Ordinary Differential Equations

Peter Philip*

Lecture Notes

Originally Created for the Class of Spring Semester 2012 at LMU Munich,
Revised and Extended for the Classes of Spring Semesters 2013 and 2014

June 23, 2015

Contents

*E-Mail: philip@math.lmu.de

1 Basic Notions

1.1 Types of Ordinary Differential Equations (ODE) and First Examples

A *differential equation* is an equation for some unknown function, involving one or more derivatives of the unknown function. Here are some first examples:

$$y' = y, \tag{1.1a}$$

$$y^{(5)} = (y')^2 + \pi\, x, \tag{1.1b}$$

$$(y')^2 = c, \tag{1.1c}$$

$$\partial_t x = e^{2\pi i t}\, x^2, \tag{1.1d}$$

$$x'' = -3x + \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \tag{1.1e}$$

One distinguishes between *ordinary* differential equations (ODE) and *partial* differential equations (PDE). While ODE contain only derivatives with respect to one variable, PDE can contain (partial) derivatives with respect to several different variables. In general, PDE are much harder to solve than ODE. The equations in (1.1) all are ODE, and only ODE are the subject of this class. We will see precise definitions shortly, but we can already use the examples in (1.1) to get some first exposure to important ODE-related terms and to discuss related issues.

As in (1.1), the notation for the unknown function varies in the literature, where the two variants presented in (1.1) are probably the most common ones: In the first three equations of (1.1), the unknown function is denoted y, usually assumed to depend on a variable denoted x, i.e. $x \mapsto y(x)$. In the last two equations of (1.1), the unknown function is denoted x, usually assumed to depend on a variable denoted t, i.e. $t \mapsto x(t)$. So one has to use some care due to the different roles of the symbol x. The notation $t \mapsto x(t)$ is typically favored in situations arising from physics applications, where t represents time. In this class, we will mostly use the notation $x \mapsto y(x)$.

There is another, in a way a slightly more serious, notational issue that one commonly encounters when dealing with ODE: Strictly speaking, the notation in (1.1b) and (1.1d) is not entirely correct, as functions and function arguments are not properly distinguished. Correctly written, (1.1b) and (1.1d) read

$$\forall_{x \in \mathcal{D}(y)} \quad y^{(5)}(x) = \big(y'(x)\big)^2 + \pi\, x, \tag{1.2a}$$

$$\forall_{t \in \mathcal{D}(x)} \quad (\partial_t x)(t) = e^{2\pi i t}\, \big(x(t)\big)^2, \tag{1.2b}$$

where $\mathcal{D}(y)$ and $\mathcal{D}(x)$ denote the respective domains of the functions y and x. However, one might also notice that the notation in (1.2) is more cumbersome and, perhaps, harder to read. In any case, the type of slight abuse of notation present in (1.1b) and (1.1d) is so common in the literature that one will have to live with it.

One speaks of *first-order* ODE if the equations involve only first derivatives such as in (1.1a), (1.1c), and (1.1d). Otherwise, one speaks of *higher-order* ODE, where the precise order is given by the highest derivative occurring in the equation, such that (1.1b) is an ODE of fifth order and (1.1e) is an ODE of second order. We will see later in Th. 3.1 that ODE of higher order can be equivalently formulated and solved as *systems* of ODE of first order, where systems of ODE obviously consist of several ODE to be solved simultaneously. Such a system of ODE can, equivalently, be interpreted as a single ODE in higher dimensions: For instance, (1.1e) can be seen as a single two-dimensional ODE of second order or as the system

$$x_1'' = -3x_1 - 1, \tag{1.3a}$$

$$x_2'' = -3x_2 + 1 \tag{1.3b}$$

of two one-dimensional ODE of second order.

One calls an ODE *explicit* if it has been solved explicitly for the highest-order derivative, otherwise *implicit*. Thus, in (1.1), all ODE except (1.1c) are explicit. In general, explicit ODE are much easier to solve than implicit ODE (which include, e.g., so-called differential-algebraic equations, cf. Ex. 1.4(g) below), and we will mostly consider explicit ODE in this class.

As the reader might already have noticed, without further information, none of the equations in (1.1) makes much sense. Every function, in particular, every function solving an ODE, needs a set as the domain where it is defined, and a set as the range it maps into. Thus, for each ODE, one needs to specify the admissible domains as well as the range of the unknown function. For an ODE, one usually requires a solution to be defined on a nontrivial (bounded or unbounded) interval $I \subseteq \mathbb{R}$. Prescribing the possible range of the solution is an integral part of setting up an ODE, and it often completely determines the ODE's meaning and/or its solvability. For example for (1.1d), (a subset of) $\mathbb{C}$ is a reasonable range. Similarly, for (1.1a)–(1.1c), one can require the range to be either $\mathbb{R}$ or $\mathbb{C}$, where requiring range $\mathbb{R}$ for (1.1c) immediately implies there is no solution for $c < 0$. However, one can also specify (a subset of) $\mathbb{R}^n$ or $\mathbb{C}^n$, $n > 1$, as range for (1.1a), turning the ODE into an n-dimensional ODE (or a system of ODE), where y now has n compoments $(y_1, \dots, y_n)$ (note that, except in cases where we are dealing with matrix multiplications, we sometimes denote elements of $\mathbb{R}^n$ as columns and sometimes as rows, switching back and forth without too much care). A reasonable range for (1.1e) is (a subset of) $\mathbb{R}^2$ or $\mathbb{C}^2$.

One of the important goals regarding ODE is to find conditions, where one can guanrantee the existence of solutions. Moreover, if possible, one would like to find conditions that guarantee the existence of a *unique* solution. Clearly, for each $a \in \mathbb{R}$, the function $y : \mathbb{R} \longrightarrow \mathbb{R}$, $y(x) = a\, e^x$, is a solution to (1.1a), showing one cannot expect uniqueness without specifying further requirements. The most common additional conditions that often (but not always) guarantee a unique solution are *initial conditions*, (e.g. requiring $y(x_0) = y_0$ (x_0, y_0 given); or *boundary conditions* (e.g. requiring $y(a) = y_a$, $y(b) = y_b$ for $y : [a, b] \longrightarrow \mathbb{C}^n$ ($y_a, y_b \in \mathbb{C}^n$ given)).

Let us now proceed to mathematically precise definitions of the abovementioned notions.

Notation 1.1. We will write $\mathbb{K}$ in situations, where we allow $\mathbb{K}$ to be $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.2. Let $k, n \in \mathbb{N}$.

(a) Given $U \subseteq \mathbb{R} \times \mathbb{K}^{(k+1)n}$ and $F : U \longrightarrow \mathbb{K}^n$, call

$$F(x, y, y', \dots, y^{(k)}) = 0 \tag{1.4}$$

an *implicit ODE* of kth order. A *solution* to this ODE is a k times differentiable function

$$\phi : I \longrightarrow \mathbb{K}^n, \tag{1.5}$$

defined on a nontrivial (bounded or unbounded, open or closed or half-open) interval $I \subseteq \mathbb{R}$ satisfying the two conditions

(i) $\left\{(x, \phi(x), \phi'(x), \dots, \phi^{(k)}(x)) \in I \times \mathbb{K}^{(k+1)n} : x \in I\right\} \subseteq U$.

(ii) $F(x, \phi(x), \phi'(x), \dots, \phi^{(k)}(x)) = 0$ for each $x \in I$.

Note that condition (i) is necessary so that one can even formulate condition (ii).

(b) Given $G \subseteq \mathbb{R} \times \mathbb{K}^{kn}$, and $f : G \longrightarrow \mathbb{K}^n$, call

$$y^{(k)} = f(x, y, y', \dots, y^{(k-1)}) \tag{1.6}$$

an *explicit ODE* of kth order. A *solution* to this ODE is a k times differentiable function ϕ as in (1.5), defined on a nontrivial (bounded or unbounded, open or closed or half-open) interval $I \subseteq \mathbb{R}$ satisfying the two conditions

(i) $\left\{\left(x, \phi(x), \phi'(x), \dots, \phi^{(k-1)}(x)\right) \in I \times \mathbb{K}^{kn} : x \in I\right\} \subseteq G$.

(ii) $\phi^{(k)}(x) = f(x, \phi(x), \phi'(x), \dots, \phi^{(k-1)}(x))$ for each $x \in I$.

Again, note that condition (i) is necessary so that one can even formulate condition (ii). Also note that ϕ is a solution to (1.6) if, and only if, ϕ is a solution to the equivalent implicit ODE $y^{(k)} - f(x, y, y', \dots, y^{(k-1)}) = 0$.

Definition 1.3. Let $k, n \in \mathbb{N}$.

(a) An *initial value problem* for (1.4) (resp. for (1.6)) consists of the ODE (1.4) (resp. of the ODE (1.6)) plus the *initial condition*

$$\underset{j=0,\dots,k-1}{\forall} \quad y^{(j)}(x_0) = y_{0,j} \tag{1.7}$$

with given $x_0 \in \mathbb{R}$ and $y_{0,0}, \dots, y_{0,k-1} \in \mathbb{K}^n$. A solution ϕ to the initial value problem is a k times differentiable function ϕ as in (1.5) that is a solution to the ODE and that also satisfies (1.7) (with y replaced by ϕ) – in particular, this requires $x_0 \in I$.

(b) A *boundary value problem* for (1.4) (resp. for (1.6)) consists of the ODE (1.4) (resp. of the ODE (1.6)) plus the *boundary condition*

$$\underset{j \in J_a}{\forall} \quad y^{(j)}(a) = y_{a,j} \quad \text{and} \quad \underset{j \in J_b}{\forall} \quad y^{(j)}(b) = y_{b,j} \tag{1.8}$$

with given $a, b \in \mathbb{R}$, $a < b$; $J_a, J_b \subseteq \{0, \dots, k-1\}$, $y_{a,j} \in \mathbb{K}^n$ for each $j \in J_a$, and $y_{b,j} \in \mathbb{K}^n$ for each $j \in J_b$. A solution ϕ to the boundary value problem is a k times differentiable function ϕ as in (1.5) that is a solution to the ODE and that also satisfies (1.8) (with y replaced by ϕ) – in particular, this requires $[a, b] \subseteq I$.

—

Under suitable hypotheses, initial and boundary value problems for ODE have unique solutions (for initial value problems, we will see some rather general results in Cor. 3.10 and Cor. 3.16 below). However, in general, they can have infinitely many solutions or no solutions, as shown by Examples 1.4(b),(c),(e) below.

Example 1.4. (a) Let $k \in \mathbb{N}$. The function $\phi : \mathbb{R} \longrightarrow \mathbb{K}$, $\phi(x) = a\, e^x$, $a \in \mathbb{K}$, is a solution to the kth order explicit initial value problem

$$y^{(k)} = y, \tag{1.9a}$$

$$\underset{j=0,\dots,k-1}{\forall} \quad y^{(j)}(0) = a. \tag{1.9b}$$

We will see later (e.g., as a consequence of Th. 4.8 combined with Th. 3.1) that ϕ is the unique solution to (1.9) on $\mathbb{R}$.

(b) Consider the one-dimensional explicit first-order initial value problem

$$y' = \sqrt{|y|}, \tag{1.10a}$$

$$y(0) = 0. \tag{1.10b}$$

Then, for every $c \geq 0$, the function

$$\phi_c : \mathbb{R} \longrightarrow \mathbb{R}, \quad \phi_c(x) := \begin{cases} 0 & \text{for } x \leq c, \\ \frac{(x-c)^2}{4} & \text{for } x \geq c, \end{cases} \tag{1.11}$$

is a solution to (1.10): Clearly, $\phi_c(0) = 0$, ϕ_c is differentiable, and

$$\phi_c' : \mathbb{R} \longrightarrow \mathbb{R}, \quad \phi_c'(x) := \begin{cases} 0 & \text{for } x \leq c, \\ \frac{x-c}{2} & \text{for } x \geq c, \end{cases} \tag{1.12}$$

solving the ODE. Thus, (1.10) is an example of an initial value problem with uncountably many different solutions, all defined on the same domain.

(c) As mentioned before, the one-dimensional implicit first-order ODE (1.1c) has no real-valued solution for $c < 0$. For $c \geq 0$, every function $\phi : \mathbb{R} \longrightarrow \mathbb{R}$, $\phi(x) := a \pm x\sqrt{c}$, $a \in \mathbb{R}$, is a solution to (1.1c). Moreover, for $c < 0$, every function $\phi : \mathbb{R} \longrightarrow \mathbb{C}$, $\phi(x) := a \pm xi\sqrt{-c}$, $a \in \mathbb{C}$, is a $\mathbb{C}$-valued solution to (1.1c). The one-dimensional implicit first-order ODE

$$e^{y'} = 0 \tag{1.13}$$

is an example of an ODE that does not even have a $\mathbb{C}$-valued solution. It is an exercise to find $f : \mathbb{R} \longrightarrow \mathbb{R}$ such that the explicit ODE $y' = f(x)$ has no solution.

(d) Let $n \in \mathbb{N}$ and let $a, c \in \mathbb{K}^n$. Then, on $\mathbb{R}$, the function

$$\phi : \mathbb{R} \longrightarrow \mathbb{K}^n, \quad \phi(x) := c + xa, \tag{1.14}$$

is the unique solution to the n-dimensional explicit first-order initial value problem

$$y' = a, \tag{1.15a}$$
$$y(0) = c. \tag{1.15b}$$

This situation is a special case of Ex. 1.6 below.

(e) Let $a, b \in \mathbb{R}$, $a < b$. We will see later in Example 4.12 that on $[a, b]$ the 1-dimensional explicit second-order ODE

$$y'' = -y \tag{1.16}$$

has precisely the set of solutions

$$\mathcal{L} = \Big\{ \Big((c_1 \sin + c_2 \cos) : [a, b] \longrightarrow \mathbb{K} \Big) : c_1, c_2 \in \mathbb{K} \Big\}. \tag{1.17}$$

In consequence, the boundary value problem

$$y(0) = 0, \quad y(\pi/2) = 1, \tag{1.18a}$$

for (1.16) has the unique solution $\phi : [0, \pi/2] \longrightarrow \mathbb{K}$, $\phi(x) := \sin x$ (using (1.18a) and (1.17) implies $c_2 = 0$ and $c_1 = 1$); the boundary value problem

$$y(0) = 0, \quad y(\pi) = 0, \tag{1.18b}$$

for (1.16) has the infinitely many different solutions $\phi_c : [0, \pi] \longrightarrow \mathbb{K}$, $\phi_c(x) := c \sin x$, $c \in \mathbb{K}$; and the boundary value problem

$$y(0) = 0, \quad y(\pi) = 1, \tag{1.18c}$$

for (1.16) has no solution (using (1.18c) and (1.17) implies the contradictory requirements $c_2 = 0$ and $c_2 = -1$).

(f) Consider

$$F: \mathbb{R} \times \mathbb{K}^2 \times \mathbb{K}^2 \longrightarrow \mathbb{K}^2, \quad F\left(x, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right) := \begin{pmatrix} z_2 \\ z_2 - 1 \end{pmatrix}. \tag{1.19a}$$

Clearly, the implicit $\mathbb{K}^2$-valued ODE

$$F(x, y, y') = \begin{pmatrix} y_2' \\ y_2' - 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{1.19b}$$

has no solution on any nontrivial interval.

(g) Consider

$$F: \mathbb{R} \times \mathbb{C}^3 \times \mathbb{C}^3 \longrightarrow \mathbb{C}^3, \quad F\left(x, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}\right) := \begin{pmatrix} z_2 + iy_3 - 2i \\ z_1 + y_2 - x^2 \\ y_1 - ie^{ix} \end{pmatrix}. \tag{1.20a}$$

It is an exercise to show the $\mathbb{C}^3$-valued implicit ODE

$$F(x, y, y') = \begin{pmatrix} y_2' + iy_3 - 2i \\ y_1' + y_2 - x^2 \\ y_1 - ie^{ix} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{1.20b}$$

has a unique solution on $\mathbb{R}$ (note that, here, we do not need to provide initial or boundary conditions to obtain uniqueness). The implicit ODE (1.20b) is an example of a differential algebraic equation, since, read in components, only its first two equations contain derivatives, whereas its third equation is purely algebraic.

1.2 Equivalent Integral Equation

It is often useful to rewrite a first-order explicit intitial value problem as an equivalent *integral equation*. We provide the details of this equivalence in the following theorem:

Theorem 1.5. *If $G \subseteq \mathbb{R} \times \mathbb{K}^n$, $n \in \mathbb{N}$, and $f : G \longrightarrow \mathbb{K}^n$ is continuous, then, for each $(x_0, y_0) \in G$, the explicit n-dimensional first-order initial value problem*

$$y' = f(x, y), \tag{1.21a}$$

$$y(x_0) = y_0, \tag{1.21b}$$

is equivalent to the integral equation

$$y(x) = y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t\,, \tag{1.22}$$

in the sense that a continuous function $\phi : I \longrightarrow \mathbb{K}^n$, with $x_0 \in I \subseteq \mathbb{R}$ being a nontrivial interval, and ϕ satisfying

$$\big\{\big(x, \phi(x)\big) \in I \times \mathbb{K}^n : x \in I\big\} \subseteq G, \tag{1.23}$$

is a solution to (1.21) *in the sense of Def.* 1.3(a) *if, and only if,*

$$\forall_{x\in I} \quad \phi(x) = y_0 + \int_{x_0}^{x} f\big(t,\phi(t)\big)\,\mathrm{d}t\,, \tag{1.24}$$

i.e. if, and only if, ϕ is a solution *to the integral equation* (1.22).

Proof. Assume $I \subseteq \mathbb{R}$ with $x_0 \in I$ to be a nontrivial interval and $\phi : I \longrightarrow \mathbb{K}^n$ to be a continuous function, satisfying (1.23). If ϕ is a solution to (1.21), then ϕ is differentiable and the assumed continuity of f implies the continuity of ϕ'. In other words, each component ϕ_j of ϕ, $j = \{1,\dots,n\}$, is in $C^1(I,\mathbb{K})$. Thus, the fundamental theorem of calculus [Phi13a, Th. G.6(b)] applies, and [Phi13a, (G.16b)] yields

$$\forall_{x\in I} \quad \forall_{j\in\{1,\dots,n\}} \quad \phi_j(x) = \phi_j(x_0) + \int_{x_0}^{x} f_j\big(t,\phi(t)\big)\,\mathrm{d}t \overset{(1.21\text{b})}{=} y_{0,j} + \int_{x_0}^{x} f_j\big(t,\phi(t)\big)\,\mathrm{d}t\,, \tag{1.25}$$

proving ϕ satisfies (1.24). Conversely, if ϕ satisfies (1.24), then the validity of the initial condition (1.21b) is immediate. Moreover, as f and ϕ are continuous, so is the integrand function $t \mapsto f\big(t,\phi(t)\big)$ of (1.24). Thus, [Phi13a, Th. G.6(a)] applies to (the components of) ϕ, proving $\phi'(x) = f\big(x,\phi(x)\big)$ for each $x \in I$, proving ϕ is a solution to (1.21). ■

Example 1.6. Consider the situation of Th. 1.5. In the particularly simple special case, where f does not actually depend on y, but merely on x, the equivalence between (1.21) and (1.22) can be directly exploited to actually *solve* the initial value problem: If $f : I \longrightarrow \mathbb{K}^n$, where $I \subseteq \mathbb{R}$ is some nontrivial interval with $x_0 \in I$, then we obtain $\phi : I \longrightarrow \mathbb{K}^n$ to be a solution of (1.21) if, and only if,

$$\forall_{x\in I} \quad \phi(x) = y_0 + \int_{x_0}^{x} f(t)\,\mathrm{d}t\,, \tag{1.26}$$

i.e. if, and only if, ϕ is the antiderivative of f that satisfies the initial condition. In particular, in the present situation, ϕ as given by (1.26) is the *unique* solution to the initial value problem. Of course, depending on f, it can still be difficult to carry out the integral in (1.26).

