n$ matrix function $A(t)$ and the $n \times 1$ matrix function $\boldsymbol{f}(t)$ are both continuous on an interval I in $\mathbb{R}$. Let $t_0 \in I$. Then for every choice of the vector $\boldsymbol{y}_0$, the initial value problem*

$$\boldsymbol{y}' = A(t)\boldsymbol{y} + \boldsymbol{f}(t), \quad \boldsymbol{y}(t_0) = \boldsymbol{y}_0,$$

has a unique solution $\boldsymbol{y}(t)$ which is defined on the same interval I.

Remark 3. How is this theorem related to existence and uniqueness theorems we have stated previously?

- If $A(t) = A$ is a constant matrix, then Theorem 2 is precisely Theorem 2 of Sect. 9.5 where we have actually provided a solution method and a formula.
- If $n = 1$, then this theorem is just Corollary 8 of Sect. 1.5. In this case, we have actually proved the result by exhibiting a formula for the unique solution. However, for general n, there is no formula like (15) of Sect. 1.5, unless $A(t)$ satisfies certain stronger conditions.
- Theorem 6 of Sect. 5.1 is a corollary of Theorem 2. Indeed, if $n = 2$,

$$A(t) = \begin{bmatrix} 0 & 1 \\ -\frac{a_0(t)}{a_2(t)} & -\frac{a_1(t)}{a_2(t)} \end{bmatrix}, \; \boldsymbol{f}(t) = \begin{bmatrix} 0 \\ f(t) \end{bmatrix}, \; \boldsymbol{y}_0 = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix}, \; \text{and } \boldsymbol{y} = \begin{bmatrix} y \\ y' \end{bmatrix},$$

[7]A proof of this result can be found in the text *An Introduction to Ordinary Differential Equations* by Earl Coddington, Prentice Hall, (1961), Page 256.

then the second order linear initial value problem

$$a_2(t)y'' + a_1(t)y' + a_0(t)y = f(t), \quad y(t_0) = y_0,\ y'(t_0) = y_1$$

has the solution $y(t)$ if and only if the first order linear system

$$\boldsymbol{y}' = A(t)\boldsymbol{y} + \boldsymbol{f}(t), \quad \boldsymbol{y}(t_0) = \boldsymbol{y}_0$$

has the solution $\boldsymbol{y}(t) = \begin{bmatrix} y(t) \\ y'(t) \end{bmatrix}$. You should convince yourself of the validity of this statement.

Linear Homogeneous Differential Equations

In Theorem 4 of Sect. 9.2, we showed that each solution to (1) takes the form $\boldsymbol{y} = \boldsymbol{y}_{\mathrm{h}} + \boldsymbol{y}_{\mathrm{p}}$, where $\boldsymbol{y}_{\mathrm{h}}$ is a solution to the ***associated homogeneous system***

$$\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t) \tag{2}$$

and $\boldsymbol{y}_{\mathrm{p}}$ is a fixed particular solution. We now focus on the solution set to (2).

Theorem 4. *If the $n \times n$ matrix $A(t)$ is continuous on an interval I, then the solution set to (2) is a linear space of dimension n. In other words,*

1. *There are n linearly independent solutions.*
2. *Given any set of n linear independent solutions $\{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_n\}$, then any other solution $\boldsymbol{\phi}$ can be written as*

$$\boldsymbol{\phi} = c_1\boldsymbol{\phi}_1 + \cdots + c_n\boldsymbol{\phi}_n$$

for some scalars $c_1, \ldots, c_n \in \mathbb{R}$.

Proof. Let $\boldsymbol{e}_i$ be the $n \times 1$ matrix with 1 in the ith position and zeros elsewhere. By the existence and uniqueness theorem, there is a unique solution, $\boldsymbol{\psi}_i(t)$, to (2) with initial condition $\boldsymbol{y}(t_0) = \boldsymbol{e}_i$. We claim $\{\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_n\}$ is linearly independent. To show this, let

$$\boldsymbol{\Psi}(t) = \begin{bmatrix} \boldsymbol{\psi}_1(t) & \boldsymbol{\psi}_2(t) & \cdots & \boldsymbol{\psi}_n(t) \end{bmatrix}. \tag{3}$$

Then $\boldsymbol{\Psi}$ is an $n \times n$ matrix of functions with $\boldsymbol{\Psi}(t_0) = \begin{bmatrix} \boldsymbol{\psi}_1(t_0) & \cdots & \boldsymbol{\psi}_n(t_0) \end{bmatrix} = \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} = I$, the $n \times n$ identity matrix. Now suppose there are scalars $c_1, \ldots, c_n$ such that $c_1\boldsymbol{\psi}_1 + \cdots + c_n\boldsymbol{\psi}_n = 0$, valid for all $t \in I$. We can reexpress this as $\boldsymbol{\Psi}(t)\boldsymbol{c} = 0$, where $\boldsymbol{c}$ is the column vector

$$\boldsymbol{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}. \tag{4}$$

Now evaluate at $t = t_0$ to get $\boldsymbol{c} = \boldsymbol{I}\boldsymbol{c} = \Psi(t_0)\boldsymbol{c} = \mathbf{0}$. This implies $\{\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_n\}$ is linearly independent. Now suppose that $\boldsymbol{\psi}$ is any solution to (2). Let $\boldsymbol{\eta}(t) = \boldsymbol{\Psi}(t)\boldsymbol{\psi}(t_0)$. Then $\boldsymbol{\eta}$ is a linear combination of $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_n$, hence a solution, and $\boldsymbol{\eta}(t_0) = \boldsymbol{\Psi}(t_0)\boldsymbol{\psi}(t_0) = \boldsymbol{\psi}(t_0)$. Since $\boldsymbol{\eta}$ and $\boldsymbol{\psi}$ satisfy the same initial condition, they are equal by the existence and uniqueness theorem. It follows that every solution, $\boldsymbol{y}(t)$, to (2) is a linear combination of $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_n$ and may be expressed as

$$\boldsymbol{y}(t) = \boldsymbol{\Psi}(t)\boldsymbol{y}(t_0), \tag{5}$$

where $\boldsymbol{y}(t_0)$ is the initial condition, expressed as a column vector.

