Chapter 17
Matrix Functions and Systems of Differential Equations

In this chapter we give an introduction to the area of matrix functions. We first define general matrix functions and derive their most important properties. Using the examples of network analysis and chemical reactions, we illustrate how matrix functions arise naturally in applications. The network analysis example involves the exponential function of matrices, and we study the properties of this important function in detail. The analysis of chemical reaction kinetics leads to a system of ordinary differential equations, whose solution again is based on the matrix exponential function.

17.1 Matrix Functions and the Matrix Exponential Function

In the following we will study functions that yield for a given $n \times n$ matrix again an $n \times n$ matrix. A possible definition of such a function is given by the entrywise application of scalar functions to the matrix. For instance, one could define for $A = [a_{ij}] \in \mathbb{C}^{n,n}$ the function $\sin(A)$ by $\sin(A) := [\sin(a_{ij})]$. However, such a definition is not compatible with the matrix multiplication, since in general already $A^2 \neq [a_{ij}^2]$.

The following definition of the *primary matrix function* from [Hig08, Definition 1.1–1.2] will turn out to be consistent with the matrix multiplication. Since this definition is based on the Jordan canonical form, we assume for simplicity that $A \in \mathbb{C}^{n,n}$. Our considerations also apply to square matrices over $\mathbb{R}$, as long as they have a Jordan canonical form.

Definition 17.1 Let $A \in \mathbb{C}^{n,n}$ have the Jordan canonical form

$$J = \operatorname{diag}(J_{d_1}(\lambda_1), \ldots, J_{d_m}(\lambda_m)) = S^{-1}AS,$$

J. Liesen and V. Mehrmann, *Linear Algebra*, Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-24346-7_17

and let $\Omega \subset \mathbb{C}$ be such that $\{\lambda_1, \ldots, \lambda_m\} \subseteq \Omega$. A function $f : \Omega \to \mathbb{C}$ is said to be *defined on the spectrum of* A, if the values

$$f^{(j)}(\lambda_i) \quad \text{for} \quad i = 1, \ldots, m \quad \text{and} \quad j = 0, 1 \ldots, d_i - 1 \tag{17.1}$$

exist. Here $f^{(j)}(\lambda_i)$, $j = 1, \ldots, d_i - 1$, is the jth derivative of the function $f(\lambda)$ with respect to λ evaluated at λ_i. If $\lambda_i \in \mathbb{R}$, then this is the real derivative, and for $\lambda_i \in \mathbb{C} \setminus \mathbb{R}$ it is the complex derivative. Moreover, we assume that equal eigenvalues that occur in different Jordan blocks are mapped to the same values in (17.1).

If f is defined on the spectrum of A then the *primary matrix function* $f(A)$ is defined by

$$f(A) := Sf(J)S^{-1} \quad \text{where} \quad f(J) := \operatorname{diag}(f(J_{d_1}(\lambda_1)), \ldots, f(J_{d_m}(\lambda_m))) \tag{17.2}$$

and

$$f\left(J_{d_i}(\lambda_i)\right) := \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{f''(\lambda_i)}{2!} & \cdots & \frac{f^{(d_i-1)}(\lambda_i)}{(d_i-1)!} \\ & f(\lambda_i) & f'(\lambda_i) & \ddots & \vdots \\ & & \ddots & \ddots & \frac{f''(\lambda_i)}{2!} \\ & & & \ddots & f'(\lambda_i) \\ & & & & f(\lambda_i) \end{bmatrix} \quad \text{for } i = 1, \ldots, m. \tag{17.3}$$

Note that for the definition of $f(A)$ in (17.2)–(17.3) only the existence of the values in (17.1) is required.

Example 17.2 Let $A = I_2 \in \mathbb{C}^{2,2}$ and let $f(z) = \sqrt{z}$ (the square root function). If we set $f(1) = \sqrt{1} = +1$, then $f(A) = \sqrt{A} = I_2$ by Definition 17.1. If we choose the other branch of the square root function, i.e., $f(1) = \sqrt{1} = -1$, then $f(A) = \sqrt{A} = -I_2$. The matrices I_2 and $-I_2$ are *primary square roots* of $A = I_2$. Taking different branches of a function for different Jordan blocks corresponding to the same eigenvalue is incompatible with Definition 17.1. For instance, the matrices

$$X_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad X_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

are incompatible with Definition 17.1, despite the fact that $X_1^2 = I_2$ and $X_2^2 = I_2$.

All solutions $X \in \mathbb{C}^{n,n}$ of the matrix equation $X^2 = A$ are called square roots of the matrix $A \in \mathbb{C}^{n,n}$. But as Example 17.2 shows, some of these may not be primary square roots according to Definition 17.1. In the following, by $f(A)$ we will always mean a primary matrix function according to Definition 17.1, and will usually omit the term "primary".

In (16.8) we have shown that for each polynomial $p \in \mathbb{C}[t]$ of degree $k \geq 0$ we have

$$p(J_{d_i}(\lambda_i)) = \sum_{j=0}^{k} \frac{p^{(j)}(\lambda_i)}{j!} (J_{d_i}(0))^j. \tag{17.4}$$

A simple comparison shows that this formula agrees with (17.3) for $f = p$. This means that the *computation* of $p(J_{d_i}(\lambda_i))$ with (17.4) leads to the same result as the *definition* of $p(J_{d_i}(\lambda_i))$ by (17.3). More generally, the following result holds.