1.3 Patching and Time Reversion

If solutions defined on different intervals fit together, then they can be patched to obtain a solution on the union of the two intervals:

Lemma 1.7 (Patching of Solutions)**.** *Let $k,n \in \mathbb{N}$. Given $G \subseteq \mathbb{R}\times\mathbb{K}^{kn}$ and $f : G \longrightarrow \mathbb{K}^n$, if $\phi : I \longrightarrow \mathbb{K}^n$ and $\psi : J \longrightarrow \mathbb{K}^n$ are both solutions to* (1.6)*, i.e. to*

$$y^{(k)} = f(x,y,y',\dots,y^{(k-1)}),$$

such that $I =]a,b]$, $J = [b,c[$, $a < b < c$, and such that

$$\forall_{j=0,\dots,k-1} \quad \phi^{(j)}(b) = \psi^{(j)}(b), \tag{1.27}$$

then

$$\sigma : I \cup J \longrightarrow \mathbb{K}^n, \quad \sigma(x) := \begin{cases} \phi(x) & \text{for } x \in I, \\ \psi(x) & \text{for } x \in J, \end{cases} \tag{1.28}$$

is also a solution to (1.6).

Proof. Since ϕ and ψ both are solutions to (1.6),

$$\Big\{ \big(x, \sigma(x), \sigma'(x), \dots, \sigma^{(k-1)}(x)\big) \in (I \cup J) \times \mathbb{K}^{kn} : x \in I \cup J \Big\} \subseteq G \tag{1.29}$$

must hold, where (1.27) guarantees that $\sigma^{(j)}(b)$ exists for each $j = 0, \dots, k-1$. Moreover, σ is k times differentiable at each $x \in I \cup J$, $x \neq b$, and

$$\underset{\substack{x \in I \cup J, \\ x \neq b}}{\forall} \quad \sigma^{(k)}(x) = f\big(x, \sigma(x), \sigma'(x), \dots, \sigma^{(k-1)}(x)\big). \tag{1.30}$$

However, at b, we also have (using the left-hand derivatives for ϕ and the right-hand derivatives for ψ)

$$\begin{aligned} \phi^{(k)}(b) &= f\big(b, \phi(b), \phi'(b), \dots, \phi^{(k-1)}(b)\big) \\ &= f\big(b, \psi(b), \psi'(b), \dots, \psi^{(k-1)}(b)\big) = \psi^{(k)}(b), \end{aligned} \tag{1.31}$$

which shows σ is k times differentiable and the equality of (1.30) also holds at $x = b$, completing the proof that σ is a solution. ■

It is sometimes useful to apply what is known as *time reversion*:

Definition 1.8. Let $k, n \in \mathbb{N}$, $G_f \subseteq \mathbb{R} \times \mathbb{K}^{kn}$, $f : G_f \longrightarrow \mathbb{K}^n$, and consider the ODE

$$y^{(k)} = f(x, y, y', \dots, y^{(k-1)}). \tag{1.32}$$

We call the ODE

$$y^{(k)} = g(x, y, y', \dots, y^{(k-1)}), \tag{1.33}$$

where

$$g : G_g \longrightarrow \mathbb{K}^n, \quad g(x, y) := (-1)^k f\big(-x, y_1, -y_2, \dots, (-1)^{k-1} y_k\big), \tag{1.34a}$$

$$G_g := \big\{(x, y) \in \mathbb{R} \times \mathbb{K}^{kn} : \big(-x, y_1, -y_2, \dots, (-1)^{k-1} y_k\big) \in G_f\big\}, \tag{1.34b}$$

the *time-reversed* version of (1.32).

Lemma 1.9 (Time Reversion)**.** *Let $k, n \in \mathbb{N}$, $G_f \subseteq \mathbb{R} \times \mathbb{K}^{kn}$, and $f : G_f \longrightarrow \mathbb{K}^n$.*

(a) *The time-reversed version of* (1.33) *is the original ODE, i.e.* (1.32).

(b) *If $-\infty \leq a < b \leq \infty$, then $\phi :]a, b[\longrightarrow \mathbb{K}^n$ is a solution to* (1.32) *if, and only if,*

$$\psi :]-b, -a[\longrightarrow \mathbb{K}^n, \quad \psi(x) := \phi(-x), \tag{1.35}$$

is a solution to the time-reversed version (1.33).

Proof. (a) is immediate from the definition of g in (1.34).

(b): Due to (a), it suffices to show if ϕ is a solution to (1.32), then ψ is a solution to (1.33). Clearly, if $x \in]-b,-a[$, then $-x \in]a,b[$. Moreover, noting

$$\underset{j=0,\dots,k}{\forall} \quad \underset{x\in]-b,-a[}{\forall} \quad \psi^{(j)}(x) = (-1)^j\,\phi^{(j)}(-x), \tag{1.36a}$$

one has

$$\underset{x\in]-b,-a[}{\forall} \quad \begin{aligned} \Big(\big(-x,\phi(-x),\phi'(-x),\dots,\phi^{(k-1)}(-x)\big) \in G_f \\ \Rightarrow \quad \big(x,\psi(x),\psi'(x),\dots,\psi^{(k-1)}(x)\big) \\ = \big(x,\phi(-x),-\phi'(-x),\dots,(-1)^{k-1}\,\phi^{(k-1)}(-x)\big) \in G_g\Big) \end{aligned} \tag{1.36b}$$

and

$$\underset{x\in]-b,-a[}{\forall} \quad \begin{aligned} \psi^{(k)}(x) &= (-1)^k\,f\big(-x,\phi(-x),\phi'(-x),\dots,\phi^{(k-1)}(-x)\big) \\ &= (-1)^k\,f\big(-x,\psi(x),-\psi'(x),\dots,(-1)^{k-1}\,\psi^{(k-1)}(x)\big) \\ &= g\big(x,\psi(x),\psi'(x),\dots,\psi^{(k-1)}(x)\big), \end{aligned} \tag{1.36c}$$

thereby establishing the case. ■

2 Elementary Solution Methods for 1-Dimensional First-Order ODE

2.1 Geometric Interpretation, Graphing

Geometrically, in the 1-dimensional real-valued case, the ODE (1.21a) provides a slope $y' = f(x,y)$ for every point (x,y). In other words, it provides a field of directions. The task is to find a differentiable function ϕ such that its graph has the prescribed slope in each point it contains. In certain simple cases, drawing the field of directions can help to *guess* the solutions of the ODE.

Example 2.1. Let $G := \mathbb{R}^+ \times \mathbb{R}$ and $f : G \longrightarrow \mathbb{R}$, $f(x,y) := y/x$, i.e. we consider the ODE $y' = y/x$. Drawing the field of directions leads to the idea that the solutions are functions whose graphs constitute rays, i.e. $\phi_c : \mathbb{R}^+ \longrightarrow \mathbb{R}$, $y = \phi_c(x) = c\,x$ with $c \in \mathbb{R}$. Indeed, one immediately verifies that each ϕ_c constitutes a solution to the ODE.

2.2 Linear ODE, Variation of Constants

Definition 2.2. Let $I \subseteq \mathbb{R}$ be an open interval and let $a, b : I \longrightarrow \mathbb{K}$ be continuous functions. An ODE of the form

$$y' = a(x)y + b(x) \tag{2.1}$$

is called a *linear ODE* of first order. It is called *homogeneous* if, and only if, $b \equiv 0$; it is called *inhomogeneous* if, and only if, it is not homogeneous.

Theorem 2.3 (Variation of Constants)**.** *Let $I \subseteq \mathbb{R}$ be an open interval and let $a, b : I \longrightarrow \mathbb{K}$ be continuous. Moreover, let $x_0 \in I$ and $c \in \mathbb{K}$. Then the linear ODE (2.1) has a unique solution $\phi : I \longrightarrow \mathbb{K}$ that satisfies the initial condition $y(x_0) = c$. This unique solution is given by*

$$\phi : I \longrightarrow \mathbb{K}, \quad \phi(x) = \phi_0(x) \left(c + \int_{x_0}^{x} \phi_0(t)^{-1}\, b(t)\, \mathrm{d}t \right), \tag{2.2a}$$

where

$$\phi_0 : I \longrightarrow \mathbb{K}, \quad \phi_0(x) = \exp\left(\int_{x_0}^{x} a(t)\, \mathrm{d}t \right) = e^{\int_{x_0}^{x} a(t)\, \mathrm{d}t}. \tag{2.2b}$$

Here, and in the following, ϕ_0^{-1} denotes $1/\phi_0$ and not the inverse function of ϕ_0 (which does not even necessarily exist).

Proof. We begin by noting that ϕ_0 according to (2.2b) is well-defined since a is assumed to be continuous, i.e., in particular, Riemann integrable on $[x_0, x]$. Moreover, the fundamental theorem of calculus [Phi13a, Th. G.6(a)] applies, showing ϕ_0 is differentiable with

$$\phi_0' : I \longrightarrow \mathbb{K}, \quad \phi_0'(x) = a(x) \exp\left(\int_{x_0}^{x} a(t)\, \mathrm{d}t \right) = a(x)\phi_0(x), \tag{2.3}$$

where Lem. A.1 of the Appendix was used as well. In particular, ϕ_0 is continuous. Since $\phi_0 \neq 0$ as well, ϕ_0^{-1} is also continuous. Moreover, as b is continuous by hypothesis, $\phi_0^{-1}\, b$ is continuous and, thus, Riemann integrable on $[x_0, x]$. Once again, [Phi13a, Th. G.6(a)] applies, yielding ϕ to be differentiable with

$$\begin{aligned} \phi' &: I \longrightarrow \mathbb{K}, \\ \phi'(x) &= \phi_0'(x) \left(c + \int_{x_0}^{x} \phi_0(t)^{-1}\, b(t)\, \mathrm{d}t \right) + \phi_0(x)\phi_0(x)^{-1}\, b(x) \\ &= a(x)\phi_0(x) \left(c + \int_{x_0}^{x} \phi_0(t)^{-1}\, b(t)\, \mathrm{d}t \right) + b(x) = a(x)\phi(x) + b(x), \end{aligned} \tag{2.4}$$

where the product rule of [Phi13a, Th. 9.6(c)] was used as well. Comparing (2.4) with (2.1) shows ϕ is a solution to (2.1). The computation

$$\phi(x_0) = \phi_0(x_0)\,(c + 0) = 1 \cdot c = c \tag{2.5}$$

verifies that ϕ satisfies the desired initial condition. It remains to prove uniqueness. To this end, let $\psi : I \longrightarrow \mathbb{K}$ be an arbitrary differentiable function that satisfies (2.1) as well as the initial condition $\psi(x_0) = c$. We have to show $\psi = \phi$. Since $\phi_0 \neq 0$, we can define $u := \psi/\phi_0$ and still have to verify

$$\forall_{x \in I} \quad u(x) = c + \int_{x_0}^{x} \phi_0(t)^{-1}\, b(t)\, \mathrm{d}t\,. \tag{2.6}$$

We obtain

$$a\,\phi_0\, u + b = a\,\psi + b = \psi' = (\phi_0\, u)' = \phi_0'\, u + \phi_0\, u' = a\,\phi_0\, u + \phi_0\, u', \tag{2.7}$$

implying $b = \phi_0 \, u'$ and $u' = \phi_0^{-1} \, b$. Thus, the fundamental theorem of calculus in the form [Phi13a, Th. G.6(b)] implies

$$\underset{x \in I}{\forall} \quad u(x) = u(x_0) + \int_{x_0}^{x} u'(t) \, \mathrm{d}t \, = c + \int_{x_0}^{x} \phi_0(t)^{-1} \, b(t) \, \mathrm{d}t \, , \tag{2.8}$$

thereby completing the proof. ∎

Corollary 2.4. *Let $I \subseteq \mathbb{R}$ be an open interval and let $a : I \longrightarrow \mathbb{K}$ be continuous. Moreover, let $x_0 \in I$ and $c \in \mathbb{K}$. Then the homogeneous linear ODE* (2.1) *(i.e. with $b \equiv 0$) has a unique solution $\phi : I \longrightarrow \mathbb{K}$ that satisfies the initial condition $y(x_0) = c$. This unique solution is given by*

$$\phi(x) = c \exp\left(\int_{x_0}^{x} a(t) \, \mathrm{d}t \right) = c \, e^{\int_{x_0}^{x} a(t) \, \mathrm{d}t} . \tag{2.9}$$

Proof. One immediately obtains (2.9) by setting $b \equiv 0$ in in (2.2). ∎

Remark 2.5. The name *variation of constants* for Th. 2.3 can be understood from comparing the solution (2.9) of the homogeneous linear ODE with the solution (2.2) of the general inhomogeneous linear ODE: One obtains (2.2) from (2.9) by *varying the constant* c, i.e. by replacing it with the function $x \mapsto c + \int_{x_0}^{x} \phi_0(t)^{-1} \, b(t) \, \mathrm{d}t$.

Example 2.6. Consider the ODE

$$y' = 2xy + x^3 \tag{2.10}$$

with initial condition $y(0) = c$, $c \in \mathbb{C}$. Comparing (2.10) with Def. 2.2, we observe we are facing an inhomogeneous linear ODE with

$$a : \mathbb{R} \longrightarrow \mathbb{R}, \quad a(x) := 2x, \tag{2.11a}$$

$$b : \mathbb{R} \longrightarrow \mathbb{R}, \quad b(x) := x^3. \tag{2.11b}$$

From Cor. 2.4, we obtain the solution $\phi_{0,c}$ to the homogeneous version of (2.10):

$$\phi_{0,c} : \mathbb{R} \longrightarrow \mathbb{C}, \quad \phi_{0,c}(x) = c \exp\left(\int_{0}^{x} a(t) \, \mathrm{d}t \right) = c e^{x^2} . \tag{2.12}$$

The solution to (2.10) is given by (2.2a):

$$\begin{aligned} &\phi : \mathbb{R} \longrightarrow \mathbb{C}, \\ \phi(x) &= e^{x^2} \left(c + \int_{0}^{x} e^{-t^2} \, t^3 \, \mathrm{d}t \right) = e^{x^2} \left(c + \left[-\frac{1}{2}(t^2 + 1) e^{-t^2} \right]_{0}^{x} \right) \\ &= e^{x^2} \left(c + \frac{1}{2} - \frac{1}{2}(x^2 + 1) e^{-x^2} \right) = \left(c + \frac{1}{2} \right) e^{x^2} - \frac{1}{2}(x^2 + 1). \end{aligned} \tag{2.13}$$

2.3 Separation of Variables

If the ODE (1.21a) has the particular form

$$y' = f(x)g(y), \tag{2.14}$$

with one-dimensional real-valued continuous functions f and g, and $g(y) \neq 0$, then it can be solved by a method known as *separation of variables*:

Theorem 2.7. *Let $I, J \subseteq \mathbb{R}$ be (bounded or unbounded) open intervals and suppose that $f : I \longrightarrow \mathbb{R}$ and $g : J \longrightarrow \mathbb{R}$ are continuous with $g(y) \neq 0$ for each $y \in J$. For each $(x_0, y_0) \in I \times J$, consider the initial value problem consisting of the ODE* (2.14) *together with the initial condition*

$$y(x_0) = y_0. \tag{2.15}$$

Define the functions

$$F : I \longrightarrow \mathbb{R}, \quad F(x) := \int_{x_0}^{x} f(t)\,\mathrm{d}t\,, \quad G : J \longrightarrow \mathbb{R}, \quad G(y) := \int_{y_0}^{y} \frac{\mathrm{d}t}{g(t)}. \tag{2.16}$$

(a) Uniqueness: *On each open interval $I' \subseteq I$ satisfying $x_0 \in I'$ and $F(I') \subseteq G(J)$, the initial value problem consisting of* (2.14) *and* (2.15) *has a unique solution. This unique solution is given by*

$$\phi : I' \longrightarrow \mathbb{R}, \quad \phi(x) := G^{-1}\big(F(x)\big), \tag{2.17}$$

where $G^{-1} : G(J) \longrightarrow J$ is the inverse function of G on $G(J)$.

(b) Existence: *There exists an open interval $I' \subseteq I$ satisfying $x_0 \in I'$ and $F(I') \subseteq G(J)$, i.e. an I' such that* (a) *applies.*

Proof. (a): We begin by proving G has a differentiable inverse function $G^{-1} : G(J) \longrightarrow J$. According to the fundamental theorem of calculus [Phi13a, Th. 10.19(a)], G is differentiable with $G' = 1/g$. Since g is continuous and nonzero, G is even C^1. If $G'(y_0) = 1/g(y_0) > 0$, then G is strictly increasing on J (due to the intermediate value theorem [Phi13a, Th. 7.57]; $g(y_0) > 0$, the continuity of g, and $g \neq 0$ imply that $g > 0$ on J). Analogously, if $G'(y_0) = 1/g(y_0) < 0$, then G is strictly decreasing on J. In each case, G has a differentiable inverse function on $G(J)$ by [Phi13a, Th. 9.8].

In the next step, we verify that (2.17) does, indeed, define a solution to (2.14) and (2.15). The assumption $F(I') \subseteq G(J)$ and the existence of G^{-1} as shown above provide that ϕ is well-defined by (2.17). Verifying (2.15) is quite simple: $\phi(x_0) = G^{-1}(F(x_0)) = G^{-1}(0) = y_0$. To see ϕ to be a solution of (2.14), notice that (2.17) implies $F = G \circ \phi$ on I'. Thus, we can apply the chain rule to obtain the derivative of $F = G \circ \phi$ on I':

$$\underset{x \in I'}{\forall} \quad f(x) = F'(x) = G'\big(\phi(x)\big)\,\phi'(x) = \frac{\phi'(x)}{g\big(\phi(x)\big)}, \tag{2.18}$$

showing ϕ satisfies (2.14).

We now proceed to show that each solution $\phi : I' \longrightarrow \mathbb{R}$ to (2.14) that satisfies (2.15) must also satisfy (2.17). Since ϕ is a solution to (2.14),

$$\frac{\phi'(x)}{g\big(\phi(x)\big)} = f(x) \quad \text{for each } x \in I'. \tag{2.19}$$

Integrating (2.19) yields

$$\int_{x_0}^{x} \frac{\phi'(t)}{g\big(\phi(t)\big)} \, \mathrm{d}t = \int_{x_0}^{x} f(t) \, \mathrm{d}t = F(x) \quad \text{for each } x \in I'. \tag{2.20}$$

Using the change of variables formula of [Phi13a, Th. 10.24] in the left-hand side of (2.20), allows one to replace $\phi(t)$ by the new integration variable u (note that each solution $\phi : I' \longrightarrow \mathbb{R}$ to (2.14) is in $C^1(I')$ since f and g are presumed continuous). Thus, we obtain from (2.20):

$$F(x) = \int_{\phi(x_0)}^{\phi(x)} \frac{\mathrm{d}u}{g(u)} = \int_{y_0}^{\phi(x)} \frac{\mathrm{d}u}{g(u)} = G\big(\phi(x)\big) \quad \text{for each } x \in I'. \tag{2.21}$$

Applying G^{-1} to (2.21) establishes ϕ satisfies (2.17).

(b): During the proof of (a), we have already seen G to be either strictly increasing or strictly decreasing. As $G(y_0) = 0$, this implies the existence of $\epsilon > 0$ such that $]-\epsilon, \epsilon[\subseteq G(J)$. The function F is differentiable and, in particular, continuous. Since $F(x_0) = 0$, there is $\delta > 0$ such that, for $I' :=]x_0 - \delta, x_0 + \delta[$, one has $F(I') \subseteq]-\epsilon, \epsilon[\subseteq G(J)$ as desired. ■

Example 2.8. Consider the ODE

$$y' = -\frac{y}{x} \quad \text{on } I \times J := \mathbb{R}^+ \times \mathbb{R}^+ \tag{2.22}$$

with the initial condition $y(1) = c$ for some given $c \in \mathbb{R}^+$. Introducing functions

$$f : \mathbb{R}^+ \longrightarrow \mathbb{R}, \quad f(x) := -\frac{1}{x}, \quad g : \mathbb{R}^+ \longrightarrow \mathbb{R}, \quad g(y) := y, \tag{2.23}$$

one sees that Th. 2.7 applies. To compute the solution $\phi = G^{-1} \circ F$, we first have to determine F and G:

$$F : \mathbb{R}^+ \longrightarrow \mathbb{R}, \qquad F(x) = \int_1^x f(t) \, \mathrm{d}t = -\int_1^x \frac{\mathrm{d}t}{t} = -\ln x, \tag{2.24a}$$

$$G : \mathbb{R}^+ \longrightarrow \mathbb{R}, \qquad G(y) = \int_c^y \frac{\mathrm{d}t}{g(t)} = \int_c^y \frac{\mathrm{d}t}{t} = \ln \frac{y}{c}. \tag{2.24b}$$

Here, we can choose $I' = I = \mathbb{R}^+$, because $F(\mathbb{R}^+) = \mathbb{R} = G(\mathbb{R}^+)$. That means ϕ is defined on the entire interval I. The inverse function of G is given by

$$G^{-1} : \mathbb{R} \longrightarrow \mathbb{R}^+, \quad G^{-1}(t) = c \, e^t. \tag{2.25}$$

Finally, we get

$$\phi : \mathbb{R}^+ \longrightarrow \mathbb{R}, \quad \phi(x) = G^{-1}\big(F(x)\big) = c\, e^{-\ln x} = \frac{c}{x}. \tag{2.26}$$

The uniqueness part of Th. 2.7 further tells us the above initial value problem can have no solution different from ϕ.