Now suppose that $\{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_n\}$ is any set of n linearly independent solutions of (2). We wish now to show that any solution may be expressed as a linear combination of $\{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_n\}$.[8] By (5), we have

$$\boldsymbol{\phi}_i(t) = \boldsymbol{\Psi}(t)\boldsymbol{\phi}_i(t_0),$$

for each $i = 1, \ldots, n$. Let $\Phi(t)$ be the $n \times n$ matrix given by

$$\boldsymbol{\Phi}(t) = \begin{bmatrix} \boldsymbol{\phi}_1(t) & \boldsymbol{\phi}_2(t) & \cdots & \boldsymbol{\phi}_n(t) \end{bmatrix}. \tag{6}$$

Now we can write

$$\boldsymbol{\Phi}(t) = \boldsymbol{\Psi}(t)\boldsymbol{\Phi}(t_0).$$

We claim $\boldsymbol{\Phi}(t_0)$ is invertible. Suppose not. Then there would be a column matrix $\boldsymbol{c}$ as in (4), where $c_1, \ldots, c_n$ are not all zero, such that $\boldsymbol{\Phi}(t_0)c = 0$. But this implies

$$c_1\boldsymbol{\phi}_1(t) + \cdots + c_n\boldsymbol{\phi}_n(t) = \boldsymbol{\Phi}(t)\boldsymbol{c} = \boldsymbol{\Psi}(t)\boldsymbol{\Phi}(t_0)\boldsymbol{c} = \mathbf{0}.$$

This contradicts the linear independence of $\{\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_n\}$. It follows that $\boldsymbol{\Phi}(t_0)$ must be invertible. We can thus write

$$\boldsymbol{\Psi}(t) = \boldsymbol{\Phi}(t)(\boldsymbol{\Phi}(t_0))^{-1}. \tag{7}$$

Now suppose $\boldsymbol{\phi}$ is a solution to (2). Then (5) and (7) give

$$\boldsymbol{\phi}(t) = \boldsymbol{\Psi}(t)\boldsymbol{\phi}(t_0) = \boldsymbol{\Phi}(t)(\boldsymbol{\Phi}(t_0))^{-1}\boldsymbol{\phi}(t_0),$$

which when multiplied out is a linear combination of the $\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_n$. $\square$

[8] This actually follows from a general result in linear algebra.

We will say that an $n \times n$ matrix function $\boldsymbol{\Phi}(t)$ is a ***fundamental matrix*** for a homogeneous system $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t)$ if its columns form a linearly independent set of solutions, as in (6). A matrix function $\boldsymbol{\Psi}(t)$ is the ***standard fundamental matrix at*** $t = t_0$ if it is a fundamental matrix and $\boldsymbol{\Psi}(t_0) = I$, the $n \times n$ identity matrix, as in (3). Given a fundamental matrix $\boldsymbol{\Phi}(t)$ Theorem 4, shows that the solution set to the homogeneous system $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t)$ is the span[9] of the columns of $\boldsymbol{\Phi}(t)$.

Theorem 5. *Suppose $A(t)$ is an $n \times n$ continuous matrix function on an interval I. A matrix function $\boldsymbol{\Phi}(t)$ is a fundamental matrix for $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t)$ if and only if*

$$\boldsymbol{\Phi}'(t) = A(t)\boldsymbol{\Phi}(t) \quad \textit{and} \quad \det \boldsymbol{\Phi}(t) \neq 0,$$

for at least one $t \in I$. If this is true for one $t \in I$, it is in fact true for all $t \in I$. The standard fundamental matrix $\boldsymbol{\Psi}(t)$ at $t = t_0$ is uniquely characterized by the equations

$$\boldsymbol{\Psi}'(t) = A(t)\boldsymbol{\Psi}(t) \quad \textit{and} \quad \boldsymbol{\Psi}(t_0) = I.$$

Furthermore, given a fundamental matrix $\boldsymbol{\Phi}(t)$, the standard fundamental matrix $\boldsymbol{\Psi}(t)$ at $t = t_0$ is given by the formula

$$\boldsymbol{\Psi}(t) = \boldsymbol{\Phi}(t)\left(\boldsymbol{\Phi}(t_0)\right)^{-1}.$$

Proof. Suppose $\Phi(t) = \begin{bmatrix} \boldsymbol{\phi}_1 & \cdots & \boldsymbol{\phi}_n \end{bmatrix}$ is a fundamental matrix for $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t)$. Then

$$\begin{aligned}
\boldsymbol{\Phi}'(t) &= \begin{bmatrix} \boldsymbol{\phi}_1'(t) & \cdots & \boldsymbol{\phi}_n'(t) \end{bmatrix} \\
&= \begin{bmatrix} A(t)\boldsymbol{\phi}_1(t) & \cdots & A(t)\boldsymbol{\phi}_n(t) \end{bmatrix} \\
&= A(t)\begin{bmatrix} \boldsymbol{\phi}_1(t) & \cdots & \boldsymbol{\phi}_n(t) \end{bmatrix} \\
&= A(t)\boldsymbol{\Phi}(t).
\end{aligned}$$

As in the proof above, $\boldsymbol{\Phi}(t_0)$ is invertible which implies $\det \boldsymbol{\Phi}(t_0) \neq 0$. Since $t_0 \in I$ is arbitrary, it follows that $\Phi(t)$ has nonzero determinant for all $t \in \mathcal{I}$. Now suppose that $\Phi(t)$ is a matrix function satisfying $\boldsymbol{\Phi}'(t) = A(t)\boldsymbol{\Phi}(t)$ and $\det \boldsymbol{\Phi}(t) \neq 0$, for some point, $t = t_0$ say. Then the above calculation gives that each column $\boldsymbol{\phi}_i(t)$ of $\boldsymbol{\Phi}(t)$ satisfies $\boldsymbol{\phi}_i'(t) = A(t)\boldsymbol{\phi}(t)$. Suppose there are scalars $c_1, \ldots, c_n$ such that $c_1\boldsymbol{\phi}_1 + \cdots + c_n\boldsymbol{\phi}_n = 0$ as a function on I. If $\boldsymbol{c}$ is the column vector given as in (4), then

$$\boldsymbol{\Phi}(t_0)\boldsymbol{c} = c_1\boldsymbol{\phi}_1(t_0) + \cdots + c_n\boldsymbol{\phi}_n(t_0) = \boldsymbol{0}.$$

[9]Recall that "span" means the set of all linear combinations.