Lemma 17.3 *Let $A \in \mathbb{C}^{n,n}$ and $p = \alpha_k t^k + \ldots + \alpha_1 t + \alpha_0 \in \mathbb{C}[t]$. Then (17.2)–(17.3) with $f = p$ yields a matrix function $f(A)$ that satisfies $f(A) = \alpha_k A^k + \ldots + \alpha_1 A + \alpha_0 I_n$.*

Proof Exercise. □

If we consider, in particular, the polynomial $f = t^2$ in (17.2)–(17.3), then the resulting $f(A)$ is equal to the product $A * A$. This shows that the definition of the primary matrix function $f(A)$ is consistent with the matrix multiplication.

The following theorem, which is of great practical and theoretical importance, shows that the matrix $f(A)$ can always be written as a polynomial in A.

Theorem 17.4 *Let $A \in \mathbb{C}^{n,n}$ have the minimal polynomial M_A, and let $f(A)$ be as in Definition 17.1. Then there exists a uniquely determined polynomial $p \in \mathbb{C}[t]$ of degree at most $\deg(M_A) - 1$ with $f(A) = p(A)$. In particular, $Af(A) = f(A)A$, $f(A^T) = f(A)^T$ as well as $f(VAV^{-1}) = Vf(A)V^{-1}$ for all $V \in GL_n(\mathbb{C})$.*

Proof We will not present the proof here since it requires advanced results from interpolation theory. Details can be found in [Hig08, Chap. 1]. □

Using Theorem 17.4 we can show that the primary matrix function $f(A)$ in Definition 17.1 is independent of the choice of the Jordan canonical form of A. We already know from Theorem 16.12, that the Jordan canonical form of A is unique up to the order of the Jordan blocks. If

$$J = \mathrm{diag}(J_{d_1}(\lambda_1), \ldots, J_{d_m}(\lambda_m)) = S^{-1}AS,$$
$$\widetilde{J} = \mathrm{diag}(J_{\widetilde{d}_1}(\widetilde{\lambda}_1), \ldots, J_{\widetilde{d}_m}(\widetilde{\lambda}_m)) = \widetilde{S}^{-1}A\widetilde{S}$$

are two Jordan canonical forms of A, then $\widetilde{J} = P^T JP$ for a permutation matrix $P \in \mathbb{R}^{n,n}$, where the matrices J and $\widetilde{J}$ are the same up to the order of diagonal blocks. Hence

$$\begin{aligned} f(J) &= \mathrm{diag}(f(J_{d_1}(\lambda_1)), \ldots, f(J_{d_m}(\lambda_m))) \\ &= P\left(P^T \mathrm{diag}(f(J_{d_1}(\lambda_1)), \ldots, f(J_{d_m}(\lambda_m)))P\right)P^T \\ &= P\left(\mathrm{diag}(f(J_{\widetilde{d}_1}(\widetilde{\lambda}_1)), \ldots, f(J_{\widetilde{d}_m}(\widetilde{\lambda}_m)))\right)P^T \\ &= Pf(\widetilde{J})P^T. \end{aligned}$$

Theorem 17.4 applied to the matrix J yields the existence of a polynomial p with $f(J) = p(J)$. Thus, we get

$$f(A) = Sf(J)S^{-1} = Sp(J)S^{-1} = p(A) = p(\widetilde{S}\widetilde{J}\widetilde{S}^{-1}) = \widetilde{S}P^T p(J)P\widetilde{S}^{-1} = \widetilde{S}P^T f(J)P\widetilde{S}^{-1}$$
$$= \widetilde{S}f(\widetilde{J})\widetilde{S}^{-1}.$$

Let us now consider the exponential function $f(z) = e^z$ that is infinitely often complex differentiable throughout $\mathbb{C}$. In particular, e^z is defined (in the sense of Definition 17.1) on the spectrum of every given matrix

$$A = S\text{diag}(J_{d_1}(\lambda_1), \ldots, J_{d_m}(\lambda_m))S^{-1} \in \mathbb{C}^{n,n}.$$

If $t \in \mathbb{C}$ is arbitrary (but fixed), then the derivatives of the function e^{tz} with respect to the variable z are given by

$$\frac{d^j}{dz^j} e^{tz} = t^j e^{tz}, \quad j = 0, 1, 2, \ldots .$$

We will use the notation $\exp(M)$ instead of e^M for the exponential function of a matrix M. For every Jordan block $J_d(\lambda)$ of A we then have, by (17.3) with $f(z) = e^z$,

$$\exp(tJ_d(\lambda)) = e^{t\lambda} \begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{d-1}}{(d-1)!} \\ & 1 & t & \ddots & \vdots \\ & & \ddots & \ddots & \frac{t^2}{2!} \\ & & & \ddots & t \\ & & & & 1 \end{bmatrix} = e^{t\lambda} \sum_{k=0}^{d-1} \frac{1}{k!} (tJ_d(0))^k, \tag{17.5}$$

and the matrix exponential function $\exp(tA)$ is given by

$$\exp(tA) = S\text{diag}(\exp(tJ_{d_1}(\lambda_1)), \ldots, \exp(tJ_{d_m}(\lambda_m)))S^{-1}. \tag{17.6}$$

The parameter t will be used in the next section in the context of linear differential equations.