—

The advantage of using Th. 2.7 as in the previous example, by computing the relevant functions F, G, and G^{-1}, is that it is mathematically rigorous. In particular, one can be sure one has found the unique solution to the ODE with initial condition. However, in practice, it is often easier to use the following heuristic (not entirely rigorous) procedure. In the end, in most cases, one can easily check by differentiation that the function found is, indeed, a solution to the ODE with initial condition. However, one does not know uniqueness without further investigations (general results such as Th. 3.15 below can often help). One also has to determine on which interval the found solution is defined. On the other hand, as one is usually interested in choosing the interval as large as possible, the optimal choice is not always obvious when using Th. 2.7, either.

The heuristic procedure is as follows: Start with the ODE (2.14) written in the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y). \tag{2.27a}$$

Multiply by $\mathrm{d}x$ and divide by $g(y)$ (i.e. *separate the variables*):

$$\frac{\mathrm{d}y}{g(y)} = f(x)\,\mathrm{d}x\,. \tag{2.27b}$$

Integrate:

$$\int \frac{\mathrm{d}y}{g(y)} = \int f(x)\,\mathrm{d}x\,. \tag{2.27c}$$

Change the integration variables and supply the appropriate upper and lower limits for the integrals (according to the initial condition):

$$\int_{y_0}^{y} \frac{\mathrm{d}t}{g(t)} = \int_{x_0}^{x} f(t)\,\mathrm{d}t\,. \tag{2.27d}$$

Solve this equation for y, set $\phi(x) := y$, check by differentiation that ϕ is, indeed, a solution to the ODE, and determine the largest interval I' such that $x_0 \in I'$ and such that ϕ is defined on I'. The use of this heuristic procedure is demonstrated by the following example:

Example 2.9. Consider the ODE

$$y' = -y^2 \quad \text{on } I \times J := \mathbb{R} \times \mathbb{R} \tag{2.28}$$

with the initial condition $y(x_0) = y_0$ for given values $x_0, y_0 \in \mathbb{R}$. We manipulate (2.28) according to the heuristic procedure described in (2.27) above:

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} = -y^2 \quad &\rightsquigarrow \quad -y^{-2}\,\mathrm{d}y \;=\; \mathrm{d}x \quad \rightsquigarrow \quad -\int y^{-2}\,\mathrm{d}y \;=\; \int \mathrm{d}x \\
&\rightsquigarrow \quad -\int_{y_0}^{y} t^{-2}\,\mathrm{d}t \;=\; \int_{x_0}^{x} \mathrm{d}t \quad \rightsquigarrow \quad \left[\frac{1}{t}\right]_{y_0}^{y} = [t]_{x_0}^{x} \quad \rightsquigarrow \quad \frac{1}{y} - \frac{1}{y_0} = x - x_0 \\
&\rightsquigarrow \quad \phi(x) = y = \frac{y_0}{1 + (x - x_0)\,y_0}. \qquad (2.29)
\end{aligned}$$

Clearly, $\phi(x_0) = y_0$. Moreover,

$$\phi'(x) = -\frac{y_0^2}{\big(1 + (x - x_0)\,y_0\big)^2} = -\big(\phi(x)\big)^2, \qquad (2.30)$$

i.e. ϕ does, indeed, provide a solution to (2.28). If $y_0 = 0$, then $\phi \equiv 0$ is defined on the entire interval $I = \mathbb{R}$. If $y_0 \neq 0$, then the denominator of $\phi(x)$ has a zero at $x = (x_0 y_0 - 1)/y_0$, and ϕ is not defined on all of $\mathbb{R}$. In that case, if $y_0 > 0$, then $x_0 > (x_0 y_0 - 1)/y_0 = x_0 - 1/y_0$ and the maximal open interval for ϕ to be defined on is $I' =]x_0 - 1/y_0, \infty[$; if $y_0 < 0$, then $x_0 < (x_0 y_0 - 1)/y_0 = x_0 - 1/y_0$ and the maximal open interval for ϕ to be defined on is $I' =]-\infty, x_0 - 1/y_0[$. Note that the formula for ϕ obtained by (2.29) works for $y_0 = 0$ as well, even though not every previous expression in (2.29) is meaningful for $y_0 = 0$ and, also, Th. 2.7 does not apply to (2.28) for $y_0 = 0$. In the present example, the subsequent Th. 3.15 does, indeed, imply ϕ to be the unique solution to the initial value problem on I'.

2.4 Change of Variables

To solve an ODE, it can be useful to transform it into an equivalent ODE, using a so-called *change of variables*. If one already knows how to solve the transformed ODE, then the equivalence allows one to also solve the original ODE. We first present the following Th. 2.10, which constitutes the base for the change of variables technique, followed by examples, where the technique is applied.

Theorem 2.10. *Let $G \subseteq \mathbb{R} \times \mathbb{K}^n$ be open, $n \in \mathbb{N}$, $f : G \longrightarrow \mathbb{K}^n$, and $(x_0, y_0) \in G$. Define*

$$\underset{x \in \mathbb{R}}{\forall} \quad G_x := \{y \in \mathbb{K}^n : (x, y) \in G\} \qquad (2.31)$$

and assume the change of variables *function $T : G \longrightarrow \mathbb{K}^n$ is differentiable and such that*

$$\underset{G_x \neq \emptyset}{\forall} \quad \Big(T_x := T(x, \cdot) : G_x \longrightarrow T_x(G_x), \quad T_x(y) := T(x, y), \quad \text{is a diffeomorphism}\Big), \qquad (2.32)$$

i.e. T_x *is invertible and both* T_x *and* T_x^{-1} *are differentiable. Then the first-order initial value problems*

$$y' = f(x,y), \tag{2.33a}$$
$$y(x_0) = y_0, \tag{2.33b}$$

and

$$y' = \left(DT_x^{-1}(y)\right)^{-1} f\big(x, T_x^{-1}(y)\big) + \partial_x T\big(x, T_x^{-1}(y)\big), \tag{2.34a}$$
$$y(x_0) = T(x_0, y_0), \tag{2.34b}$$

are equivalent in the following sense:

(a) *A differentiable function* $\phi : I \longrightarrow \mathbb{K}^n$, *where* $I \subseteq \mathbb{R}$ *is a nontrivial interval, is a solution to* (2.33a) *if, and only if, the function*

$$\mu : I \longrightarrow \mathbb{K}^n, \quad \mu(x) := (T_x \circ \phi)(x) = T\big(x, \phi(x)\big), \tag{2.35}$$

is a solution to (2.34a).

(b) *A differentiable function* $\phi : I \longrightarrow \mathbb{K}^n$, *where* $I \subseteq \mathbb{R}$ *is a nontrivial interval, is a solution to* (2.33) *if, and only if, the function of* (2.35) *is a solution to* (2.34).

Proof. We start by noting that the assumption of G being open clearly implies each G_x, $x \in \mathbb{R}$, to be open as well, which, in turn, implies $T_x(G_x)$ to be open, even though this is not as obvious[1]. Next, for each $x \in \mathbb{R}$ such that $G_x \neq \emptyset$, we can apply the chain rule [Phi13b, Th. 2.28] to $T_x \circ T_x^{-1} = \mathrm{Id}$ to obtain

$$\underset{y \in T_x(G_x)}{\forall} \quad DT_x\big(T_x^{-1}(y)\big) \circ DT_x^{-1}(y) = \mathrm{Id} \tag{2.36}$$

and, thus, each $DT_x^{-1}(y)$ is invertible with

$$\underset{y \in T_x(G_x)}{\forall} \quad \left(DT_x^{-1}(y)\right)^{-1} = DT_x\big(T_x^{-1}(y)\big). \tag{2.37}$$

Consider ϕ and μ as in (a) and notice that (2.35) implies

$$\underset{x \in I}{\forall} \quad \phi(x) = T_x^{-1}(\mu(x)). \tag{2.38}$$

Moreover, the differentiability of ϕ and T imply differentiability of μ by the chain rule, which also yields

$$\underset{x \in I}{\forall} \quad \begin{aligned} \mu'(x) &= DT\big(x, \phi(x)\big) \begin{pmatrix} 1 \\ \phi'(x) \end{pmatrix} \\ &= DT_x(\phi(x))\, \phi'(x) + \partial_x T\big(x, \phi(x)\big). \end{aligned} \tag{2.39}$$

[1] If T_x is a continuously differentiable map, then this is related to the inverse function theorem (see, e.g. [Phi13b, Cor. C.9]); it is still true if T_x is merely continuous and injective, but then it is the invariance of domain theorem of algebraic topology [Oss09, 5.6.15], which is equivalent to the Brouwer fixed-point theorem [Oss09, 5.6.10], and is much harder to prove.

To prove (a), first assume $\phi : I \longrightarrow \mathbb{K}^n$ to be a solution of (2.33a). Then, for each $x \in I$,

$$\begin{aligned}
\mu'(x) &\overset{(2.39),(2.33a)}{=} DT_x(\phi(x))\, f\big(x, \phi(x)\big) + \partial_x T\big(x, \phi(x)\big) \\
&\overset{(2.38)}{=} DT_x\big(T_x^{-1}(\mu(x))\big)\, f\big(x, T_x^{-1}(\mu(x))\big) + \partial_x T\big(x, T_x^{-1}(\mu(x))\big) \\
&\overset{(2.37)}{=} \Big(DT_x^{-1}(\mu(x))\Big)^{-1} f\big(x, T_x^{-1}(\mu(x))\big) + \partial_x T\big(x, T_x^{-1}(\mu(x))\big), \qquad (2.40)
\end{aligned}$$

showing μ satisfies (2.34a). Conversely, assume μ to be a solution to (2.34a). Then, for each $x \in I$,

$$\begin{aligned}
&\Big(DT_x^{-1}(\mu(x))\Big)^{-1} f\big(x, T_x^{-1}(\mu(x))\big) + \partial_x T\big(x, T_x^{-1}(\mu(x))\big) \\
&\overset{(2.34a)}{=} \mu'(x) \overset{(2.39)}{=} DT_x(\phi(x))\, \phi'(x) + \partial_x T\big(x, \phi(x)\big). \qquad (2.41)
\end{aligned}$$

Using (2.38), one can subtract the second summand from (2.41). Multiplying the result by $DT_x^{-1}(\mu(x))$ from the left and taking into account (2.37) then provides

$$\underset{x \in I}{\forall} \quad \phi'(x) = f\big(x, T_x^{-1}(\mu(x))\big) \overset{(2.38)}{=} f\big(x, \phi(x)\big), \qquad (2.42)$$

showing ϕ satisfies (2.33a).

It remains to prove (b). If ϕ satisfies (2.33), then μ satisfies (2.34a) by (a). Moreover, $\mu(x_0) = T\big(x_0, \phi(x_0)\big) = T(x_0, y_0)$, i.e. μ satisfies (2.34b) as well. Conversely, assume μ satisfies (2.34). Then ϕ satisfies (2.33a) by (a). Moreover, by (2.38), $\phi(x_0) = T_{x_0}^{-1}(\mu(x_0)) = T_{x_0}^{-1}(T(x_0, y_0)) = y_0$, showing ϕ satisfies (2.33b) as well. ■

As a first application of Th. 2.10, we prove the following theorem about so-called *Bernoulli differential equations*:

Theorem 2.11. *Consider the* Bernoulli differential equation

$$y' = f(x, y) := a(x)\, y + b(x)\, y^\alpha, \qquad (2.43a)$$

where $\alpha \in \mathbb{R} \setminus \{0, 1\}$, the functions $a, b : I \longrightarrow \mathbb{R}$ are continuous and defined on an open interval $I \subseteq \mathbb{R}$, and $f : I \times \mathbb{R}^+ \longrightarrow \mathbb{R}$. For (2.43a)*, we add the initial condition*

$$y(x_0) = y_0, \quad (x_0, y_0) \in I \times \mathbb{R}^+, \qquad (2.43b)$$

and, furthermore, we also consider the corresponding linear initial value problem

$$\begin{aligned}
y' &= (1 - \alpha)\big(a(x)\, y + b(x)\big), \qquad (2.44a) \\
y(x_0) &= y_0^{1-\alpha}, \qquad (2.44b)
\end{aligned}$$

with its unique solution $\psi : I \longrightarrow \mathbb{R}$ given by Th. 2.3.

(a) Uniqueness: *On each open interval $I' \subseteq I$ satisfying $x_0 \in I'$ and $\psi > 0$ on I', the Bernoulli initial value problem* (2.43) *has a unique solution. This unique solution is given by*

$$\phi : I' \longrightarrow \mathbb{R}^+, \quad \phi(x) := \big(\psi(x)\big)^{\frac{1}{1-\alpha}}. \tag{2.45}$$

(b) Existence: *There exists an open interval $I' \subseteq I$ satisfying $x_0 \in I'$ and $\psi > 0$ on I', i.e. an I' such that* (a) *applies.*

Proof. (b) is immediate from Th. 2.3, since $\psi(x_0) = y_0 > 0$ and ψ is continuous.

To prove (a), we apply Th. 2.10 with the change of variables

$$T : I \times \mathbb{R}^+ \longrightarrow \mathbb{R}^+, \quad T(x,y) := y^{1-\alpha}. \tag{2.46}$$

Then $T \in C^1(I \times \mathbb{R}^+, \mathbb{R})$ with $\partial_x T \equiv 0$ and $\partial_y T(x,y) = (1-\alpha)\, y^{-\alpha}$. Moreover,

$$\underset{x\in I}{\forall} \quad T_x = S, \quad S : \mathbb{R}^+ \longrightarrow \mathbb{R}^+, \quad S(y) := y^{1-\alpha}, \tag{2.47}$$

which is differentiable with the differentiable inverse function $S^{-1} : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$, $S^{-1}(y) = y^{\frac{1}{1-\alpha}}$, $DS^{-1}(y) = (S^{-1})'(y) = \frac{1}{1-\alpha}\, y^{\frac{\alpha}{1-\alpha}}$. Thus, (2.34a) takes the form

$$\begin{aligned} y' &= \Big(DT_x^{-1}(y)\Big)^{-1} f\big(x, T_x^{-1}(y)\big) + \partial_x T\big(x, T_x^{-1}(y)\big) \\ &= (1-\alpha)\, y^{-\frac{\alpha}{1-\alpha}} \Big(a(x)\, y^{\frac{1}{1-\alpha}} + b(x)\, \big(y^{\frac{1}{1-\alpha}}\big)^{\alpha}\Big) + 0 \\ &= (1-\alpha)\, \big(a(x)\, y + b(x)\big). \end{aligned} \tag{2.48}$$

Thus, if $I' \subseteq I$ is such that $x_0 \in I'$ and $\psi > 0$ on I', then Th. 2.10 says ϕ defined by (2.45) must be a solution to (2.43) (note that the differentiability of ψ implies the differentiability of ϕ). On the other hand, if $\lambda : I' \longrightarrow \mathbb{R}^+$ is an arbitrary solution to (2.43), then Th. 2.10 states $\mu := S \circ \lambda = \lambda^{1-\alpha}$ to be a solution to (2.44). The uniqueness part of Th. 2.3 then yields $\lambda^{1-\alpha} = \psi{\restriction}_{I'} = \phi^{1-\alpha}$, i.e. $\lambda = \phi$. ■

Example 2.12. Consider the initial value problem

$$y' = f(x,y) := i - \frac{1}{ix - y + 2}, \tag{2.49a}$$

$$y(1) = i, \tag{2.49b}$$

where $f : G \longrightarrow \mathbb{C}$, $G := \{(x,y) \in \mathbb{R} \times \mathbb{C} : ix - y + 2 \neq 0\}$ (G is open as the continuous preimage of the open set $\mathbb{C} \setminus \{0\}$). We apply the change of variables

$$T : G \longrightarrow \mathbb{C}, \quad T(x,y) := ix - y. \tag{2.50}$$

Then, $T \in C^1(G, \mathbb{C})$. Moreover, for each $x \in \mathbb{R}$,

$$G_x = \{y \in \mathbb{C} : (x,y) \in G\} = \mathbb{C} \setminus \{ix + 2\} \tag{2.51}$$

and we have the diffeomorphisms

$$T_x : \mathbb{C} \setminus \{ix+2\} \longrightarrow \mathbb{C} \setminus \{-2\}, \qquad T_x(y) = ix - y, \tag{2.52a}$$

$$T_x^{-1} : \mathbb{C} \setminus \{-2\} \longrightarrow \mathbb{C} \setminus \{ix+2\}, \qquad T_x^{-1}(y) = ix - y. \tag{2.52b}$$

To obtain the transformed equation, we compute the right-hand side of (2.34a)

$$\begin{aligned} &\Big(DT_x^{-1}(y)\Big)^{-1} f\big(x, T_x^{-1}(y)\big) + \partial_x T\big(x, T_x^{-1}(y)\big) \\ &= (-1) \cdot \left(i - \frac{1}{y+2}\right) + i = \frac{1}{y+2}. \end{aligned} \tag{2.53}$$

Thus, the transformed initial value problem is

$$y' = \frac{1}{y+2}, \tag{2.54a}$$

$$y(1) = T(1, i) = i - i = 0. \tag{2.54b}$$

Using seperation of variables, one finds the solution

$$\mu :]-1, \infty[\longrightarrow]-2, \infty[, \quad \mu(x) := \sqrt{2x+2} - 2, \tag{2.55}$$

to (2.54). Then Th. 2.10 implies that

$$\phi :]-1, \infty[\longrightarrow \mathbb{C}, \quad \phi(x) := T_x^{-1}\big(\mu(x)\big) = ix - \sqrt{2x+2} + 2, \tag{2.56}$$

is a solution to (2.49) (that ϕ is a solution to (2.49) can now also easily be checked directly). It will become clear from Th. 3.15 below that ϕ and ψ are also the unique solutions to their respective initial value problems.

—

Finding a suitable change of variables to transform a given ODE such that one is in a position to solve the transformed ODE is an art, i.e. it can be very difficult to spot a useful transformation, and it takes a lot of practise and experience.

Remark 2.13. Somewhat analogous to the situation described in the paragraph before (2.27) regarding the separation of variables technique, in practise, one frequently uses a heuristic procedure to apply a change of variables, rather than appealing to the rigorous Th. 2.10. For the initial value problem $y' = f(x, y)$, $y(x_0) = y_0$, this heuristic procedure proceeds as follows:

(1) One introduces the new variable $z := T(x, y)$ and then computes z', i.e. the derivative of the function $x \mapsto z(x) = T(x, y(x))$.

(2) In the result of (1), one eliminates all occurrences of the variable y by first replacing y' by $f(x, y)$ and then replacing y by $T_x^{-1}(z)$, where $T_x(y) := T(x, y) = z$ (i.e. one has to solve the equation $z = T(x, y)$ for y). One thereby obtains the transformed initial value problem problem $z' = g(x, z)$, $z(x_0) = T(x_0, y_0)$, with a suitable function g.