Since $\det \boldsymbol{\Phi}(t_0) \neq 0$, it follows that $\boldsymbol{\Phi}(t_0)$ is invertible. Therefore, $\boldsymbol{c} = 0$ and $\boldsymbol{\Phi}(t)$ is a fundamental matrix. Now suppose $\boldsymbol{\Psi}_1$ and $\boldsymbol{\Psi}_2$ are standard fundamental matrices at t_0. As fundamental matrices, their ith columns both are solutions to $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t)$ and evaluate to e_i at t_0. By the existence and uniqueness theorem, they are equal. It follows that the standard fundamental matrix is unique. Finally, (7) gives the final statement $\boldsymbol{\Psi}(t) = \boldsymbol{\Phi}(t)\,(\boldsymbol{\Phi}(t_0))^{-1}$. □

In the case $A(t) = A$ is a constant, the matrix exponential e^{At} is the standard fundamental matrix at $t = 0$ for the system $\boldsymbol{y}' = A\boldsymbol{y}$. This follows from Lemma 9.5.1. More generally, $\mathrm{e}^{A(t-t_0)}$ is the standard fundamental matrix at $t = t_0$.

Example 6. Show that

$$\boldsymbol{\Phi}(t) = \begin{bmatrix} \mathrm{e}^{2t} & \mathrm{e}^{-t} \\ 2\mathrm{e}^{2t} & -\mathrm{e}^{-t} \end{bmatrix}$$

is a fundamental matrix for the system $\boldsymbol{y}' = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}\boldsymbol{y}$. Find the standard fundamental matrix at $t = 0$.

► **Solution.** We first observe that

$$\boldsymbol{\Phi}'(t) = \begin{bmatrix} \mathrm{e}^{2t} & \mathrm{e}^{-t} \\ 2\mathrm{e}^{2t} & -\mathrm{e}^{-t} \end{bmatrix}' = \begin{bmatrix} 2\mathrm{e}^{2t} & -\mathrm{e}^{-t} \\ 4\mathrm{e}^{2t} & \mathrm{e}^{-t} \end{bmatrix}$$

and

$$A\boldsymbol{\Phi}(t) = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} \mathrm{e}^{2t} & \mathrm{e}^{-t} \\ 2\mathrm{e}^{2t} & -\mathrm{e}^{-t} \end{bmatrix} = \begin{bmatrix} 2\mathrm{e}^{2t} & -\mathrm{e}^{-t} \\ 4\mathrm{e}^{2t} & \mathrm{e}^{-t} \end{bmatrix}.$$

Thus, $\boldsymbol{\Phi}'(t) = A\boldsymbol{\Phi}(t)$. Observe also that

$$\boldsymbol{\Phi}(0) = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$$

and this matrix has determinant -3. By Theorem 5, $\boldsymbol{\Phi}(t)$ is a fundamental matrix. The standard fundamental matrix at $t = 0$ is given by

$$\begin{aligned}
\boldsymbol{\Psi}(t) &= \boldsymbol{\Phi}(t)\,(\boldsymbol{\Phi}(0))^{-1} \\
&= \begin{bmatrix} \mathrm{e}^{2t} & \mathrm{e}^{-t} \\ 2\mathrm{e}^{2t} & -\mathrm{e}^{-t} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}^{-1} \\
&= -\frac{1}{3}\begin{bmatrix} \mathrm{e}^{2t} & \mathrm{e}^{-t} \\ 2\mathrm{e}^{2t} & -\mathrm{e}^{-t} \end{bmatrix}\begin{bmatrix} -1 & -1 \\ -2 & 1 \end{bmatrix} \\
&= \frac{1}{3}\begin{bmatrix} \mathrm{e}^{2t} + 2\mathrm{e}^{-t} & \mathrm{e}^{2t} - \mathrm{e}^{-t} \\ 2\mathrm{e}^{2t} - 2\mathrm{e}^{-t} & 2\mathrm{e}^{2t} + \mathrm{e}^{-t} \end{bmatrix}.
\end{aligned}$$

The reader is encouraged to compute e^{At} where $A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$ and verify that $\boldsymbol{\Psi}(t) = e^{At}$. ◄

Example 7. Show that

$$\boldsymbol{\Phi}(t) = \begin{bmatrix} t^2 & t^3 \\ 2t & 3t^2 \end{bmatrix}$$

is a fundamental matrix for the system $\boldsymbol{y}' = A(t)\boldsymbol{y}$ where

$$A(t) = \begin{bmatrix} 0 & 1 \\ -6/t^2 & 4/t \end{bmatrix}.$$

Solve the system with initial condition $\boldsymbol{y}(1) = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$. Find the standard fundamental matrix at $t = 1$.

► **Solution.** Note that

$$\boldsymbol{\Phi}'(t) = \begin{bmatrix} 2t & 3t^2 \\ 2 & 6t \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -6/t^2 & 4/t \end{bmatrix} \begin{bmatrix} t^2 & t^3 \\ 2t & 3t^2 \end{bmatrix} = A(t)\boldsymbol{\Phi}(t),$$

while

$$\det \boldsymbol{\Phi}(1) = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} = 1 \neq 0.$$

Hence, $\boldsymbol{\Phi}(t)$ is a fundamental matrix. The general solution is of the form $\boldsymbol{y}(t) = \boldsymbol{\Phi}(t)\boldsymbol{c}$, where $\boldsymbol{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$. The initial condition implies

$$\begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_1 + c_2 \\ 2c_1 + 3c_2 \end{bmatrix}.$$

Solving for $\boldsymbol{c}$ gives $c_1 = 2$ and $c_2 = 1$. Thus,

$$\boldsymbol{y}(t) = 2\begin{bmatrix} t^2 \\ 2t \end{bmatrix} + \begin{bmatrix} t^3 \\ 3t^2 \end{bmatrix} = \begin{bmatrix} 2t^2 + t^3 \\ 4t + 3t^2 \end{bmatrix}.$$

The standard fundamental matrix is given by

$$\boldsymbol{\Psi}(t) = \boldsymbol{\Phi}(t)\,(\boldsymbol{\Phi}(1))^{-1} = \begin{bmatrix} t^2 & t^3 \\ 2t & 3t^2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} t^2 & t^3 \\ 2t & 3t^2 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3t^2 - 2t^3 & -t^2 + t^3 \\ 6t - 6t^2 & -2t + 3t^2 \end{bmatrix}.$$

Observe that $\boldsymbol{\Phi}(0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ which has determinant 0. Why does this not prevent $\boldsymbol{\Phi}(t)$ from being a fundamental matrix? ◀

You will recall that in Sect. 9.5, we used the inverse of the matrix exponential $\left(e^{tA}\right)^{-1} = e^{-tA}$ as an integrating factor for the constant coefficient system $\boldsymbol{y}' = A\boldsymbol{y} + \boldsymbol{f}$. As observed above, the matrix exponential e^{tA} is the standard fundamental matrix at $t = 0$. In the more general context, we will show that the inverse of any fundamental matrix $\boldsymbol{\Phi}(t)$ is an integrating factor for the system $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t) + \boldsymbol{f}(t)$. To show this, however, particular care must be taken when calculating the derivative of the inverse of a matrix-valued function.