In Analysis it is shown that for every $z \in \mathbb{C}$ the function e^z can be represented by the absolutely convergent series

$$e^z = \sum_{j=0}^{\infty} \frac{z^j}{j!}.$$

Using this series and the equation $(J_d(0))^\ell = 0$ for all $\ell \geq d$, we obtain

$$
\begin{aligned}
\exp(tJ_d(\lambda)) &= e^{t\lambda}\sum_{\ell=0}^{d-1}\frac{1}{\ell!}\,(tJ_d(0))^\ell = \left(\sum_{j=0}^{\infty}\frac{(t\lambda)^j}{j!}\right)\cdot\left(\sum_{\ell=0}^{\infty}\frac{1}{\ell!}\,(tJ_d(0))^\ell\right)\\
&= \sum_{j=0}^{\infty}\left(\sum_{\ell=0}^{j}\frac{(t\lambda)^{j-\ell}}{(j-\ell)!}\cdot\frac{1}{\ell!}\,(tJ_d(0))^\ell\right)\\
&= \sum_{j=0}^{\infty}\frac{t^j}{j!}\left(\sum_{\ell=0}^{j}\binom{j}{\ell}\lambda_i^{j-\ell}\,(J_d(0))^\ell\right)\\
&= \sum_{j=0}^{\infty}\frac{t^j}{j!}\,(\lambda I_d + J_d(0))^j\\
&= \sum_{j=0}^{\infty}\frac{1}{j!}\,(tJ_d(\lambda))^j\,. \qquad (17.7)
\end{aligned}
$$

In this derivation we have used the absolute convergence of the exponential series and the finiteness of the series with the matrix $J_d(0)$. This allows the application of the *Cauchy product formula*[1] for absolutely convergent series, which is also proven in Analysis.

Lemma 17.5 *If* $A \in \mathbb{C}^{n,n}$, $t \in \mathbb{C}$ *and* $\exp(tA)$ *is the matrix exponential function in (17.5)–(17.6), then*

$$\exp(tA) = \sum_{j=0}^{\infty}\frac{1}{j!}(tA)^j.$$

Proof In (17.7) we have shown this already for Jordan blocks. The assertion then follows from

$$\sum_{j=0}^{\infty}\frac{1}{j!}\,(tSJS^{-1})^j = S\left(\sum_{j=0}^{\infty}\frac{1}{j!}(tJ)^j\right)S^{-1}$$

and the representation (17.6) of the matrix exponential function. □

We immediately see from Lemma 17.5 that for a matrix $A \in \mathbb{R}^{n,n}$ and every real t the matrix exponential function $\exp(tA)$ is a real matrix.

The following result presents further important properties of the matrix exponential function.

Lemma 17.6 *If the two matrices* $A, B \in \mathbb{C}^{n,n}$ *commute, then* $\exp(A + B) = \exp(A)\exp(B)$. *For every matrix* $A \in \mathbb{C}^{n,n}$ *we have* $\exp(A) \in GL_n(\mathbb{C})$ *with* $(\exp(A))^{-1} = \exp(-A)$.

[1] Augustin Louis Cauchy (1789–1857).

Proof If A and B commute, then the Cauchy product formula yields

$$\begin{aligned}\exp(A)\,\exp(B) &= \left(\sum_{j=0}^{\infty}\frac{1}{j!}A^j\right)\left(\sum_{\ell=0}^{\infty}\frac{1}{\ell!}B^\ell\right) = \sum_{j=0}^{\infty}\left(\sum_{\ell=0}^{j}\frac{1}{\ell!}A^\ell\,\frac{1}{(j-\ell)!}B^{j-\ell}\right)\\ &= \sum_{j=0}^{\infty}\left(\frac{1}{j!}\sum_{\ell=0}^{j}\binom{j}{\ell}A^\ell\,B^{j-\ell}\right) = \sum_{j=0}^{\infty}\frac{1}{j!}(A+B)^j\\ &= \exp(A+B).\end{aligned}$$

Here we have used the binomial formula for commuting matrices (cp. Exercise 4.10).

Since A and $-A$ commute, we have

$$\exp(A)\exp(-A) = \exp(A-A) = \exp(0) = \sum_{j=0}^{\infty}\frac{1}{j!}0^j = I_n,$$

and hence $\exp(A) \in GL_n(\mathbb{C})$ with $(\exp(A))^{-1} = \exp(-A)$. □

For non-commuting matrices the statements in Lemma 17.6 in general do not hold (cp. Exercise 17.9).

MATLAB-Minute.
Compute the matrix exponential function $\exp(A)$ for the matrix

$$A = \begin{bmatrix} 1 & -1 & 3 & 4 & 5\\ -1 & -2 & 4 & 3 & 5\\ 2 & 0 & -3 & 1 & 5\\ 3 & 0 & 0 & -2 & -3\\ 4 & 0 & 0 & -3 & -5\end{bmatrix} \in \mathbb{R}^{5,5}$$

using the command `E1=expm(A)`. (Look at `help expm`.)
Also compute the diagonalization of A using the command `[S,D]=eig(A)`, and form the matrix exponential function $\exp(A)$ as `E2=S*expm(D)/S`.
Compare the matrices `E1` and `E2` and compute the relative error `norm(E1-E2)/norm(E2)`. (Look at `help norm`.)