(3) One solves the transformed initial value problem to obtain a solution μ, and then $x \mapsto \phi(x) := T_x^{-1}\big(\mu(x)\big)$ yields a candidate for a solution to the original initial value problem.

(4) One checks that ϕ is, indeed, a solution to $y' = f(x,y)$, $y(x_0) = y_0$.

Example 2.14. Consider

$$f : \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}, \quad f(x,y) := 1 + \frac{y}{x} + \frac{y^2}{x^2}, \tag{2.57}$$

and the initial value problem

$$y' = f(x,y), \quad y(1) = 0. \tag{2.58}$$

We introduce the change of variables $z := T(x,y) := y/x$ and proceed according to the steps of Rem. 2.13. According to (1), we compute, using the quotient rule,

$$z'(x) = \frac{y'(x)\,x - y(x)}{x^2}. \tag{2.59}$$

According to (2), we replace $y'(x)$ by $f(x,y)$ and then replace y by $T_x^{-1}(z) = xz$ to obtain the transformed initial value problem

$$z' = \frac{1}{x}\left(1 + \frac{y}{x} + \frac{y^2}{x^2}\right) - \frac{y}{x^2} = \frac{1}{x}\,(1 + z + z^2) - \frac{z}{x} = \frac{1+z^2}{x}, \quad z(1) = 0/1 = 0. \tag{2.60}$$

According to (3), we next solve (2.60), e.g. by seperation of variables, to obtain the solution

$$\mu : \left]e^{-\frac{\pi}{2}}, e^{\frac{\pi}{2}}\right[\longrightarrow \mathbb{R}, \quad \mu(x) := \tan\ln x, \tag{2.61}$$

of (2.60), and

$$\phi : \left]e^{-\frac{\pi}{2}}, e^{\frac{\pi}{2}}\right[\longrightarrow \mathbb{R}, \quad \phi(x) := x\,\mu(x) = x\,\tan\ln x, \tag{2.62}$$

as a candidate for a solution to (2.58). Finally, according to (4), we check that ϕ is, indeed, a solution to (2.58): Due to $\phi(1) = 1 \cdot \tan 0 = 0$, ϕ satisfies the initial condition, and due to

$$\begin{aligned} \phi'(x) &= \tan\ln x + x\,\frac{1}{x}\,(1 + \tan^2\ln x) = 1 + \tan\ln x + \tan^2\ln x \\ &= 1 + \frac{\phi(x)}{x} + \frac{\phi^2(x)}{x^2}, \end{aligned} \tag{2.63}$$

ϕ satisfies the ODE.

3 General Theory

3.1 Equivalence Between Higher-Order ODE and Systems of First-Order ODE

It turns out that each one-dimensional kth-order ODE is equivalent to a system of k first-order ODE; more generally, that each n-dimensional kth-order ODE is equivalent to a kn-dimensional first-order ODE (i.e. to a system of kn one-dimensional first-order ODE). Even though, in this class, we will mainly consider explicit ODE, we provide the equivalence also for the implicit case, as the proof is essentially the same (the explicit case is included as a special case).

Theorem 3.1. *In the situation of Def.* 1.2(a), *i.e.* $U \subseteq \mathbb{R} \times \mathbb{K}^{(k+1)n}$ *and* $F : U \longrightarrow \mathbb{K}^n$, *plus* $(x_0, y_{0,0}, \dots, y_{0,k-1}) \in \mathbb{R} \times \mathbb{K}^{kn}$, *consider the kth-order initial value problem*

$$F(x, y, y', \dots, y^{(k)}) = 0, \tag{3.1a}$$

$$\underset{j \in \{0,\dots,k-1\}}{\forall} \quad y^{(j)}(x_0) = y_{0,j}, \tag{3.1b}$$

and the first-order initial value problem

$$\begin{aligned} y_1' - y_2 &= 0, \\ y_2' - y_3 &= 0, \\ &\vdots \\ y_{k-1}' - y_k &= 0, \\ F(x, y_1, \dots, y_k, y_k') &= 0, \end{aligned} \tag{3.2a}$$

$$y(x_0) = \begin{pmatrix} y_{0,0} \\ \vdots \\ y_{0,k-1} \end{pmatrix} \tag{3.2b}$$

(note that the unknown function y in (3.1) *is $\mathbb{K}^n$-valued, whereas the unknown function y in* (3.2) *is $\mathbb{K}^{kn}$-valued). Then both initial value problems are equivalent in the following sense:*

(a) *If $\phi : I \longrightarrow \mathbb{K}^n$ is a solution to* (3.1), *then*

$$\Phi : I \longrightarrow \mathbb{K}^{kn}, \quad \Phi := \begin{pmatrix} \phi \\ \phi' \\ \vdots \\ \phi^{(k-1)} \end{pmatrix}, \tag{3.3}$$

is a solution to (3.2).

(b) *If $\Phi : I \longrightarrow \mathbb{K}^{kn}$ is a solution to* (3.2), *then $\phi := \Phi_1$ (which is $\mathbb{K}^n$-valued) is a solution to* (3.1).

Proof. We rewrite (3.2a) as
$$G(x,y,y')=0, \tag{3.4}$$
where
$$\begin{aligned}
G:\, V &\longrightarrow \mathbb{K}^{kn},\\
V &:= \big\{(x,y,z)\in\mathbb{R}\times\mathbb{K}^{kn}\times\mathbb{K}^{kn}:\,(x,y,z_k)\in U\big\}\subseteq \mathbb{R}\times\mathbb{K}^{kn}\times\mathbb{K}^{kn},\\
G_1(x,y,z) &:= z_1-y_2,\\
G_2(x,y,z) &:= z_2-y_3,\\
&\ \vdots\\
G_{k-1}(x,y,z) &:= z_{k-1}-y_k,\\
G_k(x,y,z) &:= F(x,y,z_k).
\end{aligned} \tag{3.5}$$

(a): As a solution to (3.1), ϕ is k times differentiable and Φ is well-defined. Then (3.1b) implies (3.2b), since
$$\Phi(x_0)=\begin{pmatrix}\phi(x_0)\\ \phi'(x_0)\\ \vdots\\ \phi^{(k-1)}(x_0)\end{pmatrix}\overset{(3.1b)}{=}\begin{pmatrix}y_{0,0}\\ \vdots\\ y_{0,k-1}\end{pmatrix}.$$

Next, Def. 1.2(a)(i) for ϕ implies Def. 1.2(a)(i) for Φ, since
$$\begin{aligned}
&\big\{\big(x,\Phi(x),\Phi'(x)\big)\in I\times\mathbb{K}^{kn}\times\mathbb{K}^{kn}:\,x\in I\big\}\\
&\overset{(3.3)}{=}\big\{\big(x,\phi(x),\phi'(x),\dots,\phi^{(k-1)}(x),\phi'(x),\dots,\phi^{(k)}(x)\big)\in I\times\mathbb{K}^{kn}\times\mathbb{K}^{kn}:\,x\in I\big\}\\
&\overset{(x,\phi(x),\dots,\phi^{(k)}(x))\in U}{\subseteq}\quad V.
\end{aligned} \tag{3.6}$$

The definition of Φ in (3.3) implies
$$\underset{j\in\{1,\dots,k-1\}}{\forall}\quad \Phi_j'=(\phi^{(j-1)})'=\phi^{(j)}=\Phi_{j+1},$$
showing Φ satisfies the first $k-1$ equations of (3.2a). As
$$\Phi_k'(x)=(\phi^{(k-1)})'(x)=\phi^{(k)}(x)$$
and, thus,
$$\underset{x\in I}{\forall}\quad G_k\big(x,\Phi(x),\Phi'(x)\big)=F\big(x,\phi(x),\phi'(x),\dots,\phi^{(k)}(x)\big)\overset{(3.1a)}{=}0, \tag{3.7}$$
Φ also satisfies the last equation of (3.2a).

(b): As Φ is a solution to (3.2), the first $k-1$ equations of (3.2a) imply
$$\underset{j\in\{1,\dots,k-1\}}{\forall}\quad \Phi_{j+1}=\Phi_j'\overset{\phi=\Phi_1}{=}\phi^{(j)},$$

i.e. ϕ is k times differentiable and Φ has, once again, the form (3.3) (note $\Phi_1 = \phi$ by the definition of ϕ). Then, clearly, (3.2b) implies (3.1b), and Def. 1.2(a)(i) for Φ implies Def. 1.2(a)(i) for ϕ:

$$\big\{\big(x, \Phi(x), \Phi'(x)\big) \in I \times \mathbb{K}^{kn} \times \mathbb{K}^{kn} : x \in I\big\} \subseteq V$$

and the definition of V in (3.5) imply

$$\big\{\big(x, \phi(x), \dots, \phi^{(k)}\big) \in I \times \mathbb{K}^{(k+1)n} : x \in I\big\} \subseteq U.$$

Finally, from the last equation of (3.2a), one obtains

$$\underset{x\in I}{\forall} \quad F\big(x, \phi(x), \dots, \phi^{(k)}\big) \overset{(3.3),(3.5)}{=} G_k\big(x, \Phi(x), \Phi'(x)\big) = 0,$$

proving ϕ satisfies (3.1a). ■

Example 3.2. The second-order initial value problem

$$y'' = -y, \qquad \begin{aligned} y(0) &= 0, \\ y'(0) &= r, \quad r \in \mathbb{R} \text{ given}, \end{aligned} \tag{3.8}$$

is equivalent to the following system of two first-order ODE:

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= -y_1, \end{aligned} \qquad y(0) = \begin{pmatrix} 0 \\ r \end{pmatrix}. \tag{3.9}$$

The solution to (3.9) is

$$\Phi : \mathbb{R} \longrightarrow \mathbb{R}^2, \quad \Phi(x) = \begin{pmatrix} \Phi_1(x) \\ \Phi_2(x) \end{pmatrix} = \begin{pmatrix} r \sin x \\ r \cos x \end{pmatrix}, \tag{3.10}$$

and, thus, the solution to (3.8) is

$$\phi : \mathbb{R} \longrightarrow \mathbb{R}, \quad \phi(x) = r \sin x. \tag{3.11}$$

—

As a consequence of Th. 3.1, one can carry out much of the general theory of ODE (such as results regarding existence and uniqueness of solutions) for systems of first-order ODE, obtaining the corresponding results for higher-order ODE as a corollary. This is the strategy usually pursued in the literature and we will follow suit in this class.

3.2 Existence of Solutions

It is a rather remarkable fact that, under the very mild assumption that $f : G \longrightarrow \mathbb{K}^n$ is a continuous function defined on an open subset G of $\mathbb{R} \times \mathbb{K}^{kn}$ with $(x_0, y_{0,0}, \dots, y_{0,k-1}) \in$

G, *every* initial value problem (1.7) for the n-dimensional explicit kth-order ODE (1.6) has at least one solution $\phi : I \longrightarrow \mathbb{K}^n$, defined on a, possibly very small, open interval. This is the contents of the Peano Th. 3.8 below and its Cor. 3.10. From Example 1.4(b), we already know that uniqueness of the solution cannot be expected without stronger hypotheses.

The proof of the Peano theorem requires some work. One of the key ingredients is the Arzelà-Ascoli Th. 3.7 that, under suitable hypotheses, guarantees a given sequence of continuous functions to have a uniformly convergent subsequence (the formulation in Th. 3.7 is suitable for our purposes – many different variants of the Arzelà-Ascoli theorem exist in the literature).

We begin with some prelimanaries from the theory of metric spaces. At this point, the reader might want to review the definition of a metric, a metric space, and basic notions on metric spaces, such as the notion of compactness and the notion of continuity of functions between metric spaces. Also recall that every normed space is a metric space via the metric induced by the norm (in particular, if we use metric notions on normed spaces, they are always meant with respect to the respective induced metric). If you are not sufficiently familiar with metrics and norms, you might want to consult the relevant subsections of [Phi13b, Sec. 1]; for compactness and some related results see, e.g., Appendix C.2.

Notation 3.3. Let (X, d) be a metric space. Given $x \in X$ and $r \in \mathbb{R}^+$, let

$$B_r(x) := \{y \in X : d(x, y) < r\}$$

denote the *open ball* with center x and radius r, also known as the r-ball with center x.

Definition 3.4. Let (X, d_X) and (Y, d_Y) be metric spaces. We say a sequence of functions $(f_m)_{m\in\mathbb{N}}$, $f_m : X \longrightarrow Y$, *converges uniformly* to a function $f : X \longrightarrow Y$ if, and only if,

$$\underset{\epsilon>0}{\forall} \; \underset{N\in\mathbb{N}}{\exists} \; \underset{\substack{m\geq N,\\ x\in X}}{\forall} \quad d_Y\big(f_m(x), f(x)\big) < \epsilon.$$

Theorem 3.5. *Let (X, d_X) and (Y, d_Y) be metric spaces. If the sequence $(f_m)_{m\in\mathbb{N}}$ of continuous functions $f_m : X \longrightarrow Y$ converges uniformly to the function $f : X \longrightarrow Y$, then f is continuous as well.*

Proof. We have to show that f is continuous at every $\xi \in X$. Thus, let $\xi \in X$ and $\epsilon > 0$. Due to the uniform convergence, we can choose $m \in \mathbb{N}$ such that $d_Y\big(f_m(x), f(x)\big) < \epsilon/3$ for every $x \in X$. Moreover, as f_m is continuous at ξ, there exists $\delta > 0$ such that $x \in B_\delta(\xi)$ implies $d_Y\big(f_m(\xi), f_m(x)\big) < \epsilon/3$. Thus, if $x \in B_\delta(\xi)$, then

$$\begin{aligned} d_Y\big(f(\xi), f(x)\big) &\leq d_Y\big(f(\xi), f_m(\xi)\big) + d_Y\big(f_m(\xi), f_m(x)\big) + d_Y\big(f_m(x), f(x)\big) \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon, \end{aligned}$$

proving f is continuous at ξ. ∎

Definition 3.6. Let (X, d_X) and (Y, d_Y) be metric spaces and let $\mathcal{F}$ be a set of functions from X into Y. Then the set $\mathcal{F}$ (or the functions in $\mathcal{F}$) are said to be *uniformly equicontinuous* if, and only if, for each $\epsilon > 0$, there is $\delta > 0$ such that

$$\underset{x,\xi \in X}{\forall} \left(d_X(x,\xi) < \delta \quad \Rightarrow \quad \underset{f \in \mathcal{F}}{\forall} \; d_Y\big(f(x), f(\xi)\big) < \epsilon \right). \tag{3.12}$$

Theorem 3.7 (Arzelà-Ascoli). *Let $n \in \mathbb{N}$, let $\|\cdot\|$ denote some norm on $\mathbb{K}^n$, and let $I \subseteq \mathbb{R}$ be some bounded interval. If $(f_m)_{m\in\mathbb{N}}$ is a sequence of functions $f_m : I \longrightarrow \mathbb{K}^n$ such that $\{f_m : m \in \mathbb{N}\}$ is uniformly equicontinuous and such that, for each $x \in I$, the sequence $\big(f_m(x)\big)_{m\in\mathbb{N}}$ is bounded, then $(f_m)_{m\in\mathbb{N}}$ has a uniformly convergent subsequence $(f_{m_j})_{j\in\mathbb{N}}$, i.e. there exists $f : I \longrightarrow \mathbb{K}^n$ such that*

$$\underset{\epsilon>0}{\forall} \; \underset{N\in\mathbb{N}}{\exists} \; \underset{\substack{j \geq N,\\ x\in I}}{\forall} \; \|f_{m_j}(x) - f(x)\| < \epsilon.$$

In particular, the limit function f is continuous.

Proof. Let $(r_1, r_2, \dots)$ be an enumeration of the set of rational numbers in I, i.e. of $\mathbb{Q} \cap I$. Inductively, we construct a sequence $(\mathcal{F}_m)_{m\in\mathbb{N}}$ of subsequences of $(f_m)_{m\in\mathbb{N}}$, $\mathcal{F}_m = (f_{m,k})_{k\in\mathbb{N}}$, such that

(i) for each $m \in \mathbb{N}$, $\mathcal{F}_m$ is a subsequence of each $\mathcal{F}_j$ with $j \in \{1, \dots, m\}$,

(ii) for each $m \in \mathbb{N}$, $\mathcal{F}_m$ converges pointwise at each of the first m rational numbers r_j; more precisely, there exists a sequence $(z_1, z_2, \dots)$ in $\mathbb{K}^n$ such that, for each $m \in \mathbb{N}$ and each $j \in \{1, \dots, m\}$:

$$\lim_{k\to\infty} f_{m,k}(r_j) = z_j.$$

Actually, we construct the $(z_m)_{m\in\mathbb{N}}$ inductively together with the $(\mathcal{F}_m)_{m\in\mathbb{N}}$: Since the sequence $(f_m(r_1))_{m\in\mathbb{N}}$ is, by hypothesis, a bounded sequence in $\mathbb{K}^n$, one can apply the Bolzano-Weierstrass theorem (cf. [Phi13b, Th. 1.16(b)]) to obtain $z_1 \in \mathbb{K}^n$ and a subsequence $\mathcal{F}_1 = (f_{1,k})_{k\in\mathbb{N}}$ of $(f_m)_{m\in\mathbb{N}}$ such that $\lim_{k\to\infty} f_{1,k}(r_1) = z_1$. To proceed by induction, we now assume to have already constructed $\mathcal{F}_1, \dots, \mathcal{F}_M$ and $z_1, \dots, z_M$ for $M \in \mathbb{N}$ such that (i) and (ii) hold for each $m \in \{1, \dots, M\}$. Since the sequence $(f_{M,k}(r_{M+1}))_{k\in\mathbb{N}}$ is a bounded sequence in $\mathbb{K}^n$, one can, once more, apply the Bolzano-Weierstrass theorem to obtain $z_{M+1} \in \mathbb{K}^n$ and a subsequence $\mathcal{F}_{M+1} = (f_{M+1,k})_{k\in\mathbb{N}}$ of $\mathcal{F}_M$ such that $\lim_{k\to\infty} f_{M+1,k}(r_{M+1}) = z_{M+1}$. Since $\mathcal{F}_{M+1}$ is a subsequence of $\mathcal{F}_M$, it is also a subsequence of all previous subsequences, i.e. (i) now also holds for $m = M+1$. In consequence, $\lim_{k\to\infty} f_{M+1,k}(r_j) = z_j$ for each $j = 1, \dots, M+1$, such that (ii) now also holds for $m = M+1$ as required.

Next, one considers the *diagonal sequence* $(g_m)_{m\in\mathbb{N}}$, $g_m := f_{m,m}$, and observes that this sequence converges pointwise at each rational number r_j ($\lim_{m\to\infty} g_m(r_j) = z_j$), since, at least for $m \geq j$, $(g_m)_{m\in\mathbb{N}}$ is a subsequence of every $\mathcal{F}_j$ (exercise) – in particular, $(g_m)_{m\in\mathbb{N}}$ is also a subsequence of the original sequence $(f_m)_{m\in\mathbb{N}}$.