Lemma 8. *Suppose $\boldsymbol{\Phi}(t)$ is a differentiable and invertible $n \times n$ matrix-valued function. Then*

$$\frac{\mathrm{d}}{\mathrm{dt}}(\boldsymbol{\Phi}(t))^{-1} = -(\boldsymbol{\Phi}(t))^{-1}\boldsymbol{\Phi}'(t)(\boldsymbol{\Phi}(t))^{-1}.$$

Remark 9. Observe the order of the matrix multiplications. We are not assuming that $\boldsymbol{\Phi}(t)$ and its derivative $\boldsymbol{\Phi}'(t)$ commute.

Proof. We apply the definition of the derivative:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{dt}}(\boldsymbol{\Phi}(t))^{-1} &= \lim_{h\to 0} \frac{(\boldsymbol{\Phi}(t+h))^{-1} - (\boldsymbol{\Phi}(t))^{-1}}{h} \\ &= \lim_{h\to 0} (\boldsymbol{\Phi}(t+h))^{-1} \frac{\boldsymbol{\Phi}(t) - \boldsymbol{\Phi}(t+h)}{h} (\boldsymbol{\Phi}(t))^{-1} \\ &= -\,(\boldsymbol{\Phi}(t+h))^{-1} \lim_{h\to 0} \frac{\boldsymbol{\Phi}(t+h) - \boldsymbol{\Phi}(t)}{h} (\boldsymbol{\Phi}(t))^{-1} \\ &= -(\boldsymbol{\Phi}(t))^{-1}\boldsymbol{\Phi}'(t)(\boldsymbol{\Phi}(t))^{-1}. \end{aligned}$$

The second line is just a careful factoring of the first line. To verify this step, simply multiply out the second line. □

If $\boldsymbol{\Phi}(t)$ is a scalar-valued function, then Lemma 8 reduces to the usual chain rule formula, $\frac{\mathrm{d}}{\mathrm{dt}}(\boldsymbol{\Phi}(t))^{-1} = -(\boldsymbol{\Phi}(t))^{-2}\boldsymbol{\Phi}'(t)$, since Φ and Φ' commute. For a matrix-valued function, we may not assume that Φ and Φ' commute.

Nonhomogeneous Linear Systems

We are now in a position to consider the solution method for the general linear system $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t) + \boldsymbol{f}(t)$ which we write in the form

$$\boldsymbol{y}'(t) - A(t)\boldsymbol{y}(t) = \boldsymbol{f}(t). \tag{8}$$

Assume $A(t)$ and $\boldsymbol{f}(t)$ are continuous on an interval I and an initial condition $\boldsymbol{y}(t_0) = \boldsymbol{y}_0$ is given. Suppose $\boldsymbol{\Phi}(t)$ is a fundamental matrix for the associated homogeneous system $\boldsymbol{y}'(t) = A(t)y(t)$. Then $(\boldsymbol{\Phi}(t))^{-1}$ will play the role of an integrating factor and we will be able mimic the procedure found in Sect. 1.4. First observe from Lemma 8 and the product rule that we get

$$\begin{aligned}\left((\boldsymbol{\Phi}(t))^{-1}\boldsymbol{y}(t)\right)' &= (\boldsymbol{\Phi}(t))^{-1}\boldsymbol{y}'(t) - (\boldsymbol{\Phi}(t))^{-1}\boldsymbol{\Phi}'(t)(\boldsymbol{\Phi}(t))^{-1}\boldsymbol{y}(t)\\ &= (\boldsymbol{\Phi}(t))^{-1}\boldsymbol{y}'(t) - (\boldsymbol{\Phi}(t))^{-1}A(t)\boldsymbol{\Phi}(t)(\boldsymbol{\Phi}(t))^{-1}\boldsymbol{y}(t)\\ &= (\boldsymbol{\Phi}(t))^{-1}\boldsymbol{y}'(t) - (\boldsymbol{\Phi}(t))^{-1}A(t)\boldsymbol{y}(t)\\ &= (\boldsymbol{\Phi}(t))^{-1}(\boldsymbol{y}'(t) - A(t)\boldsymbol{y}(t)).\end{aligned}$$

Thus, multiplying both sides of (8) by $(\boldsymbol{\Phi}(t))^{-1}$ gives

$$\left((\boldsymbol{\Phi}(t))^{-1}\boldsymbol{y}(t)\right)' = (\boldsymbol{\Phi}(t))^{-1}\boldsymbol{f}(t).$$

Now change the variable from t to u and integrate both sides from t_0 to t where $t \in I$. We get

$$\int_{t_0}^{t} \left((\boldsymbol{\Phi}(u))^{-1}\boldsymbol{y}(u)\right)' \,\mathrm{d}u = \int_{t_0}^{t} (\boldsymbol{\Phi}(u))^{-1}\boldsymbol{f}(u)\,\mathrm{d}u.$$

The left side simplifies to $\left(\boldsymbol{\Phi}(t)^{-1}\boldsymbol{y}(t)\right) - \left(\boldsymbol{\Phi}(t_0)^{-1}\boldsymbol{y}(t_0)\right)$. Solving for $\boldsymbol{y}(t)$, we get

$$\boldsymbol{y}(t) = \boldsymbol{\Phi}(t)(\boldsymbol{\Phi}(t_0))^{-1}y_0 + \boldsymbol{\Phi}(t)\int_{t_0}^{t} (\boldsymbol{\Phi}(u))^{-1}\boldsymbol{f}(u)\,\mathrm{d}u.$$

It is convenient to summarize this discussion in the following theorem.