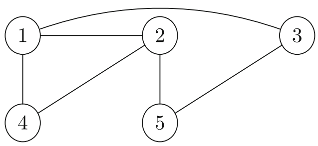
Example 17.7 Let $A = [a_{ij}] \in \mathbb{C}^{n,n}$ be a symmetric matrix with $a_{ii} = 0$ and $a_{ij} \in \{0, 1\}$ for all $i, j = 1, \dots, n$. We identify the matrix A with a *graph* $G_A = (V_A, E_A)$ consisting of a set of n *vertices* $V_A = \{1, \dots, n\}$ and a set of *edges* $E_A \subseteq V_A \times V_A$. For $i = 1, \dots, n$ the row i of A is identified with the vertex $i \in E_A$, and every entry $a_{ij} = 1$ is identified with an edge $(i, j) \in E_A$. Due to the symmetry of A, we have $a_{ij} = 1$ if and only if $a_{ji} = 1$. We therefore consider in the following the elements of E_A as *unordered pairs*, i.e., $(i, j) = (j, i)$. The following example illustrates this identification:

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

is identified with $G_A = (V_A, E_A)$, where

$$E_A = \{1, 2, 3, 4, 5\}, \quad V_A = \{(1, 2),\ (1, 3),\ (1, 4),\ (2, 4),\ (2, 5),\ (3, 5)\},$$

and the graph G_A can be displayed as follows:



A *path* of length m from the vertex k_1 to the vertex k_{m+1} is an ordered list of vertices $k_1, k_2, \ldots, k_{m+1}$, where $(k_i, k_{i+1}) \in V_A$ for $i = 1, \ldots, m$. If $k_1 = k_{m+1}$, then this is a closed path of length m. In the above example, paths from 1 to 4 are given by 1, 2, 4 and 1, 2, 5, 3, 1, 2, 4; these have the lengths 2 and 6, respectively. In the mathematical field of Graph Theory one usually assumes that the vertices in a path are pairwise distinct. Our deviation from this convention is motivated by the following interpretation of a matrix A and its powers:

An entry $a_{ij} = 1$ in the matrix A means that there exists a path of length 1 from vertex i to vertex j, i.e., the vertices i and j are adjacent. If $a_{ij} = 0$, then no such path exists. The matrix A is therefore called the *adjacency matrix* of the graph G_A. If we square the adjacency matrix, then the entry in the (i, j) position is given by

$$(A^2)_{ij} = \sum_{\ell=1}^{n} a_{i\ell} a_{\ell j}.$$

In the sum on the right hand side, we obtain for a given ℓ a 1 if and only if $(i, \ell) \in E_A$ and $(\ell, j) \in E_A$. The sum on the right had side therefore is equal to the number of vertices that are adjacent to both i and j. Hence the (i, j) entry of A^2 is equal to the number of pairwise distinct paths from i to j ($i \neq j$), or the pairwise distinct closed paths from i to i of length 2 in G_A. More generally, one can show the following (cp. Exercise 17.10):

Let $A = [a_{ij}] \in \mathbb{C}^{n,n}$ be a symmetric adjacency matrix, i.e., $A = A^T$ with $a_{ii} = 0$ and $a_{ij} \in \{0, 1\}$ for all $i, j = 1, \ldots, n$, and let G_A be the graph identified with A. Then for each $m \in \mathbb{N}$ the (i, j) entry of A^m is equal to the number of pairwise distinct paths from i to j ($i \neq j$) or the pairwise distinct closed paths from i to i of length m in G_A.

For the above matrix A we obtain

$$A^2 = \begin{bmatrix} 3 & 1 & 0 & 1 & 2 \\ 1 & 3 & 2 & 1 & 0 \\ 0 & 2 & 2 & 1 & 0 \\ 1 & 1 & 1 & 2 & 1 \\ 2 & 0 & 0 & 1 & 2 \end{bmatrix} \quad \text{and} \quad A^3 = \begin{bmatrix} 2 & 6 & 5 & 4 & 1 \\ 6 & 2 & 1 & 4 & 5 \\ 5 & 1 & 0 & 2 & 4 \\ 4 & 4 & 2 & 2 & 2 \\ 1 & 5 & 4 & 2 & 0 \end{bmatrix}.$$

The 3 pairwise distinct closed paths of length 2 from 1 to 1 are

$$1, 2, 1, \quad 1, 3, 1, \quad 1, 4, 1$$

and the 4 pairwise distinct paths of length 3 from 1 to 4 are

$$1, 2, 1, 4, \quad 1, 3, 1, 4, \quad 1, 4, 1, 4, \quad 1, 4, 2, 4.$$

Numerous real world applications involve networks that can be modeled mathematically using graphs. Examples include social, biological, telecommunication or airline networks. The properties of such networks are studied in the interdisciplinary area of *Network Science*. An important task is to identify participants in the network that are central in the sense that their functionality has a significant impact on the entire network. If the network has been modeled by a graph, then we can study the *centrality* of the vertices. For example, a vertex can be considered central if it is connected to a large part of the graph via many short closed paths. Longer connections are usually less important, and thus paths should be scaled down according to their length. If we use the scaling factor $1/m!$ for a path of length m, then for the vertex i in the graph G_A with the adjacency matrix A we obtain a centrality measure of the form

$$\left(\frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \ldots\right)_{ii}.$$

The relative ordering of the vertices according to this formula is not changed when we add the constant 1. We then obtain the *centrality* of the vertex i as

$$\left(I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \ldots\right)_{ii} = (\exp(A))_{ii}.$$

Another important quantity is the so-called *communicability* between the vertices i and j for $i \neq j$, which is given by the weighted sum of the pairwise distinct paths from i to j, i.e., by

$$\left(I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \ldots\right)_{ij} = (\exp(A))_{ij}.$$

For the above matrix A the MATLAB function `expm` yields

$$\exp(A) = \begin{bmatrix} 3.7630 & 3.1953 & 2.2500 & 2.7927 & 1.8176 \\ 3.1953 & 3.7630 & 1.8176 & 2.7927 & 2.2500 \\ 2.2500 & 1.8176 & 2.4881 & 1.2749 & 1.9204 \\ 2.7927 & 2.7927 & 1.2749 & 2.8907 & 1.2749 \\ 1.8176 & 2.2500 & 1.9204 & 1.2749 & 2.4881 \end{bmatrix}.$$

The vertices 1 and 2 have the largest centrality, followed by 4, 3 and 5. If we would define the centrality of a vertex as the number of adjacent vertices, then in this example we could not distinguish between the vertices 3, 4 and 5. The largest communicability in this example exists between the vertices 1 and 2.