In the last step of the proof, we show that $(g_m)_{m\in\mathbb{N}}$ converges uniformly on the entire interval I to some $f : I \longrightarrow \mathbb{K}^n$. To this end, fix $\epsilon > 0$. Since $\{g_m : m \in \mathbb{N}\} \subseteq \{f_m : m \in \mathbb{N}\}$, the assumed uniform equicontinuity of $\{f_m : m \in \mathbb{N}\}$ yields $\delta > 0$ such that

$$\forall_{x,\xi\in I} \quad \left(|x-\xi| < \delta \quad \Rightarrow \quad \forall_{m\in\mathbb{N}} \quad \big\|g_m(x) - g_m(\xi)\big\| < \frac{\epsilon}{3} \right).$$

Since I is bounded, it has finite length and, thus, it can be covered with finitely many intervals $I_1, \dots, I_N$, $I = \bigcup_{j=1}^N I_j$, $N \in \mathbb{N}$, such that each I_j has length less than δ. Moreover, since $\mathbb{Q}$ is dense in $\mathbb{R}$, for each $j \in \{1, \dots, N\}$, there exists $k(j) \in \mathbb{N}$ such that $r_{k(j)} \in I_j$. Define $M := \max\{k(j) : j = 1, \dots, N\}$. We note that each of the finitely many sequences $\big((g_m(r_1)\big)_{m\in\mathbb{N}}, \dots, \big((g_m(r_M)\big)_{m\in\mathbb{N}}$ is a Cauchy sequence. Thus,

$$\exists_{K\in\mathbb{N}} \quad \forall_{k,l\geq K} \quad \forall_{\alpha=1,\dots,M} \quad \big\|g_k(r_\alpha) - g_l(r_\alpha)\big\| < \frac{\epsilon}{3}. \tag{3.13}$$

We now consider an arbitrary $x \in I$ and $k, l \geq K$. Let $j \in \{1, \dots, N\}$ such that $x \in I_j$. Then $r_{k(j)} \in I_j$, $|r_{k(j)} - x| < \delta$, and the estimate in (3.13) holds for $\alpha = k(j)$. In consequence, we obtain the crucial estimate

$$\forall_{k,l\geq K} \quad \begin{aligned} &\big\|g_k(x) - g_l(x)\big\| \\ &\leq \big\|g_k(x) - g_k(r_{k(j)})\big\| + \big\|g_k(r_{k(j)}) - g_l(r_{k(j)})\big\| + \big\|g_l(r_{k(j)}) - g_l(x)\big\| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned} \tag{3.14}$$

The estimate (3.14) shows $\big((g_m(x)\big)_{m\in\mathbb{N}}$ is a Cauchy sequence for each $x \in I$, and we can define

$$f : I \longrightarrow \mathbb{K}^n, \quad f(x) := \lim_{m\to\infty} g_m(x). \tag{3.15}$$

Since K in (3.14) does not depend on $x \in I$, passing to the limit $k \to \infty$ in the estimate of (3.14) implies

$$\forall_{\substack{l\geq K,\\ x\in I}} \quad \|g_l(x) - f(x)\| \leq \epsilon,$$

proving uniform convergence of the subsequence $(g_m)_{m\in\mathbb{N}}$ of $(f_m)_{m\in\mathbb{N}}$ as desired. The continuity of f is now a consequence of Th. 3.5. ∎

At this point, we have all preparations in place to state and prove the existence theorem.

Theorem 3.8 (Peano)**.** *If $G \subseteq \mathbb{R}\times\mathbb{K}^n$ is open, $n \in \mathbb{N}$, and $f : G \longrightarrow \mathbb{K}^n$ is continuous, then, for each $(x_0, y_0) \in G$, the explicit n-dimensional first-order initial value problem*

$$y' = f(x, y), \tag{3.16a}$$

$$y(x_0) = y_0, \tag{3.16b}$$

has at least one solution. More precisely, given an arbitrary norm $\|\cdot\|$ on $\mathbb{K}^n$, (3.16) *has a solution $\phi : I \longrightarrow \mathbb{K}^n$, defined on the open interval*

$$I :=]x_0 - \alpha, x_0 + \alpha[, \tag{3.17}$$

$\alpha = \alpha(b) > 0$, where $b > 0$ is such that

$$B := \big\{(x,y) \in \mathbb{R} \times \mathbb{K}^n : |x - x_0| \le b \text{ and } \|y - y_0\| \le b\big\} \subseteq G, \tag{3.18}$$

$$M := M(b) := \max\{\|f(x,y)\| : (x,y) \in B\} < \infty, \tag{3.19}$$

and

$$\alpha := \alpha(b) := \begin{cases} \min\{b, b/M\} & \text{for } M > 0, \\ b & \text{for } M = 0. \end{cases} \tag{3.20}$$

In general, the choice of the norm $\|\cdot\|$ on $\mathbb{K}^n$ will influence the possible sizes of α and, thus, of I.

Proof. The proof will be conducted in several steps. In the first step, we check $\alpha = \alpha(b) > 0$ is well-defined: Since G is open, there always exists $b > 0$ such that (3.18) holds. Since B is a closed and bounded subset of the finite-dimensional space $\mathbb{R} \times \mathbb{K}^n$, B is compact (cf. [Phi13b, Cor. 3.5]). Since f and, thus, $\|f\|$ is continuous (every norm is even Lipschitz continuous due to the inverse triangle inequality), it must assume its maximum on the compact set B (cf. [Phi13b, Th. 3.8]), showing $M \in \mathbb{R}_0^+$ is well-defined by (3.19) and α is well-defined by (3.20).

In the second step of the proof, we note that it suffices to prove (3.16) has a solution ϕ_+, defined on $[x_0, x_0 + \alpha[$: One can then apply the time reversion Lem. 1.9(b): The proof providing the solution ϕ_+ also provides a solution $\psi_+ : [-x_0, -x_0 + \alpha[\longrightarrow \mathbb{K}^n$ to the time-reversed initial value problem, consisting of $y' = -f(-x, y)$ and $y(-x_0) = y_0$ (note that the same M and α work for the time-reversed problem). Then, according to Lem. 1.9(b), $\phi_- :]x_0 - \alpha, x_0] \longrightarrow \mathbb{K}^n$, $\phi_-(x) := \psi_+(-x)$, is a solution to (3.16). According to Lem. 1.7, we can patch ϕ_- and ϕ_+ together to obtain the desired solution

$$\phi : I \longrightarrow \mathbb{K}^n, \quad \phi(x) := \begin{cases} \phi_-(x) & \text{for } x \le x_0, \\ \phi_+(x) & \text{for } x \ge x_0, \end{cases} \tag{3.21}$$

defined on all of I. It is noted that one can also conduct the proof with the second step omitted, but then one has to perform the following steps on all of I, which means one has to consider additional cases in some places.

In the third step of the proof, we will define a sequence $(\phi_m)_{m\in\mathbb{N}}$ of functions

$$\phi_m : I_+ \longrightarrow \mathbb{K}^n, \quad I_+ := [x_0, x_0 + \alpha], \tag{3.22}$$

that constitute *approximate solutions* to (3.16). To begin the construction of ϕ_m, fix $m \in \mathbb{N}$. Since B is compact and f is continuous, we know f is even uniformly continuous on B (cf. [Phi13b, Th. 3.9]). In particular,

$$\underset{\delta_m > 0}{\exists} \quad \underset{(x,y),(\tilde{x},\tilde{y}) \in B}{\forall} \quad \Big(|x - \tilde{x}| < \delta_m,\ \|y - \tilde{y}\| < \delta_m \quad \Rightarrow \quad \big\|f(x,y) - f(\tilde{x},\tilde{y})\big\| < \frac{1}{m}\Big). \tag{3.23}$$

We now form what is called a *discretization* of the interval I_+, i.e. a partition of I_+ into sufficiently many small intervals: Let $N \in \mathbb{N}$ and

$$x_0 < x_1 < \cdots < x_{N-1} < x_N := x_0 + \alpha \tag{3.24}$$

such that

$$\underset{j\in\{1,\dots,N\}}{\forall} \quad x_j - x_{j-1} < \beta := \begin{cases} \min\{\delta_m,\, \delta_m/M,\, 1/m\} & \text{for } M > 0, \\ \min\{\delta_m,\, 1/m\} & \text{for } M = 0 \end{cases} \tag{3.25}$$

(for example one could make the equidistant choice $x_j := x_0 + jh$ with $h = \alpha/N$ and $N > \alpha/\beta$, but it does not matter how the x_j are defined as long as (3.24) and (3.25) both hold). Note that we get a different discretization of I_+ for each $m \in \mathbb{N}$; however, the dependence on m is suppressed in the notation for the sake of readability. We now define recursively

$$\begin{aligned} \phi_m : I_+ &\longrightarrow \mathbb{K}^n, \\ \phi_m(x_0) &:= y_0, \\ \phi_m(x) &:= \phi_m(x_j) + (x - x_j)\, f\big(x_j, \phi_m(x_j)\big) \quad \text{for each } x \in [x_j, x_{j+1}]. \end{aligned} \tag{3.26}$$

Note that there is no conflict between the two definitions given for $x = x_j$ with $j \in \{1, \dots, N-1\}$. Each function ϕ_m defines a polygon in $\mathbb{K}^n$. This construction is known as *Euler's method* and it can be used to obtain numerical approximations to the solution of the initial value problem (while simple, this method is not very efficient, though). We still need to verify that the definition (3.26) does actually make sense: We need to check that f can, indeed, be applied to $(x_j, \phi_m(x_j))$, i.e. we have to check $(x_j, \phi_m(x_j)) \in G$. We can actually show the stronger statement

$$\underset{x\in I_+}{\forall} \quad (x, \phi_m(x)) \in B, \tag{3.27}$$

where B is as defined in (3.18). First, it is pointed out that (3.20) implies $\alpha \le b$, such that $x \in I_+$ implies $|x - x_0| \le \alpha \le b$ as required in (3.18). One can now prove (3.27) by showing by induction on $j \in \{0, \dots, N-1\}$:

$$\underset{x\in[x_j,x_{j+1}]}{\forall} \quad (x, \phi_m(x)) \in B. \tag{3.28}$$

To start the induction, note $\phi_m(x_0) = y_0$ and $(x_0, y_0) \in B$ by (3.18). Now let $j \in \{0, \dots, N-1\}$ and $x \in [x_j, x_{j+1}]$. We estimate

$$\begin{aligned} \|\phi_m(x) - y_0\| &\le \|\phi_m(x) - \phi_m(x_j)\| + \sum_{k=1}^{j} \|\phi_m(x_k) - \phi_m(x_{k-1})\| \\ &\overset{(3.26)}{=} (x - x_j)\, \big\|f\big(x_j, \phi_m(x_j)\big)\big\| + \sum_{k=1}^{j} (x_k - x_{k-1})\, \big\|f\big(x_{k-1}, \phi_m(x_{k-1})\big)\big\| \\ &\overset{(*)}{\le} (x - x_j)\, M + \sum_{k=1}^{j} (x_k - x_{k-1})\, M = (x - x_0)\, M \\ &\le \alpha\, M \overset{(3.20)}{\le} b, \end{aligned} \tag{3.29}$$

where, at $(*)$, it was used that $(x_k, \phi_m(x_k)) \in B$ by induction hypothesis for each $k = 0, \dots, j$, and, thus, $\big\|f\big(x_k, \phi_m(x_k)\big)\big\| \leq M$ by (3.19). Estimate (3.29) completes the induction and the third step of the proof.

In the fourth step of the proof, we establish several properties of the functions ϕ_m. The first two properties are immediate from (3.26), namely that ϕ_m is continuous on I^+ and differentiable at each $x \in]x_j, x_{j+1}[$, $j \in \{0, \dots, N-1\}$, where

$$\underset{j\in\{0,\dots,N-1\}}{\forall} \quad \underset{x\in]x_j,x_{j+1}[}{\forall} \quad \phi_m'(x) = f\big(x_j, \phi_m(x_j)\big). \tag{3.30}$$

The next property to establish is

$$\underset{s,t\in I_+}{\forall} \quad \|\phi_m(t) - \phi_m(s)\| \leq |t-s|\, M. \tag{3.31}$$

To prove (3.31), we may assume $s < t$ without loss of generality. If $s, t \in [x_j, x_{j+1}]$, $j \in \{0, \dots, N-1\}$, then

$$\begin{aligned}
&\|\phi_m(t) - \phi_m(s)\| \\
&\overset{(3.26)}{=} \quad \big\|\phi_m(x_j) + (t - x_j)\, f\big(x_j, \phi_m(x_j)\big) - \phi_m(x_j) - (s - x_j)\, f\big(x_j, \phi_m(x_j)\big)\big\| \\
&= \quad |t-s|\, \big\|f\big(x_j, \phi_m(x_j)\big)\big\| \overset{(3.19)}{\leq} |t-s|\, M
\end{aligned} \tag{3.32a}$$

as desired. If s, t are not contained in the same interval $[x_j, x_{j+1}]$, then fix $j < k$ such that $s \in [x_j, x_{j+1}]$ and $t \in [x_k, x_{k+1}]$. Then (3.31) follows from an estimate analogous to the one in (3.29):

$$\begin{aligned}
&\|\phi_m(t) - \phi_m(s)\| \\
&\leq \quad \|\phi_m(s) - \phi_m(x_{j+1})\| + \sum_{l=j+1}^{k-1} \|\phi_m(x_l) - \phi_m(x_{l+1})\| + \|\phi_m(x_k) - \phi_m(t)\| \\
&\overset{(3.32a)}{\leq} \quad |s - x_{j+1}|\, M + \sum_{l=j+1}^{k-1} |x_l - x_{l+1}|\, M + |t - x_k|\, M \\
&= \quad |t-s|\, M,
\end{aligned} \tag{3.32b}$$

completing the proof of (3.31). The following property of the ϕ_m is the justification for calling them *approximate solutions* to our initial value problem (3.16):

$$\underset{j\in\{0,\dots,N-1\}}{\forall} \quad \underset{x\in]x_j,x_{j+1}[}{\forall} \quad \big\|\phi_m'(x) - f\big(x, \phi_m(x)\big)\big\| < \frac{1}{m}. \tag{3.33}$$

Indeed, (3.33) is a consequence of (3.23), i.e. of the uniform continuity of f on B: First, if $M = 0$, then $f \equiv \phi_m' \equiv 0$ and there is nothing to prove. So let $M > 0$. If $x \in]x_j, x_{j+1}[$, then, according to (3.30), we have $\phi_m'(x) = f\big(x_j, \phi_m(x_j)\big)$. Thus, by (3.25),

$$|x - x_j| < \beta \leq \min\{\delta_m, \delta_m/M\} \quad \Rightarrow \quad \|\phi_m(x) - \phi_m(x_j)\| \overset{(3.31)}{\leq} |x - x_j|\, M < \delta_m, \tag{3.34a}$$

and

$$\left\|\phi_m'(x) - f\big(x, \phi_m(x)\big)\right\| = \left\|f\big(x_j, \phi_m(x_j)\big) - f\big(x, \phi_m(x)\big)\right\| \overset{(3.34a),(3.23)}{<} \frac{1}{m}, \tag{3.34b}$$

proving (3.33).

The last property of the ϕ_m we need is

$$\underset{x \in I_+}{\forall} \quad \begin{aligned} \|\phi_m(x)\| &\le \|\phi_m(x) - \phi_m(x_0)\| + \|\phi_m(x_0)\| \overset{(3.31)}{\le} |x - x_0|\, M + \|\phi_m(x_0)\| \\ &\le \alpha M + \|y_0\|, \end{aligned} \tag{3.35}$$

which says that the ϕ_m are pointwise and even uniformly bounded.

In the fifth and last step of the proof, we use the Arzelà-Ascoli Th. 3.7 to obtain a function $\phi_+ : I_+ \longrightarrow \mathbb{K}^n$, and we show that ϕ constitutes a solution to (3.16). According to (3.31), the ϕ_m are uniformly equicontinuous (given $\epsilon > 0$, condition (3.12) is satisfied with $\delta := \epsilon/M$ for $M > 0$ and arbitrary $\delta > 0$ for $M = 0$), and according to (3.35) the ϕ_m are bounded such that the Arzelà-Ascoli Th. 3.7 applies to yield a subsequence $(\phi_{m_j})_{j\in\mathbb{N}}$ of $(\phi_m)_{m\in\mathbb{N}}$ converging uniformly to some continuous function $\phi_+ : I_+ \longrightarrow \mathbb{K}^n$. So it merely remains to verify that ϕ_+ is a solution to (3.16).

As the uniform convergence of the $(\phi_{m_j})_{j\in\mathbb{N}}$ implies pointwise convergence, we have $\phi_+(x_0) = \lim_{j\to\infty} \phi_{m_j}(x_0) = y_0$, showing ϕ_+ satisfies the initial condition (3.16b).

Next,

$$\underset{x \in I}{\forall} \quad (x, \phi_+(x)) = \lim_{j\to\infty} (x, \phi_{m_j}(x)) \in B,$$

since each $(x, \phi_{m_j}(x))$ is in B and B is closed. In particular, $f(x, \phi_+(x))$ is well-defined for each $x \in I_+$.

To prove that ϕ_+ also satisfies the ODE (3.16a), by Th. 1.5, it suffices to show

$$\underset{x \in I_+}{\forall} \quad \phi_+(x) - \phi_+(x_0) - \int_{x_0}^{x} f\big(t, \phi_+(t)\big)\,\mathrm{d}t = 0. \tag{3.36}$$

Fixing $x \in I_+$ and using the triangle inequality for the umpteenth time, one obtains

$$\begin{aligned} &\left\|\phi_+(x) - \phi_+(x_0) - \int_{x_0}^{x} f\big(t, \phi_+(t)\big)\,\mathrm{d}t\right\| \\ &\le \|\phi_+(x) - \phi_{m_j}(x)\| + \left\|\phi_{m_j}(x) - \phi_+(x_0) - \int_{x_0}^{x} f\big(t, \phi_{m_j}(t)\big)\,\mathrm{d}t\right\| \\ &\quad + \left\|\int_{x_0}^{x} \Big(f\big(t, \phi_{m_j}(t)\big) - f\big(t, \phi_+(t)\big)\Big)\,\mathrm{d}t\right\|, \end{aligned} \tag{3.37}$$

holding for every $j \in \mathbb{N}$. We will conclude the proof by showing that all three summands on the right-hand side of (3.37) tend to 0 for $j \to \infty$. As already mentioned above, the uniform convergence of the $(\phi_{m_j})_{j\in\mathbb{N}}$ implies pointwise convergence, implying the convergence of the first summand. We tackle the third summand next, using

$$\left\|\int_{x_0}^{x} \Big(f\big(t, \phi_{m_j}(t)\big) - f\big(t, \phi_+(t)\big)\Big)\,\mathrm{d}t\right\| \le \int_{x_0}^{x} \left\|f\big(t, \phi_{m_j}(t)\big) - f\big(t, \phi_+(t)\big)\right\|\mathrm{d}t\,, \tag{3.38}$$

which holds for every norm (cf. Appendix B), but can easily be checked directly for the 1-norm, where $\|(z_1,\dots,z_n)\|_1 := \sum_{j=1}^n |z_j|$ (exercise). Given $\epsilon > 0$, the uniform continuity of f on B provides $\delta > 0$ such that $\big\|f\big(t,\phi_{m_j}(t)\big) - f\big(t,\phi_+(t)\big)\big\| < \epsilon/\alpha$ for $\|\phi_{m_j}(t) - \phi_+(t)\| < \delta$. The uniform convergence of $(\phi_{m_j})_{j\in\mathbb{N}}$ then yields $K \in \mathbb{N}$ such that $\|\phi_{m_j}(t) - \phi_+(t)\| < \delta$ for every $j \geq K$ and each $t \in I$. Thus,

$$\underset{j\geq K}{\forall} \quad \int_{x_0}^{x} \Big\|f\big(t,\phi_{m_j}(t)\big) - f\big(t,\phi_+(t)\big)\Big\| \,\mathrm{d}t \leq \frac{|x-x_0|\,\epsilon}{\alpha} \leq \epsilon,$$

thereby establishing the convergence of the third summand from the right-hand side of (3.37). For the remaining second summand, we note that the fact that each ϕ_m is continuous and piecewise differentiable (with piecewise constant derivative) allows to apply the fundamental theorem of calculus in the form [Phi13a, Th. G.6(b)] to obtain

$$\underset{x\in I_+}{\forall} \quad \phi_m(x) = \phi_m(x_0) + \int_{x_0}^{x} \phi_m'(t)\,\mathrm{d}t\,. \tag{3.39}$$

Using (3.39) in the second summand of the right-hand side of (3.37) provides

$$\begin{aligned} \left\|\phi_{m_j}(x) - \phi_+(x_0) - \int_{x_0}^{x} f\big(t,\phi_{m_j}(t)\big)\,\mathrm{d}t\right\| \quad &\leq \quad \int_{x_0}^{x} \Big\|\phi_{m_j}'(t) - f\big(t,\phi_{m_j}(t)\big)\Big\|\,\mathrm{d}t \\ &\overset{(3.33)}{\leq} \quad \int_{x_0}^{x} \frac{1}{m_j} \leq \frac{\alpha}{m_j}, \end{aligned}$$

showing the convergence of the second summand, which finally concludes the proof. ■