Theorem 10. *Suppose $A(t)$ is an $n \times n$ matrix-valued function on an interval I and $\boldsymbol{f}(t)$ is a $\mathbb{R}^n$-valued valued function, both continuous on an interval I. Let $\boldsymbol{\Phi}(t)$ be any fundamental matrix for $\boldsymbol{y}'(t) = A(t)y(t)$ and let $t_0 \in I$. Then*

$$\boldsymbol{y}(t) = \boldsymbol{\Phi}(t)(\boldsymbol{\Phi}(t_0))^{-1}y_0 + \boldsymbol{\Phi}(t)\int_{t_0}^{t} (\boldsymbol{\Phi}(u))^{-1}\boldsymbol{f}(u)\,\mathrm{d}u, \tag{9}$$

is the unique solution to

$$\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t) + \boldsymbol{f}(t), \quad \boldsymbol{y}(t_0) = \boldsymbol{y}_0.$$

Analysis of the General Solution Set

The Particular Solution

When we set $\boldsymbol{y}_0 = 0$, we obtain a single fixed solution which we denote by $\boldsymbol{y}_{\mathrm{p}}$. Specifically,

$$\boldsymbol{y}_{\mathrm{p}}(t) = \boldsymbol{\Phi}(t) \int_{t_0}^{t} (\boldsymbol{\Phi}(u))^{-1} \boldsymbol{f}(u)\, \mathrm{d}u. \tag{10}$$

This is called a ***particular solution***.

The Homogeneous Solution

When we set $\boldsymbol{f} = 0$, we get the ***homogeneous solution*** $\boldsymbol{y}_{\mathrm{h}}$. Specifically,

$$\boldsymbol{y}_{\mathrm{h}}(t) = \boldsymbol{\Phi}(t)(\boldsymbol{\Phi}(t_0))^{-1}\boldsymbol{y}_0. \tag{11}$$

Recall from Theorem 5 that the standard fundamental matrix at t_0 is $\boldsymbol{\Psi}(t) = \boldsymbol{\Phi}(t)(\boldsymbol{\Phi}(t_0))^{-1}$. Thus, the homogeneous solution with initial value $\boldsymbol{y}_{\mathrm{h}}(t_0) = \boldsymbol{y}_0$ is given by

$$\boldsymbol{y}_{\mathrm{h}}(t) = \boldsymbol{\Psi}(t)\boldsymbol{y}_0.$$

On the other hand, if we are interested in the set of homogeneous solutions, then we let y_0 vary. However, Theorem 5 states the $\boldsymbol{\Phi}(t_0)$ is invertible so as $\boldsymbol{y}_0$ varies over $\mathbb{R}^n$, so does $(\boldsymbol{\Phi}(t_0))^{-1}\boldsymbol{y}_0$. Thus, if we let $\boldsymbol{c} = (\boldsymbol{\Phi}(t_0))^{-1}y_0$, then we have that the set of homogeneous solution is

$$\left\{\boldsymbol{\Phi}(t)\boldsymbol{c} : \boldsymbol{c} \in \mathbb{R}^n\right\}.$$

In other words, the set of homogeneous solution is the set of all linear combinations of the columns of $\boldsymbol{\Phi}(t)$, as we observed earlier.

Example 11. Solve the system $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t) + \boldsymbol{f}(t)$ where

$$A(t) = \begin{bmatrix} 0 & 1 \\ -6/t^2 & 4/t \end{bmatrix} \quad \text{and} \quad \boldsymbol{f}(t) = \begin{bmatrix} t^2 \\ t \end{bmatrix}$$

► **Solution.** In Example 7, we verified that

$$\boldsymbol{\Phi}(t) = \begin{bmatrix} t^2 & t^3 \\ 2t & 3t^2 \end{bmatrix}$$

is a fundamental matrix. It follows that the homogeneous solutions are of the form

$$\boldsymbol{y}_{\mathrm{h}}(t) = c_1 \begin{bmatrix} t^2 \\ 2t \end{bmatrix} + c_2 \begin{bmatrix} t^3 \\ 3t^2 \end{bmatrix},$$

where c_1 and c_2 are real scalars. To find the particular solution, it is convenient to set $t_0 = 1$ and use (10). We get

$$\begin{aligned}
\boldsymbol{y}_{\mathrm{p}}(t) &= \boldsymbol{\Phi}(t) \int_1^t (\boldsymbol{\Phi}(u))^{-1} \boldsymbol{f}(u)\, \mathrm{d}u \\
&= \boldsymbol{\Phi}(t) \int_1^t \begin{bmatrix} 3/u^2 & -1/u \\ -2/u^3 & 1/u^2 \end{bmatrix} \begin{bmatrix} u^2 \\ u \end{bmatrix} \mathrm{d}u \\
&= \boldsymbol{\Phi}(t) \int_1^t \begin{bmatrix} 2 \\ -1/u \end{bmatrix} \mathrm{d}u \\
&= \begin{bmatrix} t^2 & t^3 \\ 2t & 3t^2 \end{bmatrix} \begin{bmatrix} 2(t-1) \\ -\ln t \end{bmatrix} \\
&= \begin{bmatrix} 2t^3 - 2t^2 - t^3 \ln t \\ 4t^2 - 4t - 3t^2 \ln t \end{bmatrix} = -2 \begin{bmatrix} t^2 \\ 2t \end{bmatrix} + 2 \begin{bmatrix} t^3 \\ 3t^2 \end{bmatrix} + \begin{bmatrix} -t^3 \ln t \\ -2t^2 - 3t^2 \ln t \end{bmatrix}.
\end{aligned}$$

It follows that the general solution may be written as

$$\begin{aligned}
\boldsymbol{y}(t) &= \boldsymbol{y}_{\mathrm{h}}(t) + \boldsymbol{y}_{\mathrm{p}}(t) \\
&= (c_1 - 2) \begin{bmatrix} t^2 \\ 2t \end{bmatrix} + (c_2 + 2) \begin{bmatrix} t^3 \\ 3t^2 \end{bmatrix} + \begin{bmatrix} -t^3 \ln t \\ -2t^2 - 3t^2 \ln t \end{bmatrix} \\
&= C_1 \begin{bmatrix} t^2 \\ 2t \end{bmatrix} + C_2 \begin{bmatrix} t^3 \\ 3t^2 \end{bmatrix} + \begin{bmatrix} -t^3 \ln t \\ -2t^2 - 3t^2 \ln t \end{bmatrix},
\end{aligned}$$

where the last line is just a relabeling of the coefficients. In Example (7), we computed the standard fundamental matrix $\boldsymbol{\Psi}(t)$. It could have been used in place of $\boldsymbol{\Phi}(t)$ in the above computations. However, the simplicity of $\boldsymbol{\Phi}(t)$ made it a better choice. ◄

This example emphasized the need to have a fundamental matrix in order to go forward with the calculations of the particular and homogeneous solutions, (10) and (11), respectively. Computing a fundamental matrix is not always an easy task. In order to find a closed expression for it, we must place strong restrictions on the coefficient matrix $A(t)$. Let us consider one such restriction.