Further information concerning the analysis of networks using adjacency matrices and matrix functions can be found in the article [EstH10].

17.2 Systems of Linear Ordinary Differential Equations

A differential equation describes a relationship between a desired function and its derivatives. Such equations are used in all areas of science and engineering for modeling physical phenomena. Ordinary differential equations involve a function of one variable and its derivatives, while partial differential equations involve functions of several variables and their partial derivatives. In this section we focus on ordinary differential equations of first order, i.e., those in which only the function and its first derivative occur.

A simple example for the modeling with ordinary differential equations of first order is the increase or decrease of a biological population, such as bacteria in a petri dish. Let $y = y(t)$ be the size of the population at time t. If there is enough food and if the external conditions (e.g. temperature or pressure) are constant, then the population grows with a (real) rate $k > 0$, that is proportional to the current number of individuals. This can be described by the equation

$$\dot{y} := \frac{d}{dt} y = ky. \tag{17.8}$$

Clearly, one can also take $k < 0$, and then the population shrinks.

We are then looking for a function $y : D \subset \mathbb{R} \to \mathbb{R}$ that satisfies (17.8). The general solution of (17.8) is given by the exponential function

$$y = ce^{tk},$$

where $c \in \mathbb{R}$ is an arbitrary constant. For a unique solution of (17.8) we need to know the size of the population at a given initial time t_0. In this way we obtain the *initial value problem*

$$\dot{y} = ky, \quad y(t_0) = y_0,$$

which, as we will show below, is solved uniquely by the function

$$y = e^{(t-t_0)k} y_0.$$

Example 17.8 In a chemical reaction certain initial substances (called educts or reactants) are transformed into other substances (called products). Reactions can be distinguished concerning their order. Here we only discuss reactions of first order, where the reaction rate is determined by only one educt. In reactions of second and higher order one typically obtains *nonlinear* differential equations, which are beyond our focus in this chapter.

If, for example, the educt A_1 is transformed into the product A_2 with the rate $-k_1 < 0$, then we write this reaction symbolically as

$$A_1 \xrightarrow{k_1} A_2\,,$$

and we model it mathematically by the ordinary differential equation

$$\dot{y}_1 = -k_1 y_1.$$

Here the value $y_1(t)$ is the concentration of the substance A_1 at time t. For the concentration of the product A_2, which grows with the rate $k_1 > 0$, we have the corresponding equation $\dot{y}_2 = k_1 y_1$.

It may happen that a reaction of first order develops in both directions. If A_1 transforms into A_2 with the rate $-k_1$, and A_2 transforms into A_1 with the rate $-k_2$, i.e.,

$$A_1 \underset{k_2}{\overset{k_1}{\rightleftarrows}} A_2\,,$$

then we can model this reaction mathematically by the system of linear ordinary differential equations

$$\begin{aligned}\dot{y}_1 &= -k_1 y_1 + k_2 y_2,\\ \dot{y}_2 &= k_1 y_1 - k_2 y_2.\end{aligned}$$

Combining the functions y_1 and y_2 in a vector valued function $y = [y_1,\ y_2]^T$, we can write this system as

$$\dot{y} = Ay, \quad \text{where} \quad A = \begin{bmatrix} -k_1 & k_2 \\ k_1 & -k_2 \end{bmatrix}.$$

The derivative of the function $y(t)$ is always considered entrywise,

$$\dot{y} = \begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \end{bmatrix}.$$

Reactions can also have several steps. For example, a reaction of the form

$$A_1 \xrightarrow{k_1} A_2 \underset{k_3}{\overset{k_2}{\rightleftarrows}} A_3 \xrightarrow{k_4} A_4$$

leads to the differential equations

$$\begin{aligned}
\dot{y}_1 &= -k_1 y_1, \\
\dot{y}_2 &= k_1 y_1 - k_2 y_2 + k_3 y_3, \\
\dot{y}_3 &= k_2 y_2 - (k_3 + k_4) y_3, \\
\dot{y}_4 &= k_4 y_3,
\end{aligned}$$

and thus to the system

$$\dot{y} = Ay, \quad \text{where} \quad A = \begin{bmatrix} -k_1 & 0 & 0 & 0 \\ k_1 & -k_2 & k_3 & 0 \\ 0 & k_2 & -(k_3 + k_4) & 0 \\ 0 & 0 & k_4 & 0 \end{bmatrix}.$$

The sum of the entries in each column of A is equal to zero, since for every decrease in a substance with a certain rate other substances increase with the same rate.

In summary, a chemical reaction of first order leads to a system of linear ordinary differential equations of first order that can be written as $\dot{y} = Ay$ with a (real) square matrix A.