Corollary 3.9. *If $G \subseteq \mathbb{R}\times\mathbb{K}^n$ is open, $n \in \mathbb{N}$, $f : G \longrightarrow \mathbb{K}^n$ is continuous, and $C \subseteq G$ is compact, then there exists $\alpha > 0$, such that, for each $(x_0, y_0) \in C$, the explicit n-dimensional first-order initial value problem* (3.16) *has a solution $\phi : I \longrightarrow \mathbb{K}^n$, defined on the open interval $I :=]x_0-\alpha, x_0+\alpha[$, i.e. always on an interval of the* same *length 2α.*

Proof. Exercise. ■

Corollary 3.10. *If $G \subseteq \mathbb{R}\times\mathbb{K}^{kn}$ is open, $k, n \in \mathbb{N}$, and $f : G \longrightarrow \mathbb{K}^n$ is continuous, then, for each $(x_0, y_{0,0}, \dots, y_{0,k-1}) \in G$, the explicit n-dimensional kth-order initial value problem consisting of* (1.6) *and* (1.7)*, which, for convenience, we rewrite*

$$y^{(k)} = f\big(x, y, y', \dots, y^{(k-1)}\big), \tag{3.40a}$$

$$\underset{j\in\{0,\dots,k-1\}}{\forall} \quad y^{(j)}(x_0) = y_{0,j}, \tag{3.40b}$$

has at least one solution. More precisely, there exists an open interval $I \subseteq \mathbb{R}$ with $x_0 \in I$ and $\phi : I \longrightarrow \mathbb{K}^n$ such that ϕ is a solution to (3.40)*. If $C \subseteq G$ is compact, then there exists $\alpha > 0$ such that, for each $(x_0, y_{0,0}, \dots, y_{0,k-1}) \in C$,* (3.40) *has a solution $\phi : I \longrightarrow \mathbb{K}^n$, defined on the open interval $I :=]x_0-\alpha, x_0+\alpha[$, i.e. always on an interval of the* same *length 2α.*

Proof. If f is continuous, then the right-hand side of the equivalent first-order system (3.2a) (written in explicit form) is given by the continuous function

$$\tilde{f} : G \longrightarrow \mathbb{K}^{kn}, \quad \tilde{f}(x, y_1, \dots, y_k) := \begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ y_{k-1} \\ f(x, y_1, \dots, y_k) \end{pmatrix}. \tag{3.41}$$

Thus, Th. 3.8 provides a solution $\Phi : I \longrightarrow \mathbb{K}^{kn}$ to (3.2) and, then, Th. 3.1(b) yields $\phi := \Phi_1$ to be a solution to (3.40). Moreover, if $C \subseteq G$ is compact, then Cor. 3.9 provides $\alpha > 0$ such that, for each $(x_0, y_{0,0}, \dots, y_{0,k-1}) \in C$, (3.2) has a solution $\Phi : I \longrightarrow \mathbb{K}^{kn}$, defined on the same open interval $I :=]x_0 - \alpha, x_0 + \alpha[$. In particular, $\phi := \Phi_1$, the corresponding solution to (3.40) is also defined on the same I. ■

While the Peano theorem is striking in its generality, it does have several drawbacks: (a) the interval, where the existence of a solution is proved can be unnecessarily short; (b) the selection of the subsequence using the Arzelà-Ascoli theorem makes the proof nonconstructive; (c) uniqueness of solutions is not provided, even in cases, where unique solutions exist; (d) it does not provide information regarding how the solution changes with a change of the initial condition. We will subsequently address all these points, namely (b) and (c) in Sec. 3.3 (we will see that the proof of the Peano theorem becomes constructive in situations, where the solution is unique – in general, a constructive proof is not available), (a) in Sec. 3.4, and (d) in Sec. 3.5.

3.3 Uniqueness of Solutions

Example 1.4(b) shows that the hypotheses of the Peano Th. 3.8 are not strong enough to guarantee the initial value problem (3.16) has a unique solution, not even in some neighborhood of x_0. The additional condition that will yield uniqueness is local Lipschitz continuity of f with respect to y.

Definition 3.11. Let $m, n \in \mathbb{N}$, $G \subseteq \mathbb{R} \times \mathbb{K}^m$, and $f : G \longrightarrow \mathbb{K}^n$.

(a) The function f is called *(globally) Lipschitz continuous* or just *(globally) Lipschitz* with respect to y if, and only if,

$$\underset{L \geq 0}{\exists} \quad \underset{(x,y),(x,\bar{y}) \in G}{\forall} \quad \big\| f(x,y) - f(x,\bar{y}) \big\| \leq L \|y - \bar{y}\|. \tag{3.42}$$

(b) The function f is called *locally Lipschitz continuous* or just *locally Lipschitz* with respect to y if, and only if, for each $(x_0, y_0) \in G$, there exists a (relative) open set $U \subseteq G$ such that $(x_0, y_0) \in U$ (i.e. U is a (relative) open neighborhood of (x_0, y_0)) and f is Lipschitz continuous with respect to y on U, i.e. if, and only if,

$$\underset{(x_0,y_0) \in G}{\forall} \quad \underset{(x_0,\, y_0) \,\in\, U \,\subseteq\, G \text{ open}}{\exists} \quad \underset{L \geq 0}{\exists} \quad \underset{(x,y),(x,\bar{y}) \in U}{\forall} \quad \big\| f(x,y) - f(x,\bar{y}) \big\| \leq L \|y - \bar{y}\|. \tag{3.43}$$

The number L occurring in (a),(b) is called *Lipschitz constant.* The norms on $\mathbb{K}^m$ and $\mathbb{K}^n$ in (a),(b) are arbitrary. If one changes the norms, then one will, in general, change L, but not the property of f being (locally) Lipschitz.

Caveat 3.12. It is emphasized that $f : G \longrightarrow \mathbb{K}^n$, $(x, y) \mapsto f(x, y)$, being Lipschitz with respect to y does *not* imply f to be continuous: Indeed, if $I \subseteq \mathbb{R}$, $\emptyset \neq A \subseteq \mathbb{K}^m$, and $g : I \longrightarrow \mathbb{K}^n$ is an arbitrary *discontinuous* function, then $f : I \times A \longrightarrow \mathbb{K}^n$, $f(x, y) := g(x)$ is not continuous, but satisfies (3.42) with $L = 0$.

—

While the local neighborhoods U, where a function locally Lipschitz (with respect to y) is actually Lipschitz continuous (with respect to y) can be very small, we will now show that a continuous function is locally Lipschitz (with respect to y) on G if, and only if, it is Lipschitz continuous (with respect to y) on *every compact* set $K \subseteq G$.

Proposition 3.13. *Let $m, n \in \mathbb{N}$, $G \subseteq \mathbb{R} \times \mathbb{K}^m$, and $f : G \longrightarrow \mathbb{K}^n$ be continuous. Then f is locally Lipschitz with respect to y if, and only if, f is (globally) Lipschitz with respect to y on every compact subset K of G.*

Proof. First, assume f is not locally Lipschitz with respect to y. Then there exists $(x_0, y_0) \in G$ such that

$$\underset{N\in\mathbb{N}}{\forall} \quad \underset{\substack{(x_N,y_{N,1}),(x_N,y_{N,2})\\ \in G\cap B_{1/N}(x_0,y_0)}}{\exists} \quad \big\|f(x_N, y_{N,1}) - f(x_N, y_{N,2})\big\| > N\|y_{N,1} - y_{N,2}\|. \tag{3.44}$$

The set

$$K := \big\{(x_0, y_0)\big\} \cup \big\{(x_N, y_{N,j}) : N \in \mathbb{N}, \, j \in \{1, 2\}\big\}$$

is clearly a compact subset of G (e.g. by the Heine-Borel property of compact sets (see Th. C.19), since every open set containing (x_0, y_0) must contain all, but finitely many, of the elements of K). Due to (3.44), f is not (globally) Lipschitz with respect to y on the compact set K (so, actually, continuity of f was not used for this direction).

Conversely, assume f to be locally Lipschitz with respect to y, and consider a compact subset K of G. Then, for each $(x, y) \in K$, there is some (relatively) open $U_{(x,y)} \subseteq G$ with $(x, y) \in U_{(x,y)}$ and such that f is Lipschitz with respect to y in $U_{(x,y)}$. By the Heine-Borel property of compact sets (see Th. C.19), there are finitely many $U_1 := U_{(x_1,y_1)}, \dots, U_N := U_{(x_N,y_N)}$, $N \in \mathbb{N}$, such that

$$K \subseteq \bigcup_{j=1}^{N} U_j. \tag{3.45}$$

For each $j = 1, \dots, N$, let L_j denote the Lipschitz constant for f on U_j and set $L' := \max\{L_1, \dots, L_N\}$. As f is assumed continuous and K is compact, we have

$$M := \max\{\|f(x, y)\| : (x, y) \in K\} < \infty. \tag{3.46}$$

Using the compactness of K once again, there exists a *Lebesgue number* $\delta > 0$ for the open cover $(U_j)_{j\in\{1,\dots,N\}}$ of K (cf. Th. C.21), i.e. $\delta > 0$ such that

$$\underset{(x,y),(x,\bar{y})\in K}{\forall} \Big(\|y-\bar{y}\| < \delta \quad \Rightarrow \quad \underset{j\in\{1,\dots,N\}}{\exists} \quad \{(x,y),(x,\bar{y})\} \subseteq U_j\Big). \tag{3.47}$$

Define $L := \max\{L', 2M/\delta\}$. Then, for every $(x,y),(x,\bar{y}) \in K$:

$$\|y-\bar{y}\| < \delta \quad \Rightarrow \quad \|f(x,y)-f(x,\bar{y})\| \le L_j\|y-\bar{y}\| \le L\|y-\bar{y}\|, \tag{3.48a}$$

$$\|y-\bar{y}\| \ge \delta \quad \Rightarrow \quad \|f(x,y)-f(x,\bar{y})\| \le 2M = \frac{2M\delta}{\delta} \le L\|y-\bar{y}\|, \tag{3.48b}$$

completing the proof that f is Lipschitz with respect to y on K. ■

While, in general, the assertion of Prop. 3.13 becomes false if the continuity of f is omitted, for *convex* G, it does hold without the continuity assumption on f (see Appendix D). The following Prop. 3.14 provides a useful sufficient condition for $f : G \longrightarrow \mathbb{K}^n$, $G \subseteq \mathbb{R}\times\mathbb{K}^m$ open, to be locally Lipschitz with respect to y:

Proposition 3.14. *Let $m,n \in \mathbb{N}$, let $G \subseteq \mathbb{R}\times\mathbb{K}^m$ be open, and $f : G \longrightarrow \mathbb{K}^n$. A sufficient condition for f to be locally Lipschitz with respect to y is f being continuously (real) differentiable with respect to y, i.e., f is locally Lipschitz with respect to y provided that all partials $\partial_{y_k} f_l$; $k,l = 1,\dots,n$ ($\partial_{y_{k,1}} f_l$, $\partial_{y_{k,2}} f_l$ for $\mathbb{K} = \mathbb{C}$) exist and are continuous.*

Proof. We consider the case $\mathbb{K} = \mathbb{R}$; the case $\mathbb{K} = \mathbb{C}$ is included by using the identifications $\mathbb{C}^m \cong \mathbb{R}^{2m}$ and $\mathbb{C}^n \cong \mathbb{R}^{2n}$. Given $(x_0,y_0) \in G$, we have to show f is Lipschitz with respect to y on some open set $U \subseteq G$ with $(x_0,y_0) \in U$. As in the Peano Th. 3.8, since G is open,

$$\underset{b>0}{\exists} \quad B := \big\{(x,y) \in \mathbb{R}\times\mathbb{R}^m : \, |x-x_0| \le b \text{ and } \|y-y_0\|_1 \le b\big\} \subseteq G,$$

where $\|\cdot\|_1$ denotes the 1-norm on $\mathbb{R}^m$. Since the $\partial_{y_k} f_l$, $(k,l) \in \{1,\dots,m\}\times\{1,\dots,n\}$, are all continuous on the compact set B,

$$M := \max\big\{|\partial_{y_k} f_l(x,y)| : \, (x,y) \in B, \; (k,l) \in \{1,\dots,m\}\times\{1,\dots,n\}\big\} < \infty. \tag{3.49}$$

Applying the mean value theorem (cf. [Phi13b, Th. 2.32]) to the n components of the function

$$f_x : \big\{y \in \mathbb{R}^m : \, (x,y) \in B\big\} \longrightarrow \mathbb{R}^n, \quad f_x(y) := f(x,y),$$

we obtain $\eta_1,\dots,\eta_n \in \mathbb{R}^m$ such that

$$f_l(x,y) - f_l(x,\bar{y}) = \sum_{k=1}^{m} \partial_{y_k} f_l(x,\eta_l)(y_k - \bar{y}_k), \tag{3.50}$$

and, thus,

$$\underset{(x,y),(x,\bar{y})\in B}{\forall} \quad \begin{aligned} \big\|f(x,y)-f(x,\bar{y})\big\|_1 &= \sum_{l=1}^{n} |f_l(x,y) - f_l(x,\bar{y})| \\ &\overset{(3.49),(3.50)}{\le} \sum_{l=1}^{n}\sum_{k=1}^{m} M|y_k-\bar{y}_k| = \sum_{l=1}^{n} M\|y-\bar{y}\|_1 = nM\|y-\bar{y}\|_1, \end{aligned} \tag{3.51}$$

i.e. f is Lipschitz with respect to y on B (where

$$\Big\{(x,y) \in \mathbb{R} \times \mathbb{R}^m : |x - x_0| < b \text{ and } \|y - y_0\|_1 < b\Big\} \subseteq B$$

is an open neighborhood of (x_0, y_0)), showing f is locally Lipschitz with respect to y. ■

Theorem 3.15. *If $G \subseteq \mathbb{R} \times \mathbb{K}^n$ is open, $n \in \mathbb{N}$, and $f : G \longrightarrow \mathbb{K}^n$ is continuous and locally Lipschitz with respect to y, then, for each $(x_0, y_0) \in G$, the explicit n-dimensional first-order initial value problem*

$$y' = f(x,y), \tag{3.52a}$$

$$y(x_0) = y_0, \tag{3.52b}$$

has a unique solution. More precisely, if $I \subseteq \mathbb{R}$ is an open interval and $\phi, \psi : I \longrightarrow \mathbb{K}^n$ are both solutions to (3.52a)*, then $\phi(x_0) = \psi(x_0)$ for one $x_0 \in I$ implies $\phi(x) = \psi(x)$ for all $x \in I$:*

$$\left(\underset{x_0 \in I}{\exists} \quad \phi(x_0) = \psi(x_0) \right) \quad \Rightarrow \quad \left(\underset{x \in I}{\forall} \quad \phi(x) = \psi(x) \right). \tag{3.53}$$

Proof. We first show that ϕ and ψ must agree in a small neighborhood of x_0:

$$\underset{\epsilon > 0}{\exists} \quad \underset{x \in]x_0 - \epsilon, x_0 + \epsilon[}{\forall} \quad \phi(x) = \psi(x). \tag{3.54}$$

Since f is continuous and both ϕ and ψ are solutions to the initial value problem (3.52), we can use Th. 1.5 to obtain

$$\underset{x \in I}{\forall} \quad \phi(x) - \psi(x) = \int_{x_0}^{x} \Big(f\big(t, \phi(t)\big) - f\big(t, \psi(t)\big) \Big) \, \mathrm{d}t \,. \tag{3.55}$$

As f is locally Lipschitz with respect to y, there exists $\delta > 0$ such that f is Lipschitz with Lipschitz constant $L \geq 0$ with respect to y on

$$U := \{(x,y) \in G : |x - x_0| < \delta, \ \|y - y_0\| < \delta\},$$

where we have chosen some arbitrary norm $\|\cdot\|$ on $\mathbb{K}^n$. The continuity of ϕ, ψ implies the existence of $\tilde{\epsilon} > 0$ such that $\overline{B}_{\tilde{\epsilon}}(x_0) \subseteq I$, $\phi(B_{\tilde{\epsilon}}(x_0)) \subseteq B_\delta(y_0)$ and $\psi(B_{\tilde{\epsilon}}(x_0)) \subseteq B_\delta(y_0)$, implying

$$\underset{x \in B_{\tilde{\epsilon}}(x_0)}{\forall} \quad \Big\| f\big(x, \phi(x)\big) - f\big(x, \psi(x)\big) \Big\| \leq L \, \big\| \phi(x) - \psi(x) \big\|. \tag{3.56}$$

Next, define

$$\epsilon := \min\{\tilde{\epsilon}, \ 1/(2L)\}$$

and, using the compactness of $\overline{B}_\epsilon(x_0) = [x_0 - \epsilon, x_0 + \epsilon]$ plus the continuity of ϕ, ψ,

$$M := \max\Big\{ \|\phi(x) - \psi(x)\| : \ x \in \overline{B}_\epsilon(x_0) \Big\} < \infty.$$

From (3.55) and (3.56), we obtain

$$\underset{x \in B_\epsilon(x_0)}{\forall} \quad \|\phi(x) - \psi(x)\| \leq L \left| \int_{x_0}^{x} \|\phi(t) - \psi(t)\| \, \mathrm{d}t \right| \leq L \, |x - x_0| \, M \leq \frac{M}{2} \tag{3.57}$$

(note that the integral in (3.57) can be negative for $x < x_0$). The definition of M together with (3.57) yields $M \leq M/2$, i.e. $M = 0$, finishing the proof of (3.54).

To prove $\phi(x) = \psi(x)$ for each $x \geq x_0$, let

$$s := \sup\{\xi \in I : \phi(x) = \psi(x) \text{ for each } x \in [x_0, \xi]\}.$$

One needs to show $s = \sup I$. If $s = \sup I$ does not hold, then there exists $\alpha > 0$ such that $[s, s+\alpha] \subseteq I$. Then the continuity of ϕ, ψ implies $\phi(s) = \psi(s)$, i.e. ϕ and ψ satisfy the same initial value problem at s such that (3.54) must hold with s instead of x_0, in contradiction to the definition of s. Finally, $\phi(x) = \psi(x)$ for each $x \leq x_0$ follows in an completely analogous fashion, which concludes the proof of the theorem. ■

Corollary 3.16. *If $G \subseteq \mathbb{R} \times \mathbb{K}^{kn}$ is open, $k, n \in \mathbb{N}$, and $f : G \longrightarrow \mathbb{K}^n$ is continuous and locally Lipschitz with respect to y, then, for each $(x_0, y_{0,0}, \dots, y_{0,k-1}) \in G$, the explicit n-dimensional kth-order initial value problem consisting of* (1.6) *and* (1.7)*, i.e.*

$$y^{(k)} = f\big(x, y, y', \dots, y^{(k-1)}\big),$$
$$\underset{j \in \{0,\dots,k-1\}}{\forall} \quad y^{(j)}(x_0) = y_{0,j},$$

has a unique solution. More precisely, if $I \subseteq \mathbb{R}$ is an open interval and $\phi, \psi : I \longrightarrow \mathbb{K}^n$ are both solutions to (1.6)*, then*

$$\underset{j \in \{0,\dots,k-1\}}{\forall} \quad \phi^{(j)}(x_0) = \psi^{(j)}(x_0) \tag{3.58}$$

holding for one $x_0 \in I$ implies $\phi(x) = \psi(x)$ for all $x \in I$.

Proof. Exercise. ■

Remark 3.17. According to Th. 3.15, the condition of f being continuous and locally Lipschitz with respect to y is *sufficient* for each initial value problem (3.52) to have a unique solution. However, this condition is not *necessary*: It is an exercise to show that the continuous function

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad f(x, y) := \begin{cases} 1 & \text{for } y \leq 0, \\ 1 + \sqrt{y} & \text{for } y \geq 0, \end{cases} \tag{3.59}$$

is not locally Lipschitz with respect to y, but that, for each $(x_0, y_0) \in \mathbb{R}^2$, the initial value problem (3.52) still has a unique solution in the sense that (3.53) holds for each solution ϕ to (3.52a). And one can (can you?) even find simple examples of f being defined on an open domain such that f is discontinuous at every point in its domain and every initial value problem (3.52) still has a unique solution.