The Coefficient Matrix is a Functional Multiple of a Constant Matrix

Proposition 12. *Suppose $a(t)$ is a continuous function on an interval I. Let*

$$A(t) = a(t)A,$$

where A is a fixed $n \times n$ constant matrix. Then a fundamental matrix, $\boldsymbol{\Phi}(t)$, for $\boldsymbol{y}'(t) = A(t)y(t)$ is given by the formula

$$\boldsymbol{\Phi}(t) = \mathrm{e}^{b(t)A}, \tag{12}$$

where $b(t)$ is an antiderivative of $a(t)$. If $b(t) = \int_{t_0}^{t} a(u)\,\mathrm{d}u$, that is, $b(t)$ is chosen so that $b(t_0) = 0$, then (12) is the standard fundamental matrix at t_0.

Proof. Let

$$\boldsymbol{\Phi}(t) = \mathrm{e}^{b(t)A} = I + b(t)A + \frac{b^2(t)A^2}{2!} + \cdots .$$

Termwise differentiation gives

$$\begin{aligned}
\boldsymbol{\Phi}'(t) &= b'(t)A + 2b(t)b'(t)\frac{A^2}{2!} + 3b^2(t)b'(t)\frac{A^3}{3!} + \cdots \\
&= a(t)A\left(I + b(t)A + \frac{b^2(t)A^2}{2!} + \cdots\right) \\
&= A(t)\boldsymbol{\Phi}(t).
\end{aligned}$$

Since the matrix exponential is always invertible, it follows from Theorem 5 that $\boldsymbol{\Phi}(t)$ is a fundamental matrix. If $b(t)$ is chosen so that $b(t_0) = 0$, then $\boldsymbol{\Phi}(t_0) = \mathrm{e}^{Au}|_{u=0} = I$, and hence, $\boldsymbol{\Phi}(t)$ is the standard fundamental matrix at t_0. □

Remark 13. We observe that $\boldsymbol{\Phi}(t)$ may be computed by replacing u in e^{uA} with $b(t)$.

Example 14. Suppose $A(t) = (\tan t)A$, where

$$A = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}.$$

Find a fundamental matrix for $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t)$. Assume $t \in (-\frac{\pi}{2}, \frac{\pi}{2})$.

► **Solution.** We first compute e^{Au}. The characteristic polynomial is

$$c_A(s) = \det(sI - A) = \det \begin{bmatrix} s-2 & -3 \\ 1 & s+2 \end{bmatrix} = s^2 - 1 = (s-1)(s+1).$$

It follows that $\mathcal{B}_{c_A} = \{e^u, e^{-u}\}$. Using Fulmer's method, we have $e^{Au} = e^u M_1 + e^{-u} M_2$. Differentiating and evaluating at $u = 0$ gives

$$\begin{aligned} I &= M_1 + M_2, \\ A &= M_1 - M_2. \end{aligned}$$

It follows that $M_1 = \frac{1}{2}(A + I) = \frac{1}{2}\begin{bmatrix} 3 & 3 \\ -1 & -1 \end{bmatrix}$ and $M_2 = \frac{1}{2}(I - A) = \frac{1}{2}\begin{bmatrix} -1 & -3 \\ 1 & 3 \end{bmatrix}$. Thus,

$$e^{Au} = \frac{1}{2}e^u \begin{bmatrix} 3 & 3 \\ -1 & -1 \end{bmatrix} + e^{-u}\begin{bmatrix} -1 & -3 \\ 1 & 3 \end{bmatrix}.$$

Since $\ln \sec t$ is an antiderivative of $\tan t$, we have by Proposition 12

$$\begin{aligned} \boldsymbol{\Phi}(t) &= \frac{1}{2}e^{\ln \sec t}\begin{bmatrix} 3 & 3 \\ -1 & -1 \end{bmatrix} + \frac{1}{2}e^{-\ln \sec t}\begin{bmatrix} -1 & -3 \\ 1 & 3 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 3\sec t - \cos t & 3\sec t - 3\cos t \\ -\sec t + \cos t & -\sec t + 3\cos t \end{bmatrix}, \end{aligned}$$

is the fundamental matrix for $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t)$. ◄

Example 15. Solve the mixing problem introduced at the beginning of this section. Specifically, solve

$$\boldsymbol{y}' = A(t)\boldsymbol{y}(t) + \boldsymbol{f}(t), \quad \boldsymbol{y}(0) = \boldsymbol{y}_0,$$

where

$$A(t) = \begin{bmatrix} -\frac{2}{1+t} & \frac{1}{1+t} \\ \frac{1}{1+t} & -\frac{2}{1+t} \end{bmatrix}, \quad \boldsymbol{f}(t) = \begin{bmatrix} 12 \\ 20 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{y}(0) = \begin{bmatrix} 8 \\ 0 \end{bmatrix}.$$

Determine the concentration of salt in each tank after 3 minutes. In the long term, what are the concentrations of salt in each tank?