We now derive the general theory for systems of linear (real or complex) ordinary differential equations of first order of the form

$$\dot{y} = Ay + g, \quad t \in [0, a]. \tag{17.9}$$

Here $A \in K^{n,n}$ is a given matrix, a is a given positive real number, $g : [0, a] \to K^{n,1}$ is a given function, $y : [0, a] \to K^{n,1}$ is the desired solution, and we assume that $K = \mathbb{R}$ or $K = \mathbb{C}$. If $g(t) = 0 \in K^{n,1}$ for all $t \in [0, a]$, then the system (17.9) is called *homogeneous*, otherwise it is called *non-homogeneous*. For a given system of the form (17.9), the system

$$\dot{y} = Ay, \quad t \in [0, a], \tag{17.10}$$

is called the *associated homogeneous system*.

Lemma 17.9 *The solutions of the homogeneous system (17.10) form a subspace of the (infinite dimensional) K-vector space of the continuously differentiable functions from the interval* $[0, a]$ *to* $K^{n,1}$.

Proof We will show the required properties according to Lemma 9.5. The function $w = 0$ is continuously differentiable on $[0, a]$ and solves the homogeneous system (17.10). Thus, the solution set of this system is not empty. If

$$w_1, w_2 \,:\, [0, a] \to K^{n,1}$$

are continuously differentiable solutions and if $\alpha_1, \alpha_2 \in K$, then $w = \alpha_1 w_1 + \alpha_2 w_2$ is continuously differentiable on $[0, a]$, and

$$\dot{w} = \alpha_1 \dot{w}_1 + \alpha_2 \dot{w}_2 = \alpha_1 A w_1 + \alpha_2 A w_2 = A w,$$

i.e., the function w is a solution of the homogeneous system. □

The following characterization of the solutions of the non-homogeneous system (17.9) is analogous to the characterization of the solution set of a non-homogeneous linear system of equations in Lemma 6.2 (also cp. (8) in Lemma 10.7).

Lemma 17.10 *If* $w_1 : [0, a] \to K^{n,1}$ *is a solution of the non-homogeneous system (17.9), then every other solution* y *can be written as* $y = w_1 + w_2$*, where* w_2 *is a solution of the associated homogeneous system (17.10).*

Proof If w_1 and y are solutions of (17.9), then $\dot{y} - \dot{w}_1 = (Ay + g) - (Aw_1 + g) = A(y - w_1)$. The difference $w_2 := y - w_1$ thus is a solution of the associated homogeneous system and $y = w_1 + w_2$. □

In order to describe the solutions of systems of ordinary differential equations, we consider for a given matrix $A \in K^{n,n}$ the matrix exponential function $\exp(tA)$ from Lemma 17.5 or (17.5)–(17.6), where we now consider $t \in [0, a]$ as *real variable*. The power series of the matrix exponential function in Lemma 17.5 converges, and it can be differentiated termwise with respect to the variable t, where again the derivative of a matrix with respect to the variable t is considered entrywise. This yields

$$\begin{aligned}\frac{d}{dt}\exp(tA) &= \frac{d}{dt}\left(I + (tA) + \frac{1}{2}(tA)^2 + \frac{1}{6}(tA)^3 + \ldots\right)\\ &= A + tA^2 + \frac{1}{2}t^2A^3 + \ldots\\ &= A\exp(tA).\end{aligned}$$

The same result is obtained by the entrywise differentiation of the matrix $\exp(tA)$ in (17.5)–(17.6) with respect to t. With

$$M(t) := \begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{d-1}}{(d-1)!} \\ & 1 & t & \ddots & \vdots \\ & & \ddots & \ddots & \frac{t^2}{2!} \\ & & & \ddots & t \\ & & & & 1 \end{bmatrix}$$

we obtain

$$\begin{aligned} \frac{d}{dt}\exp(tJ_d(\lambda)) &= \frac{d}{dt}\left(e^{t\lambda}M(t)\right) \\ &= \lambda e^{t\lambda}M(t) + e^{t\lambda}\dot{M}(t) \\ &= \lambda e^{t\lambda}M(t) + e^{t\lambda}J_d(0)M(t) \\ &= (\lambda I_d + J_d(0))\, e^{t\lambda}M(t) \\ &= J_d(\lambda)\exp(tJ_d(\lambda)), \end{aligned}$$

which also gives $\frac{d}{dt}\exp(tA) = A\exp(tA)$.

Theorem 17.11

(1) The unique solution of the homogeneous differential equation system (17.10) for a given initial condition $y(0) = y_0 \in K^{n,1}$ is given by the function $y = \exp(tA)y_0$.

(2) The set of all solutions of the homogeneous differential equation system (17.10) forms an n-dimensional K-vector space with the basis $\{\exp(tA)e_1, \ldots, \exp(tA)e_n\}$.

Proof

(1) If $y = \exp(tA)y_0$, then

$$\begin{aligned} \dot{y} &= \frac{d}{dt}(\exp(tA)y_0) = \left(\frac{d}{dt}\exp(tA)\right)y_0 = (A\exp(tA))y_0 \\ &= A(\exp(tA)y_0) = Ay, \end{aligned}$$

and $y(0) = \exp(0)y_0 = I_n y_0 = y_0$. Hence y is a solution of (17.10) that satisfies the initial condition. If w is another such solution and $u := \exp(-tA)w$, then

$$\begin{aligned} \dot{u} &= \frac{d}{dt}(\exp(-tA)w) = -A\exp(-tA)w + \exp(-tA)\dot{w} \\ &= \exp(-tA)(\dot{w} - Aw) = 0 \in K^{n,1}, \end{aligned}$$

which shows that the function u has constant entries. In particular, we then have $u = u(0) = w(0) = y_0 = y(0)$ and $w = \exp(tA)y_0$, where we have used that $\exp(-tA) = (\exp(tA))^{-1}$ (cp. Lemma 17.6).