—

At the end of Sec. 3.2, it was pointed out that the proof of the Peano Th. 3.8 is non-constructive due to the selection of a subsequence. The following Th. 3.18 shows that, whenever the initial value problem has a unique solution, it becomes unnecessary to select a subsequence, and the construction procedure (namely Euler's method) used in the proof of Th. 3.8 becomes an effective (if not necessarily efficient) numerical approximation procedure for the unique solution.

Theorem 3.18. *Consider the situation of the Peano Th.* 3.8*. Under the additional assumption that the solution to the explicit n-dimensional first-order initial value problem* (3.16) *is unique on some interval $J \subseteq [x_0, x_0+\alpha[$, $x_0 \in J$, and where $\alpha > 0$ is constructed as in Th.* 3.8 *(i.e. given by* (3.18) – (3.20)*), every sequence $(\phi_m)_{m\in\mathbb{N}}$ of functions defined on J according to Euler's method as in the proof of Th.* 3.8 *(i.e. defined as in* (3.26)*) converges uniformly to the unique solution $\phi : J \longrightarrow \mathbb{K}^n$. An analogous statement also holds for $J \subseteq]x_0-\alpha, x_0]$, $x_0 \in J$.*

Proof. Seeking a contradiction, assume $(\phi_m)_{m\in\mathbb{N}}$ does not converge uniformly to the unique solution ϕ. Then there exists $\epsilon > 0$ and a subsequence $(\phi_{m_j})_{j\in\mathbb{N}}$ such that

$$\underset{j\in\mathbb{N}}{\forall} \quad \|\phi_{m_j} - \phi\|_{\sup} = \sup\left\{\|\phi_{m_j}(x) - \phi(x)\| : x \in J\right\} \geq \epsilon. \tag{3.60}$$

However, as a subsequence, $(\phi_{m_j})_{j\in\mathbb{N}}$ still has all the properties of the $(\phi_m)_{m\in\mathbb{N}}$ (namely pointwise boundedness, uniform equicontinuity, piecewise differentiability, being approximate solutions according to (3.33)) that guanranteed the existence of a subsequence, converging to a solution. Thus, since the solution is unique on J, $(\phi_{m_j})_{j\in\mathbb{N}}$ must, in turn, have a subsequence, converging uniformly to ϕ, which is in contradiction to (3.60). This shows the assumption that $(\phi_m)_{m\in\mathbb{N}}$ does not converge uniformly to ϕ must have been false. The proof of the analogous statement for $J \subseteq]x_0-\alpha, x_0]$, $x_0 \in J$ one obtains, e.g., via time reversion (cf. the second step of the proof of Th. 3.8). ∎

Remark 3.19. The argument used to prove Th. 3.18 is of a rather general nature: It can be applied whenever a sequence is known to have a subsequence converging to some solution of some equation (or some other problem), provided the same still holds for every subsequence of the original sequence – in that case, the additional knowledge that the solution is unique implies the convergence of the original sequence without the need to select a subsequence.

3.4 Extension of Solutions, Maximal Solutions

The Peano Th. 3.8 and Cor. 3.10 show the existence of *local* solutions to explicit initial value problems, i.e. the solution's existence is proved on some, possibly small, interval containing the initial point x_0. In the current section, we will address the question in which circumstances such local solutions can be extended, we will prove the existence of maximal solutions (solutions that can not be extended), and we will learn how such maximal solutions can be identified.

Definition 3.20. Let $\phi : I \longrightarrow \mathbb{K}^n$, $n \in \mathbb{N}$, be a solution to some ODE (such as (1.6) or (1.4) in the most general case), defined on some open interval $I \subseteq \mathbb{R}$.

(a) We say ϕ has an *extension* or *continuation* to the right (resp. to the left) if, and only if, there exists a solution $\psi : J \longrightarrow \mathbb{K}^n$ to the same ODE, defined on some open interval $J \supseteq I$ such that $\psi{\restriction_I} = \phi$ and

$$\sup J > \sup I \quad (\text{resp.} \quad \inf J < \inf I). \tag{3.61}$$

An *extension* or *continuation* of ϕ is a function that is an extension to the right or an extension to the left (or both).

(b) The solution ϕ is called a *maximal* solution if, and only if, it does not admit any extensions in the sense of (a) (note that we require maximal solutions to be defined on *open* intervals, cf. Appendix E).

Remark 3.21. As an immediate consequence of the time reversion Lem. 1.9(b), if a solution $\phi : I \longrightarrow \mathbb{K}^n$, $n \in \mathbb{N}$, to (1.6), defined on some open interval $I \subseteq \mathbb{R}$, has an extension to the right (resp. to the left) if, and only if, $\psi : (-I) \longrightarrow \mathbb{K}^n$, $\psi(x) := \phi(-x)$, (solution to $y' = -f(-x, y)$) has an extension to the left (resp. to the right).

—

The existence of maximal solutions is not trivial – a priori it could be that every solution had an extension (analogous to the fact that to every $x \in [0, 1[$ (or every $x \in \mathbb{R}$) there is some bigger element in $[0, 1[$ (respectively in $\mathbb{R}$)).

Theorem 3.22. *Every solution $\phi_0 : I_0 \longrightarrow \mathbb{K}^n$ to* (1.4) *(resp. to* (1.6)*), defined on an open interval $I_0 \subseteq \mathbb{R}$, can be extended to a maximal solution of* (1.4) *(resp. of* (1.6)*).*

Proof. The proof is carried out for solutions to (1.4) (the implicit ODE) – the proof for solutions to the explicit ODE (1.6) is analogous and can also be seen as a special case. The idea is to apply Zorn's lemma. To this end, define a partial order on the set

$$\mathcal{S} := \{(I_0, \phi_0)\} \cup \{(I, \phi) : \phi : I \longrightarrow \mathbb{K}^n \text{ is solution to (1.4), extending } \phi_0\} \tag{3.62}$$

by letting

$$(I, \phi) \leq (J, \psi) \quad :\Leftrightarrow \quad I \subseteq J, \quad \psi\restriction_I = \phi. \tag{3.63}$$

Every chain $\mathcal{C}$, i.e. every totally ordered subset of $\mathcal{S}$, has an upper bound, namely $(I_\mathcal{C}, \phi_\mathcal{C})$ with $I_\mathcal{C} := \bigcup_{(I,\phi)\in\mathcal{C}} I$ and $\phi_\mathcal{C}(x) := \phi(x)$, where $(I, \phi) \in \mathcal{C}$ is chosen such that $x \in I$ (since $\mathcal{C}$ is a chain, the value of $\phi_\mathcal{C}(x)$ does not actually depend on the choice of $(I, \phi) \in \mathcal{C}$ and is, thus, well-defined).

Clearly, $I_\mathcal{C}$ is an open interval, $I_0 \subseteq I_\mathcal{C}$, and $\phi_\mathcal{C}$ extends ϕ_0 as a function; we still need to see that $\phi_\mathcal{C}$ is a solution to (1.4). For this, we, once again, use that $x \in I_\mathcal{C}$ means there exists $(I, \phi) \in \mathcal{C}$ such that $x \in I$ and ϕ is a solution to (1.4). Thus, using the notation from Def. 1.2(a),

$$(x, \phi_\mathcal{C}(x), \phi'_\mathcal{C}(x), \dots, \phi_\mathcal{C}^{(k)}(x)) = (x, \phi(x), \phi'(x), \dots, \phi^{(k)}(x)) \in U$$

and

$$F(x, \phi_\mathcal{C}(x), \phi'_\mathcal{C}(x), \dots, \phi_\mathcal{C}^{(k)}(x)) = F(x, \phi(x), \phi'(x), \dots, \phi^{(k)}(x)) = 0,$$

showing $\phi_\mathcal{C}$ is a solution to (1.4) as defined in Def. 1.2(a). In particular, $(I_\mathcal{C}, \phi_\mathcal{C}) \in \mathcal{S}$. To verify $(I_\mathcal{C}, \phi_\mathcal{C})$ is an upper bound for $\mathcal{C}$, note that the definition of $(I_\mathcal{C}, \phi_\mathcal{C})$ immediately implies $I \subseteq I_\mathcal{C}$ for each $(I, \phi) \in \mathcal{C}$ and $\phi_\mathcal{C}\restriction_I = \phi$ for each $(I, \phi) \in \mathcal{C}$.

To conclude the proof, we note that all hypotheses of Zorn's lemma have been verified such that it yields the existence of a maximal element of $(I_{\max}, \phi_{\max}) \in \mathcal{S}$, i.e. $\phi_{\max} : I_{\max} \longrightarrow \mathbb{K}^n$ must be a maximal solution extending ϕ_0. ■

Proposition 3.23. *Let $k, n \in \mathbb{N}$. Given $G \subseteq \mathbb{R} \times \mathbb{K}^{kn}$ open and $f : G \longrightarrow \mathbb{K}^n$ continuous, if $\phi : I \longrightarrow \mathbb{K}^n$ is a solution to* (1.6) *such that $I =]a, b[$, $a < b$, $b < \infty$ (resp. $-\infty < a$), then ϕ has an extension to the right (resp. to the left) if, and only if,*

$$\underset{(b,\eta_0,\dots,\eta_{k-1})\in G}{\exists} \quad \lim_{x\uparrow b} \big(\phi(x), \phi'(x), \dots, \phi^{(k-1)}(x)\big) = (\eta_0, \dots, \eta_{k-1}), \tag{3.64a}$$

$$\left(\textit{resp.} \quad \underset{(a,\eta_0,\dots,\eta_{k-1})\in G}{\exists} \quad \lim_{x\downarrow a} \big(\phi(x), \phi'(x), \dots, \phi^{(k-1)}(x)\big) = (\eta_0, \dots, \eta_{k-1}) \right). \tag{3.64b}$$

Proof. That the respective part of (3.64) is necessary for the existence of the respective extension is immediate from the fact that, for each solution to (1.6), the solution and all its derivatives up to order $k-1$ must exist and must be continuous.

We now prove that (3.64a) is also sufficient for the existence of an extension to the right (the sufficiency of (3.64b) for the existence of an extension to the left is then immediate from Rem. 3.21). So assume (3.64a) to hold and consider the initial value problem consisting of (1.6) and the initial conditions

$$\underset{j=0,\dots,k-1}{\forall} \quad y^{(j)}(b) = \eta_j.$$

By Cor. 3.10, there must exist $\epsilon > 0$ such that this initial value problem has a solution $\psi :]b-\epsilon, b+\epsilon[\longrightarrow \mathbb{K}^n$. We now show that ϕ extended to b via (3.64a) is still a solution to (1.6). First note the mean value theorem (cf. [Phi13a, Th. 9.17]) yields that $\phi^{(j)}(b) = \eta_j$ exists for $j = 1, \dots, k-1$ as a left-hand derivative. Moreover,

$$\lim_{x\uparrow b} \phi^{(k)}(x) = \lim_{x\uparrow b} f\big(x, \phi(x), \phi'(x), \dots, \phi^{(k-1)}(x)\big) = f(b, \eta_0, \dots, \eta_{k-1}),$$

showing $\phi^{(k)}(b) = f\big(b, \phi(b), \phi'(b), \dots, \phi^{(k-1)}(b)\big)$ (again employing the mean value theorem), which proves ϕ extended to b is a solution to (1.6). Finally, Lem. 1.7 ensures

$$\sigma :]a, b+\epsilon[\longrightarrow \mathbb{K}^n, \quad \sigma(x) := \begin{cases} \phi(x) & \text{for } x \le b, \\ \psi(x) & \text{for } x \ge b, \end{cases}$$

is a solution to (1.6) that extends ϕ to the right. ■

Proposition 3.24. *Let $k, n \in \mathbb{N}$, let $G \subseteq \mathbb{R} \times \mathbb{K}^{kn}$ be open, let $f : G \longrightarrow \mathbb{K}^n$ be continuous, and let $\phi : I \longrightarrow \mathbb{K}^n$ be a solution to* (1.6) *defined on the open interval I. Consider $x_0 \in I$ and let $\mathrm{gr}_+(\phi)$ (resp. $\mathrm{gr}_-(\phi)$) denote the graph of $(\phi, \dots, \phi^{(k-1)})$ for $x \ge x_0$ (resp. for $x \le x_0$):*

$$\mathrm{gr}_+(\phi) := \mathrm{gr}_+(\phi, x_0) := \big\{ \big(x, \phi(x), \dots, \phi^{(k-1)}(x)\big) \in G : x \in I, \, x \ge x_0 \big\}, \tag{3.65a}$$

$$\mathrm{gr}_-(\phi) := \mathrm{gr}_-(\phi, x_0) := \big\{ \big(x, \phi(x), \dots, \phi^{(k-1)}(x)\big) \in G : x \in I, \, x \le x_0 \big\}. \tag{3.65b}$$

If there exists a compact *set $K \subseteq G$ such that $\mathrm{gr}_+(\phi) \subseteq K$ (resp. $\mathrm{gr}_-(\phi) \subseteq K$), then ϕ has an extension $\psi : J \longrightarrow \mathbb{K}^n$ to the right (resp. to the left) such that*

$$\underset{\tilde{x}\in J}{\exists} \quad \big(\tilde{x}, \psi(\tilde{x}), \dots, \psi^{(k-1)}(\tilde{x})\big) \notin K. \tag{3.66}$$

The statement can be rephrased by saying that $\mathrm{gr}_+(\phi)$ (resp. $\mathrm{gr}_-(\phi)$) of each maximal *solution ϕ to* (1.6) *escapes from every compact subset of G when x appoaches the right (resp. the left) boundary of I (where the boundary of I can contain $-\infty$ and/or $+\infty$).*

Proof. We conduct the proof for extensions to the right; extensions to the left can be handled completely analogously (alternatively, one can apply the time reversion Lem. 1.9(b) as demonstrated in the last paragraph of the proof below). The proof for extensions to the right is divided into three steps. Let $K \subseteq G$ be compact.

Step 1: We show that $\mathrm{gr}_+(\phi) \subseteq K$ implies ϕ has an extension to the right: Since K is bounded, so is $\mathrm{gr}_+(\phi)$, implying

$$b := \sup I < \infty \tag{3.67}$$

as well as

$$M_1 := \sup\Big\{\big\|\phi^{(j)}(x)\big\| : \, j \in \{0,\dots,k-1\}, \; x \in [x_0, b[\Big\} < \infty.$$

In the usual way, K compact and f continuous imply

$$M_2 := \max\Big\{\|f(x,y)\| : \, (x,y) \in K\Big\} < \infty.$$

Set

$$M := \max\{M_1, M_2\}.$$

According to Prop. 3.23, we need to show (3.64a) holds. To this end, notice

$$\underset{j=0,\dots,k-1}{\forall} \quad \underset{x,\bar{x} \in [x_0,b[}{\forall} \quad \|\phi^{(j)}(x) - \phi^{(j)}(\bar{x})\| \leq M\,|x - \bar{x}| : \tag{3.68}$$

Indeed,

$$\underset{x,\bar{x} \in [x_0,b[}{\forall} \quad \|\phi^{(k-1)}(x) - \phi^{(k-1)}(\bar{x})\| = \left\| \int_x^{\bar{x}} f\big(t, \phi(t), \dots, \phi^{(k-1)}(t)\big)\,\mathrm{d}t \right\| \leq M\,|x - \bar{x}|,$$

and, for $0 \leq j < k-1$,

$$\underset{x,\bar{x} \in [x_0,b[}{\forall} \quad \|\phi^{(j)}(x) - \phi^{(j)}(\bar{x})\| = \left\| \int_x^{\bar{x}} \phi^{(j+1)}(t)\,\mathrm{d}t \right\| \leq M\,|x - \bar{x}|,$$

proving (3.68). Since K is compact, there exists a sequence $(x_m)_{m\in\mathbb{N}}$ in $[x_0, b[$ such that

$$\underset{(b,\eta_0,\dots,\eta_{k-1}) \in K}{\exists} \quad \lim_{m\to\infty} \big(x_m, \phi(x_m), \phi'(x_m), \dots, \phi^{(k-1)}(x_m)\big) = (b, \eta_0, \dots, \eta_{k-1}). \tag{3.69}$$

Using $\bar{x} := x_m$ in (3.68) yields, for $m \to \infty$,

$$\underset{j=0,\dots,k-1}{\forall} \quad \underset{x \in [x_0,b[}{\forall} \quad \|\phi^{(j)}(x) - \eta_j\| \leq M\,|x - b|,$$

implying

$$\underset{j=0,\dots,k-1}{\forall} \quad \lim_{x\uparrow b} \phi^{(j)}(x) = \eta_j,$$

i.e. (3.64a) holds, completing the proof of Step 1.

Step 2: We show that $\mathrm{gr}_+(\phi) \subseteq K$ implies ϕ can be extended to the right to $I \cup]x_0, b+\alpha[$, where $\alpha > 0$ does not depend on $b := \sup I$: Since K is compact, Cor. 3.9 guarantees every initial value problem

$$y^{(k)} = f(x, y, y', \dots, y^{(k-1)}), \tag{3.70a}$$

$$\underset{j=0,\dots,k-1}{\forall} \quad y^{(j)}(\xi_0) = y_{0,j}, \quad (\xi_0, y_0) \in K, \tag{3.70b}$$

has a solution defined on $]\xi_0 - \alpha, \xi_0 + \alpha[$ with the same $\alpha > 0$. As shown in Step 1, the solution ϕ can be extended into $b = \sup I$ such that it satisfies (3.70b) with $(\xi_0, y_0) = (b, \eta) \in K$. Thus, using Lem. 1.7, it can be pieced together with the solution to (3.70) given on $[b, b+\alpha[$ by Cor. 3.9, completing the proof of Step 2.

Step 3: We finally show that $\mathrm{gr}_+(\phi) \subseteq K$ implies ϕ has an extension $\psi : J \longrightarrow \mathbb{K}^n$ to the right such that (3.66) holds: We set $a := \inf I$ and $\phi_0 := \phi$. Then, by Step 2, ϕ_0 has an extension ϕ_1 defined on $]a, b+\alpha[$. Inductively, for each $m \geq 1$, either there exists $m_0 \leq m$ such that $\phi_{m_0} :]a, b + m_0\alpha[\longrightarrow \mathbb{K}^n$ is an extension of ϕ that can be used as ψ to conclude the proof (i.e. $\psi := \phi_{m_0}$ satisfies (3.66)) or ϕ_m can, once more, be extended to $]a, b + (m+1)\alpha[$. As K is bounded, $\{x \geq x_0 : (x, y) \in K\} \subseteq \mathbb{R}$ must also be bounded, say by $\mu \in \mathbb{R}$. Thus, (3.66) must be satisfied for some $\psi := \phi_m$ with $1 \leq m \leq (\mu - x_0)/\alpha$.

As mentioned above, one can argue completely analogous to the above proof to obtain that $\mathrm{gr}_-(\phi) \subseteq K$ implies ϕ to have an extension to the left, satisfying (3.66). Here we show how one, alternatively, can use the time reversion Lem. 1.9 to this end: Consider the map

$$h : \mathbb{R} \times \mathbb{K}^{kn} \longrightarrow \mathbb{R} \times \mathbb{K}^{kn}, \quad h(x, y_1, \dots, y_k) := (-x, y_1, \dots, (-1)^{k-1} y_k),$$

which clearly constitutes an $\mathbb{R}$-linear isomophism. Noting (1.6) and (1.32) are the same, we consider the time-reversed version (1.33) and observe $G_g = h(G)$ to be open, $h(K) \subseteq G_g$ to be compact, $g : G_g \longrightarrow \mathbb{K}^n$, $g = (-1)^k (f \circ h)$, to be continuous. If $\mathrm{gr}_-(\phi, x_0) \subseteq K$, then $\mathrm{gr}_+(\psi, -x_0) \subseteq h(K)$, where ψ is the solution to the time-reversed version (1.33), given by Lem. 1.9(b). Then ψ has an extension $\tilde{\psi}$ to the right, satisfying (3.66) with ψ replaced by $\tilde{\psi}$ and K replaced by $h(K)$. Then, by Rem. 3.21, ϕ must have an extension $\tilde{\phi}$ to the left, satisfying (3.66) with ψ replaced by $\tilde{\phi}$. ∎

In Th. 3.28 below, we will show that, for continuous $f : G \longrightarrow \mathbb{K}^n$, each maximal solution to (1.6) must go to the boundary of G in the sense of the following definition.