► **Solution.** Let $a(t) = \frac{1}{t+1}$. Then $A(t) = a(t)A$ where $A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$. The characteristic polynomial of A is

$$c_A(s) = \det(sI - A) = \det \begin{bmatrix} s+2 & -1 \\ -1 & s+2 \end{bmatrix} = (s+1)(s+3).$$

A short calculation gives the resolvent matrix

$$\begin{aligned} (sI - A)^{-1} &= \frac{1}{(s+1)(s+3)} \begin{bmatrix} s+2 & 1 \\ 1 & s+2 \end{bmatrix} \\ &= \frac{1}{2(s+1)} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{1}{2(s+3)} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \end{aligned}$$

and hence

$$e^{Au} = \frac{e^{-u}}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{e^{-3u}}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Since $\ln(t+1)$ is an antiderivative of $a(t) = \frac{1}{t+1}$, we have by Proposition 12

$$\begin{aligned} \boldsymbol{\Phi}(t) &= e^{\ln(t+1)A} \\ &= \frac{e^{-\ln(t+1)}}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{e^{-3(\ln(t+1))}}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{2(t+1)} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{1}{2(t+1)^3} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{2(t+1)^3} \begin{bmatrix} (t+1)^2+1 & (t+1)^2-1 \\ (t+1)^2-1 & (t+1)^2+1 \end{bmatrix} \end{aligned}$$

a fundamental matrix for $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t)$. Observe that $\boldsymbol{\Phi}(0) = I$ is the 2×2 identity matrix so that, in fact, $\boldsymbol{\Phi}(t)$ is the standard fundamental matrix at $t = 0$. The homogeneous solution is now easily calculated:

$$\begin{aligned} \boldsymbol{y}_{\mathrm{h}}(t) &= \boldsymbol{\Phi}(t)\boldsymbol{y}_0 \\ &= \frac{1}{2(t+1)^3} \begin{bmatrix} (t+1)^2+1 & (t+1)^2-1 \\ (t+1)^2-1 & (t+1)^2+1 \end{bmatrix} \begin{bmatrix} 8 \\ 0 \end{bmatrix} \\ &= \frac{4}{(t+1)^3} \begin{bmatrix} (t+1)^2+1 \\ (t+1)^2-1 \end{bmatrix}. \end{aligned}$$

A straightforward calculation gives

$$(\boldsymbol{\Phi}(u))^{-1} = \frac{u+1}{2} \begin{bmatrix} (u+1)^2+1 & 1-(u+1)^2 \\ 1-(u+1)^2 & (u+1)^2+1 \end{bmatrix}$$

and

$$(\boldsymbol{\Phi}(u))^{-1}\boldsymbol{f}(u) = \frac{u+1}{2}\begin{bmatrix}(u+1)^2+1 & 1-(u+1)^2\\ 1-(u+1)^2 & (u+1)^2+1\end{bmatrix}\begin{bmatrix}12\\20\end{bmatrix}$$

$$= \begin{bmatrix}16(u+1)-4(u+1)^3\\ 16(u+1)+4(u+1)^3\end{bmatrix}.$$

For the particular solution, we have

$$\begin{aligned}
\boldsymbol{y}_{\mathrm{p}}(t) &= \Phi(t)\int_0^t (\boldsymbol{\Phi}(u))^{-1}\boldsymbol{f}(u)\,\mathrm{d}u\\
&= \boldsymbol{\Phi}(t)\begin{bmatrix}8(u+1)^2-(u+1)^4\\ 8(u+1)^2+(u+1)^4\end{bmatrix}\Bigg|_0^t\\
&= \frac{1}{2(t+1)^3}\begin{bmatrix}(t+1)^2+1 & (t+1)^2-1\\ (t+1)^2-1 & (t+1)^2+1\end{bmatrix}\begin{bmatrix}8(t+1)^2-(t+1)^4-7\\ 8(t+1)^2+(t+1)^4-9\end{bmatrix}\\
&= \frac{1}{(t+1)^3}\begin{bmatrix}7(t+1)^4-8(t+1)^2+1\\ 9(t+1)^4-8(t+1)^2-1\end{bmatrix}.
\end{aligned}$$

Putting the homogeneous and particular solutions together gives

$$\begin{aligned}
\boldsymbol{y}(t) &= \boldsymbol{y}_{\mathrm{h}}(t)+\boldsymbol{y}_{\mathrm{p}}(t)\\
&= \frac{4}{(t+1)^3}\begin{bmatrix}(t+1)^2+1\\ (t+1)^2-1\end{bmatrix}+\frac{1}{(t+1)^3}\begin{bmatrix}7(t+1)^4-8(t+1)^2+1\\ 9(t+1)^4-8(t+1)^2-1\end{bmatrix}\\
&= \frac{1}{(t+1)^3}\begin{bmatrix}7(t+1)^4-4(t+1)^2+5\\ 9(t+1)^4-4(t+1)^2-5\end{bmatrix}.
\end{aligned}$$

Finally, the amount of fluid in each tank is $v_1(t) = v_2(t) = t+1$. Thus, the concentration of salt in each tank is given by

$$\frac{1}{t+1}\boldsymbol{y}(t) = \frac{1}{(t+1)^4}\begin{bmatrix}7(t+1)^4-4(t+1)^2+5\\ 9(t+1)^4-4(t+1)^2-5\end{bmatrix}.$$

Evaluating at $t=3$ gives concentrations

$$\begin{bmatrix}\frac{1733}{256}\\ \frac{2235}{256}\end{bmatrix} = \begin{bmatrix}6.77\\ 8.73\end{bmatrix}.$$

In the long term, the concentrations are obtained by taking limits

$$\lim_{t\to\infty} \frac{1}{(t+1)^4} \begin{bmatrix} 7(t+1)^4 - 4(t+1)^2 + 5 \\ 9(t+1)^4 - 4(t+1)^2 - 5 \end{bmatrix} = \begin{bmatrix} 7 \\ 9 \end{bmatrix}.$$

Of course, the tank overflows in the long term, but for a sufficiently large tank, we can expect that the concentrations in each tank will be near 7g/L and 9g/L, respectively, for large values of t. ◀

Exercises

1–7. For each of the following pairs of matrix functions $\boldsymbol{\Phi}(t)$ and $A(t)$, verify that $\boldsymbol{\Phi}(t)$ is a fundamental matrix for the system $\boldsymbol{y}' = A(t)\boldsymbol{y}$. Use this fact to solve $\boldsymbol{y}' = A(t)\boldsymbol{y}$ with the given initial conditions $\boldsymbol{y}(t_0) = \boldsymbol{y}_0$. Next, determine the standard fundamental matrix at t_0.