(2) Each of the functions $\exp(tA)e_j, \ldots, \exp(tA)e_n : [0, a] \to K^{n,1}$, $j = 1, \ldots, n$, solves the homogeneous system $\dot{y} = Ay$. Since the matrix $\exp(tA) \in K^{n,n}$ is invertible for every $t \in [0, a]$ (cp. Lemma 17.6), these functions are linearly independent.

If $\widetilde{y}$ is an arbitrary solution of $\dot{y} = Ay$, then $\widetilde{y}(0) = y_0$ for some $y_0 \in K^{n,1}$. By (1) then $\widetilde{y}$ is the unique solution of the initial value problem with $y(0) = y_0$, so that $\widetilde{y} = \exp(tA)y_0$. As a consequence, $\widetilde{y}$ is a linear combination of the functions $\exp(tA)e_1, \ldots, \exp(tA)e_n$. □

To describe the solution of the non-homogeneous system (17.9), we need the integral of functions of the form

$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} : [0, a] \to K^{n,1}.$$

For every fixed $t \in [0, a]$ we define

$$\int_0^t w(s)ds := \begin{bmatrix} \int_0^t w_1(s)ds \\ \vdots \\ \int_0^t w_n(s)ds \end{bmatrix} \in K^{n,1},$$

i.e., we apply the integral entrywise to the function w. By this definition we have

$$\frac{d}{dt}\left(\int_0^t w(s)ds\right) = w(t)$$

for all $t \in [0, a]$. We can now determine an explicit solution formula for systems of linear differential equations based on the so-called *Duhamel integral*.[2]

Theorem 17.12 *The unique solution of the non-homogeneous differential equation system (17.9) with the initial condition* $y(0) = y_0 \in K^{n,1}$ *is given by*

$$y = \exp(tA)y_0 + \exp(tA)\int_0^t \exp(-sA)g(s)ds. \tag{17.11}$$

Proof The derivative of the function y defined in (17.11) is

$$\begin{aligned}\dot{y} &= \frac{d}{dt}(\exp(tA)y_0) + \frac{d}{dt}\left(\exp(tA)\int_0^t \exp(-sA)g(s)ds\right)\\ &= A\exp(tA)y_0 + A\exp(tA)\int_0^t \exp(-sA)g(s)ds + \exp(tA)\exp(-tA)g\end{aligned}$$

[2]Jean-Marie Constant Duhamel (1797–1872).

$$= A\exp(tA)y_0 + A\exp(tA)\int_0^t \exp(-sA)g(s)ds + g$$
$$= Ay + g.$$

Furthermore, we have

$$y(0) = \exp(0)y_0 + \exp(0)\int_0^0 \exp(-sA)g(s)ds = y_0,$$

so that y also satisfies the initial condition.

Let now $\widetilde{y}$ be another solution of (17.9) that satisfies the initial condition. By Lemma 17.10 we then have $\widetilde{y} = y + w$, where w solves the homogeneous system (17.10). Therefore, $w = \exp(tA)c$ for some $c \in K^{n,1}$ (cp. (2) in Theorem 17.11). For $t = 0$ we obtain $y_0 = y_0 + c$, where $c = 0$ and hence $\widetilde{y} = y$. □

In the above theorems we have shown that for the explicit solution of systems of linear ordinary differential equations of first order, we have to compute the matrix exponential function. While we have introduced this function using the Jordan canonical form of the given matrix, numerical computations based on the Jordan canonical form are not advisable (cp. Example 16.20). Because of its significant practical relevance, numerous different algorithms for computing the matrix exponential function have been proposed. But, as shown in the article [MolV03], no existing algorithm is completely satisfactory.

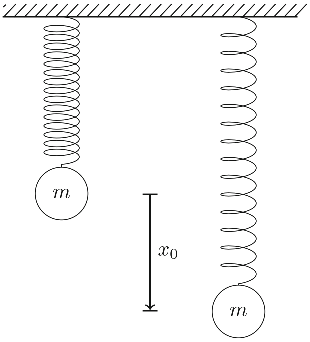
Example 17.13 The example from circuit simulation presented in Sect. 1.5 lead to the system of ordinary differential equations

$$\frac{d}{dt}I = -\frac{R}{L}I - \frac{1}{L}V_C + \frac{1}{L}V_S,$$
$$\frac{d}{dt}V_C = -\frac{1}{C}I.$$

Using (17.11) and the initial values $I(0) = I^0$ and $V_C(0) = V_C^0$, we obtain the solution

$$\begin{bmatrix} I \\ V_C \end{bmatrix} = \exp\left(t\begin{bmatrix} -R/L & -1/L \\ -1/C & 0 \end{bmatrix}\right)\begin{bmatrix} I^0 \\ V_C^0 \end{bmatrix} + \int_0^t \exp\left((t-s)\begin{bmatrix} -R/L & -1/L \\ -1/C & 0 \end{bmatrix}\right)\begin{bmatrix} V_S(s) \\ 0 \end{bmatrix} ds.$$