1. $\boldsymbol{\Phi}(t) = \begin{bmatrix} e^{-t} & e^{2t} \\ e^{-t} & 4e^{2t} \end{bmatrix}$, $A(t) = \begin{bmatrix} -2 & 1 \\ -4 & 3 \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

2. $\boldsymbol{\Phi}(t) = \begin{bmatrix} e^{2t} & 3e^{3t} \\ e^{2t} & 2e^{3t} \end{bmatrix}$, $A(t) = \begin{bmatrix} 5 & -3 \\ 2 & 0 \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

3. $\boldsymbol{\Phi}(t) = \begin{bmatrix} \sin(t^2/2) & \cos(t^2/2) \\ \cos(t^2/2) & -\sin(t^2/2) \end{bmatrix}$, $A(t) = \begin{bmatrix} 0 & t \\ -t & 0 \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

4. $\boldsymbol{\Phi}(t) = \begin{bmatrix} 1+t^2 & 3+t^2 \\ 1-t^2 & -1-t^2 \end{bmatrix}$, $A(t) = \begin{bmatrix} t & t \\ -t & -t \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$

5. $\boldsymbol{\Phi}(t) = \begin{bmatrix} -t\cos t & -t\sin t \\ t\sin t & -t\cos t \end{bmatrix}$, $A(t) = \begin{bmatrix} 1/t & 1 \\ -1 & 1/t \end{bmatrix}$, $\boldsymbol{y}(\pi) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

6. $\boldsymbol{\Phi}(t) = e^{t}\begin{bmatrix} 1 & t^2 \\ -1 & 1-t^2 \end{bmatrix}$, $A(t) = \begin{bmatrix} 1+2t & 2t \\ -2t & 1-2t \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

7. $\boldsymbol{\Phi}(t) = \begin{bmatrix} 1 & (t-1)e^{t} \\ -1 & e^{t} \end{bmatrix}$, $A(t) = \begin{bmatrix} 1 & 1 \\ 1/t & 1/t \end{bmatrix}$, $\boldsymbol{y}(1) = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$

8–12. For each problem below, compute the standard fundamental matrix for the system $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t)$ at the point given in the initial value. Then solve the initial value problem $\boldsymbol{y}'(t) = A(t)\boldsymbol{y}(t) + \boldsymbol{f}(t)$, $\boldsymbol{y}(t_0) = \boldsymbol{y}_0$.

8. $A(t) = \dfrac{1}{t}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} 1 \\ t \end{bmatrix}$, $\boldsymbol{y}(1) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

9. $A(t) = \dfrac{1}{t}\begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\boldsymbol{y}(1) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

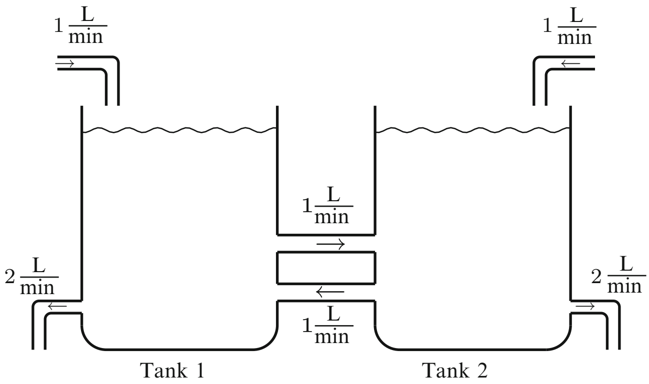
10. $A(t) = \dfrac{2t}{t^2+1}\begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} -3t \\ t \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

11. $A(t) = \begin{bmatrix} 3\sec t & 5\sec t \\ -\sec t & -3\sec t \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

12. $A(t) = \begin{bmatrix} t & t \\ -t & -t \end{bmatrix}$, $\boldsymbol{f}(t) = \begin{bmatrix} 4t \\ 4t \end{bmatrix}$, $\boldsymbol{y}(0) = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$

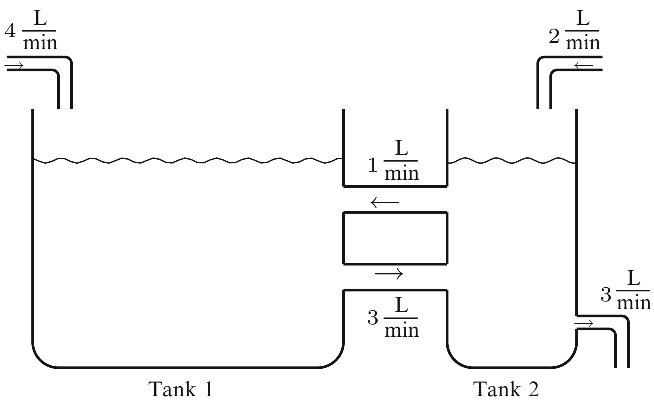
13–14. Solve the following mixing problems.

13. Two tanks are interconnected as illustrated below.



Assume that Tank 1 contains 2 Liters of pure water and Tank 2 contains 2 Liters of brine in which 20 grams of salt is initially dissolved. Moreover, the mixture is pumped from each tank to the other at a rate of 1 L/min. Assume that a brine mixture containing 6 grams salt/L enters Tank 1 at a rate of 1 L/min and pure water enters Tank 2 at a rate of 1 L/min. Assume the tanks are well stirred. Brine is removed from Tank 1 at the rate 2 L/min and from Tank 2 at a rate of 2 L/min. Let $y_1(t)$ be the amount of salt in Tank 1 at time t and let $y_2(t)$ be the amount of salt in Tank 2 at time t. Determine y_1 and y_2. What is the concentration of salt in each tank after 1 minute?

14. Two tanks are interconnected as illustrated below.



Assume initially that Tank 1 contains 2 Liters of pure water and Tank 2 contains 1 Liter of pure water. Moreover, the mixture from Tank 1 is pumped into Tank 2 at a rate of 3 L/min and the mixture from Tank 2 is pumped into Tank 1 at a rate of 1 L/min. Assume that a brine mixture containing 7 grams salt/L enters Tank 1 at a rate of 4 L/min and a brine mixture containing 14 grams salt/L enters Tank 2 at a rate of 2 L/min. Assume the tanks are well stirred. Brine is removed from Tank 2 at the rate 3 L/min. Let $y_1(t)$ be the amount of salt in Tank 1 at time t and let $y_2(t)$ be the amount of salt in Tank 2 at time t. Determine y_1 and y_2. Find the concentration of salt in each tank after 3 minutes. Assuming the tanks are large enough, what are the long-term concentrations of brine?