Example 17.14 Let us also consider an example from Mechanics. A weight with mass $m > 0$ is attached to a spring with the spring constant $\mu > 0$. Let $x_0 > 0$ be the distance of the weight from its equilibrium position, as illustrated in the following figure:



We want to determine the position $x(t)$ of the weight at time $t \geq 0$, where $x(0) = x_0$. The extension of the spring is described by *Hooke's law*.[3] The corresponding ordinary differential equation of second order is

$$\ddot{x} = \frac{d^2}{dt^2}x = -\frac{\mu}{m}x,$$

with initial conditions $x(0) = x_0$ and $\dot{x}(0) = v_0$, where $v_0 > 0$ is the initial velocity of the weight. We can write this differential equation of second order for x as a system of first order by introducing the velocity v as new variable. The velocity is given by the derivative of the position with respect to time, i.e., $v = \dot{x}$, and thus for the acceleration we have $\dot{v} = \ddot{x}$, which yields the system

$$\dot{y} = Ay, \quad \text{where} \quad A = \begin{bmatrix} 0 & 1 \\ -\frac{\mu}{m} & 0 \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} x \\ v \end{bmatrix}.$$

The initial condition then is $y(0) = y_0 = [x_0,\ v_0]^T$.

By Theorem 17.11, the unique solution of this homogeneous initial value problem is given by the function $y = \exp(tA)y_0$. We consider A as an element of $\mathbb{C}^{2,2}$. The eigenvalues of A are the two complex (non-real) numbers $\lambda_1 = \mathbf{i}\rho$ and $\lambda_2 = -\mathbf{i}\rho = \overline{\lambda}_1$, where $\rho := \sqrt{\frac{\mu}{m}}$. Corresponding eigenvectors are

$$s_1 = \begin{bmatrix} 1 \\ \mathbf{i}\rho \end{bmatrix} \in \mathbb{C}^{2,1}, \quad s_2 = \begin{bmatrix} 1 \\ -\mathbf{i}\rho \end{bmatrix} \in \mathbb{C}^{2,1}$$

and thus

$$\exp(tA)y_0 = S \begin{bmatrix} e^{\mathbf{i}t\rho} & 0 \\ 0 & e^{-\mathbf{i}t\rho} \end{bmatrix} S^{-1}y_0, \quad S = \begin{bmatrix} 1 & 1 \\ \mathbf{i}\rho & -\mathbf{i}\rho \end{bmatrix} \in \mathbb{C}^{2,2}.$$

[3] Sir Robert Hooke (1635–1703).

Exercises

17.1 Construct a matrix $A = [a_{ij}] \in \mathbb{C}^{2,2}$ with $A^3 \neq \left[a_{ij}^3\right]$.

17.2 Determine all solutions $X \in \mathbb{C}^{2,2}$ of the matrix equation $X^2 = I_2$, and classify which of these solutions are primary square roots of I_2.

17.3 Determine a matrix $X \in \mathbb{C}^{2,2}$ with real entries and $X^2 = -I_2$.

17.4 Prove Lemma 17.3.

17.5 Prove the following assertions for $A \in \mathbb{C}^{n,n}$:

(a) $\det(\exp(A)) = \exp(\operatorname{trace}(A))$.
(b) If $A^H = -A$, then $\exp(A)$ is unitary.
(c) If $A^2 = I$, then $\exp(A) = \frac{1}{2}(e + \frac{1}{e})I + \frac{1}{2}(e - \frac{1}{e})A$.

17.6 Let $A = S\,\operatorname{diag}(J_{d_1}(\lambda_1), \ldots, J_{d_m}(\lambda_m))\,S^{-1} \in \mathbb{C}^{n,n}$ with $\operatorname{rank}(A) = n$. Determine the primary matrix function $f(A)$ for $f(z) = z^{-1}$. Does this function also exist if $\operatorname{rank}(A) < n$?

17.7 Let $\log : \{z = re^{i\varphi} \,|\, r > 0, -\pi < \varphi < \pi\} \to \mathbb{C}$, $re^{i\varphi} \mapsto \ln(r) + i\varphi$, be the principle branch of the complex logarithm (where ln denotes the real natural logarithm). Show that this function is defined on the spectrum of

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathbb{C}^{2,2},$$

and compute $\log(A)$ as well as $\exp(\log(A))$.

17.8 Compute

$$\exp\left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right), \quad \exp\left(\begin{bmatrix} -1 & 1 \\ -1 & -3 \end{bmatrix}\right), \quad \sin\left(\begin{bmatrix} \pi & 1 & 1 \\ 0 & \pi & 1 \\ 0 & 0 & \pi \end{bmatrix}\right).$$

17.9 Construct two matrices $A, B \in \mathbb{C}^{2,2}$ with $\exp(A + B) \neq \exp(A)\exp(B)$.

17.10 Prove the assertion on the entries of A^d in Example 17.7.

17.11 Let

$$A = \begin{bmatrix} 5 & 1 & 1 \\ 0 & 5 & 1 \\ 0 & 0 & 4 \end{bmatrix} \in \mathbb{R}^{3,3}.$$

Compute $\exp(tA)$ for $t \in \mathbb{R}$ and solve the homogeneous system of differential equations $\dot{y} = Ay$ with the initial condition $y(0) = [1, 1, 1]^T$.

17.12 Compute the matrix $\exp(tA)$ from Example 17.14 explicitly and thus show that $\exp(tA) \in \mathbb{R}^{2,2}$ (for $t \in \mathbb{R}$), despite the fact that the eigenvalues and eigenvectors of A are not real.