Definition 3.25. Let $k, n \in \mathbb{N}$, let $G \subseteq \mathbb{R} \times \mathbb{K}^{kn}$ be open, let $f : G \longrightarrow \mathbb{K}^n$, and let $\phi :]a, b[\longrightarrow \mathbb{K}^n$, $-\infty \leq a < b \leq \infty$, be a solution to (1.6). We say that the solution ϕ *goes to the boundary* of G for $x \to b$ (resp. for $x \to a$) if, and only if,

$$\underset{K \subseteq G \text{ compact}}{\forall} \quad \underset{x_0 \in]a,b[}{\exists} \quad \mathrm{gr}_+(\phi, x_0) \cap K = \emptyset \quad (\text{resp. } \mathrm{gr}_-(\phi, x_0) \cap K = \emptyset), \tag{3.71}$$

where $\mathrm{gr}_+(\phi, x_0)$ and $\mathrm{gr}_-(\phi, x_0)$ are defined as in (3.65) (with $I =]a, b[$). In other words, ϕ goes to the boundary of G for $x \to b$ (resp. for $x \to a$) if, and only if, the graph of $(\phi, \dots, \phi^{(k-1)})$ escapes every compact subset K of G forever for $x \to b$ (resp. for $x \to a$).

Proposition 3.26. *In the situation of Def.* 3.25, *if the solution ϕ goes to the boundary of G for $x \to b$, then one of the following conditions must hold:*

(i) $b = \infty$,

(ii) $b < \infty$ *and* $L := \limsup_{x \uparrow b} \left\| \left(\phi(x), \dots, \phi^{(k-1)}(x)\right) \right\| = \infty$,

(iii) $b < \infty$, $L < \infty$ *(L as defined in* (ii)*),* $G \neq \mathbb{R} \times \mathbb{K}^{kn}$ *(i.e.* $\partial G \neq \emptyset$*), and*

$$\lim_{x \uparrow b} \operatorname{dist}\Big(\big(x, \phi(x), \dots, \phi^{(k-1)}(x)\big), \partial G\Big) = 0. \tag{3.72}$$

An analogous statement is valid for the solution ϕ going to the boundary of G for $x \to a$.

Proof. The proof is carried out for $x \to b$; the proof for $x \to a$ is analogous. Assume (i) – (iii) are all false. Choose $c \in]a, b[$. Since (i) and (ii) are false,

$$\underset{0 \le M < \infty}{\exists} \quad \underset{x \in [c,b[}{\forall} \quad \left\| \left(\phi(x), \dots, \phi^{(k-1)}(x)\right) \right\| \le M.$$

If (iii) is false because $G = \mathbb{R} \times \mathbb{K}^{kn}$, then $K := \{(x, y) \in \mathbb{R} \times \mathbb{K}^{kn} : x \in [c, b], \|y\| \le M\}$ is a compact subset of G that shows (3.71) does not hold. In the only remaining case, (iii) must be false, since (3.72) does not hold. Thus,

$$\underset{\delta > 0}{\exists} \quad \underset{x_0 \in]a,b[}{\forall} \quad \underset{x_1 \in]x_0,b[}{\exists} \quad \operatorname{dist}\Big(\big(x_1, \phi(x_1), \dots, \phi^{(k-1)}(x_1)\big), \partial G\Big) \ge \delta.$$

Clearly, the set

$$A := \big\{(x, y) \in G : \operatorname{dist}\big((x, y), \partial G\big) \ge \delta\big\}$$

is closed (e.g. as the distance function $d : \mathbb{R} \times \mathbb{K}^{kn} \longrightarrow \mathbb{R}_0^+$, $d(\cdot) := \operatorname{dist}(\cdot, \partial G)$ is continuous (see Th. C.4) and $A = (d^{-1}[\delta, \infty[) \cap (G \cup \partial G))$. In consequence, $K \cap A$ with K as defined above is a compact subset of G that shows (3.71) does not hold. ∎

Remark 3.27. (a) Examples such as the second ODE of Ex. 3.30(b) below show that the lim sup in Prop. 3.26(ii) can not be replaced with a lim.

(b) If $f : G \longrightarrow \mathbb{K}^n$ is continuous, then the three conditions of Prop. 3.26 are also *sufficient* for ϕ to go to the boundary of G (cf. Cor. 3.29 below).

(c) For discontinuous $f : G \longrightarrow \mathbb{K}^n$, in general, (ii) of Prop. 3.26 is no longer sufficient for ϕ to go to the boundary of G as is shown by simple examples, whereas (i) and (iii) remain sufficient, even for discontinuous f (exercise). Similarly, simple examples show Prop. 3.24 becomes false without the assumption of f being continuous; and it can also happen that a maximal solution escapes every compact set, but still does not go to the boundary of G (exercise).

Theorem 3.28. *In the situation of Def.* 3.25, *if $f : G \longrightarrow \mathbb{K}^n$ is continuous and $\phi :]a, b[\longrightarrow \mathbb{K}^n$ is a* maximal *solution to* (1.6)*, then ϕ must go to the boundary of G for both $x \to a$ and $x \to b$, i.e., for both $x \to a$ and $x \to b$, it must escape every compact subset K of G forever and it must satisfy one of the conditions specified in Prop.* 3.26 *(and one of the analogous conditions for $x \to a$).*

Proof. We carry out the proof for $x \to b$ – the proof for $x \to a$ can be done analogously or by applying the time reversion Lem. 1.9, as indicated at the end of the proof below. Let $\phi :]a, b[\longrightarrow \mathbb{K}^n$ be a maximal solution to (1.6). Seeking a contradiction, we assume ϕ does not go to the boundary of G for $x \to b$, i.e. (3.71) does not hold and there exists a compact subset K of G and a strictly increasing sequence $(x_m)_{m\in\mathbb{N}}$ in $]a, b[$ such that $\lim_{m\to\infty} x_m = b < \infty$ and

$$\underset{m\in\mathbb{N}}{\forall} \quad \big(x_m, \phi(x_m), \dots, \phi^{(k-1)}(x_m)\big) \in K. \tag{3.73}$$

We now define C to be another compact subset of G that is strictly between K and G, i.e. $K \subsetneq C \subsetneq G$: More precisely, we choose $r > 0$ such that

$$C := \Big\{(x, y) \in \mathbb{R} \times \mathbb{K}^{kn} : \operatorname{dist}\big((x, y), K\big) \le r\Big\} \subseteq G,$$

where

$$\operatorname{dist}\big((x, y), K\big) = \inf\{\|(x, y) - (\tilde{x}, \tilde{y})\|_2 : (\tilde{x}, \tilde{y}) \in K\},$$

$\|\cdot\|_2$ denoting the Euclidean norm on $\mathbb{R}^{kn+1}$ for $\mathbb{K} = \mathbb{R}$ and the Euclidean norm on $\mathbb{R}^{2kn+1}$ for $\mathbb{K} = \mathbb{C}$ (this choice of norm is different from previous choices and will be convenient later during the current proof). As ϕ is a maximal solution, Prop. 3.24 guarantees the existence of another strictly increasing sequence $(\xi_m)_{m\in\mathbb{N}}$ in $]a, b[$ such that $\lim_{m\to\infty} \xi_m = b < \infty$, $x_1 < \xi_1 < x_2 < \xi_2 < \dots$ (i.e. $x_m < \xi_m < x_{m+1}$ for each $m \in \mathbb{N}$) and such that

$$\underset{m\in\mathbb{N}}{\forall} \quad \big(\xi_m, \phi(\xi_m), \dots, \phi^{(k-1)}(\xi_m)\big) \notin C.$$

Noting $\big(x_m, \phi(x_m), \dots, \phi^{(k-1)}(x_m)\big) \in K$ by (3.73) and $K \subseteq C$, define

$$\underset{m\in\mathbb{N}}{\forall} \quad s_m := \sup\Big\{s \ge x_m : \big(x, \phi(x), \dots, \phi^{(k-1)}(x)\big) \in C \text{ for each } x \in [x_m, s]\Big\}.$$

By the definition of s_m as a sup, $s_m < x_{m+1} < b < \infty$, and by the continuity of the distance function $d : \mathbb{R} \times \mathbb{K}^{kn} \longrightarrow \mathbb{R}_0^+$, $d(\cdot) := \operatorname{dist}(\cdot, K)$ (see Th. C.4), one obtains

$$\underset{m\in\mathbb{N}}{\forall} \quad \operatorname{dist}\Big(\big(s_m, \phi(s_m), \dots, \phi^{(k-1)}(s_m)\big),\, K\Big) = r,$$

in particular,

$$\underset{x\in[x_m, s_m]}{\forall} \quad \big(x, \phi(x), \dots, \phi^{(k-1)}(x)\big) \in C \tag{3.74}$$

and

$$\underset{m\in\mathbb{N}}{\forall} \quad \Big\|\big(x_m, \phi(x_m), \dots, \phi^{(k-1)}(x_m)\big) - \big(s_m, \phi(s_m), \dots, \phi^{(k-1)}(s_m)\big)\Big\|_2 \ge r. \tag{3.75}$$

We use the boundedness of the compact set C and (3.74) to provide

$$M_1 := \sup\Big\{\big\|\big(\phi(x), \dots, \phi^{(k-1)}(x)\big)\big\|_2 : x \in [x_m, s_m],\ m \in \mathbb{N}\Big\} < \infty,$$

$$M_2 := \max\Big\{\|f(x, y)\|_2 : (x, y) \in C\Big\} < \infty$$

(as C is compact and f continuous),

$$M := \max\{M_1, M_2\}.$$

We now notice that each function

$$J_m : [x_m, s_m] \longrightarrow \mathbb{R} \times \mathbb{K}^{kn}, \quad J_m(x) := \big(x, \phi(x), \dots, \phi^{(k-1)}(x)\big),$$

is a continuously differentiable curve or path (using the continuity of f), cf. Def. F.1 (for $\mathbb{K} = \mathbb{C}$, we consider J_m as a path in $\mathbb{R}^{2kn+1}$). To finish the proof, we will have to make use of the notion of *arc length* (cf. Def. F.5) of such a continuously differentiable curve: Recall that each such continuously differentiable path is rectifyable, i.e. it has a well-defined finite arc length $l(J_m)$ (cf. Th. F.7). Moreover, $l(J_m)$ satisfies

$$\begin{aligned}
\|J_m(x_m) - J_m(s_m)\|_2 &\overset{\text{(F.4)}}{\leq} l(J_m) \overset{\text{(F.17)}}{=} \int_{x_m}^{s_m} \|J_m'(x)\|_2 \,\mathrm{d}x \\
&= \int_{x_m}^{s_m} \sqrt{1 + \sum_{j=1}^{k} \|\phi^{(j)}(x)\|_2^2} \,\mathrm{d}x \\
&\leq \int_{x_m}^{s_m} \sqrt{1 + \big\|\big(\phi(x), \dots, \phi^{(k-1)}(x)\big)\big\|_2^2 + \big\|f(J_m(x))\big\|_2^2} \,\mathrm{d}x \\
&\leq \int_{x_m}^{s_m} \sqrt{1 + 2M^2} \,\mathrm{d}x \,,
\end{aligned} \tag{3.76}$$

where it was used that $\|\cdot\|_2$ was chosen to be the Euclidean norm. For each $m \in \mathbb{N}$, we estimate

$$\begin{aligned}
0 < r &\overset{\text{(3.75)}}{\leq} \Big\|\big(x_m, \phi(x_m), \dots, \phi^{(k-1)}(x_m)\big) - \big(s_m, \phi(s_m), \dots, \phi^{(k-1)}(s_m)\big)\Big\|_2 \\
&= \|J_m(x_m) - J_m(s_m)\|_2 \overset{\text{(3.76)}}{\leq} \int_{x_m}^{s_m} \sqrt{1 + 2M^2} \,\mathrm{d}x \\
&= (s_m - x_m)\sqrt{1 + 2M^2}.
\end{aligned} \tag{3.77}$$

However, $\lim_{m\to\infty}(s_m - x_m)\sqrt{1 + 2M^2} = 0$ due to $\lim_{m\to\infty} s_m = \lim_{m\to\infty} x_m = b$, in contradiction to $r > 0$. This contradiction shows our initial assumption that ϕ does not go to the boundary of G for $x \to b$ must have been wrong.

To obtain the remaining assertion that ϕ must go to the boundary of G for $x \to a$, one can proceed as in the last paragraph of the proof of Prop. 3.23, making use of the function h defined there and of the time reversion Lem. 1.9: If $K \subseteq G$ is a compact set and ψ is the solution to the time-reversed version given by Lem. 1.9(b), then ψ must be maximal as ϕ is maximal. Thus, for $x \to -a$, ψ must escape the compact set $h(K)$ forever by the first part of the proof above, implying ϕ must escape K forever for $x \to a$. ■

Corollary 3.29. *Let $k, n \in \mathbb{N}$, let $G \subseteq \mathbb{R} \times \mathbb{K}^{kn}$ be open, and let $f : G \longrightarrow \mathbb{K}^n$ be continuous. If $\phi :]a, b[\longrightarrow \mathbb{K}^n$, $a < b$, is a solution to* (1.6)*, then the following statements are equivalent:*

(i) *ϕ is a maximal solution.*

(ii) *ϕ must go to the boundary of G for both $x \to a$ and $x \to b$ in the sense defined in Def.* 3.25.

(iii) *ϕ satisfies one of the conditions specified in Prop.* 3.26 *and one of the analogous conditions for $x \to a$.*

Proof. (i) implies (ii) by Th. 3.28, (ii) implies (iii) by Prop. 3.26, and it is an exercise to show (iii) implies (i) (here, Prop. 3.23 is the clue). ■

Example 3.30. The following examples illustrate the different kinds of possible bahavior of maximal solutions listed in Prop. 3.26 (the different kinds of bahavior can already be seen for 1-dimensional ODE of first order):

(a) The initial value problem

$$y' = 0, \quad y(0) = -1,$$

has the maximal solution $\phi : \mathbb{R} \longrightarrow \mathbb{R}$, $\phi(x) = -1$ – here we have

$$G = \mathbb{R}^2, \quad f : G \longrightarrow \mathbb{R}, \quad f(x, y) = 0,$$

solution interval $I = \mathbb{R}$, $b := \sup I = \infty$, i.e. we are in Case (i) of Prop. 3.26.

(b) The initial value problem

$$y' = x^{-2}, \quad y(-1) = 1,$$

has the maximal solution $\phi :]-\infty, 0[\longrightarrow \mathbb{R}$, $\phi(x) = -x^{-1}$ – here we have

$$G =]-\infty, 0[\times \mathbb{R}, \quad f : G \longrightarrow \mathbb{R}, \quad f(x, y) = x^{-2},$$

solution interval $I =]-\infty, 0[$, $b := \sup I = 0$, $\lim_{x\uparrow 0} |\phi(x)| = \infty$ i.e. we are in Case (ii) of Prop. 3.26.

To obtain an example, where we are also in Case (ii) of Prop. 3.26, but where $\lim_{x\uparrow b} |\phi(x)|$, $b := \sup I$, does not exist, consider the initial value problem

$$y' = -\frac{1}{x^2} \sin\frac{1}{x} + \frac{1}{x^3} \cos\frac{1}{x}, \quad y\left(-\frac{1}{\pi}\right) = 0,$$

which has the maximal solution $\phi :]-\infty, 0[\longrightarrow \mathbb{R}$, $\phi(x) = x^{-1} \sin(x^{-1})$ (here $\limsup_{x\uparrow 0} |\phi(x)| = \infty$, but, as $\phi(-1/(k\pi)) = 0$ for each $k \in \mathbb{N}$, $\lim_{x\uparrow 0} |\phi(x)|$ does not exist) – here we have

$$G = (\mathbb{R} \setminus \{0\}) \times \mathbb{R}, \quad f : G \longrightarrow \mathbb{R}, \quad f(x, y) = -\frac{1}{x^2} \sin\frac{1}{x} + \frac{1}{x^3} \cos\frac{1}{x}.$$

To obtain an example, where we are again in Case (ii) of Prop. 3.26, but where $G = \mathbb{R}^2$, consider the initial value problem

$$y' = y^2, \quad y(-1) = 1,$$

which, as in the first example of (b), has the maximal solution $\phi :]-\infty, 0[\longrightarrow \mathbb{R}$, $\phi(x) = -x^{-1}$ – here we have

$$G = \mathbb{R}^2, \quad f : G \longrightarrow \mathbb{R}, \quad f(x,y) = y^2.$$

(c) The initial value problem

$$y' = -y^{-1}, \quad y(-1) = 1,$$

has the maximal solution $\phi :]-\infty, -\frac{1}{2}[\longrightarrow \mathbb{R}$, $\phi(x) = \sqrt{-2x-1}$ – here we have

$$G = \mathbb{R} \times (\mathbb{R} \setminus \{0\}), \quad f : G \longrightarrow \mathbb{R}, \quad f(x,y) = -y^{-1},$$

solution interval $I =]-\infty, -\frac{1}{2}[$, $b := \sup I = -\frac{1}{2}$, $\partial G = \mathbb{R} \times \{0\}$,

$$\lim_{x \uparrow b} (x, \phi(x)) = \left(-\frac{1}{2}, 0\right) \in \partial G,$$

i.e. we are in Case (iii) of Prop. 3.26.

An example, where we are in Case (iii) of Prop. 3.26, but where $\lim_{x \uparrow b} (x, \phi(x))$ does not exist, is given by the initial value problem

$$y' = -\frac{1}{x^2} \cos \frac{1}{x}, \quad y\left(-\frac{1}{\pi}\right) = 0,$$

has the maximal solution $\phi :]-\infty, 0[\longrightarrow \mathbb{R}$, $\phi(x) = \sin(1/x)$ – here we have

$$G = (\mathbb{R} \setminus \{0\}) \times \mathbb{R}, \quad f : G \longrightarrow \mathbb{R}, \quad f(x,y) = -\frac{1}{x^2} \cos \frac{1}{x},$$

solution interval $I =]-\infty, 0[$, $b := \sup I = 0$, $\partial G = \{0\} \times \mathbb{R}$,

$$\lim_{x \uparrow 0} \operatorname{dist}\big((x, \phi(x)), \partial G\big) = \lim_{x \uparrow 0} |x| = 0.$$

As a final example, where we are again in Case (iii) of Prop. 3.26, reconsider the initial value problem from (a), but this time with

$$G =]-1, 1[\times]-3, 5[, \quad f : G \longrightarrow \mathbb{R}, \quad f(x,y) = 0.$$

Now the maximal solution is $\phi :]-1, 1[\longrightarrow \mathbb{R}$, $\phi(x) = -1$, solution interval $I =]-1, 1[$, $b := \sup I = 1$, and $\lim_{x \uparrow 1} (x, \phi(x)) = (1, -1) \in \partial G$. This last example also illustrates that, even though it is quite common to omit an explicit specification of the domain G when writing an ODE (as we did in (a)) – where it is usually assumed that the intended domain can be guessed from the context – the maximal solution will typically depend on the specification of G.

Example 3.31. We have already seen examples of initial value problems that admit more than one maximal solution – for instance, the initial value problem of Ex. 1.4(b) had infinitely many different maximal solutions, all of them defined on all of $\mathbb{R}$. The following example shows that an initial value problem can have maximal solutions defined on different intervals: Let

$$G := \mathbb{R}\times]-1,1[, \quad f : G \longrightarrow \mathbb{R}, \quad f(x,y) := \frac{\sqrt{|y|}}{1-\sqrt{|y|}},$$

and consider the initial value problem

$$y' = f(x,y) = \frac{\sqrt{|y|}}{1-\sqrt{|y|}}, \quad y(0) = 0. \tag{3.78}$$

An obvious maximal solution is

$$\phi : \mathbb{R} \longrightarrow \mathbb{R}, \quad \phi(0) = 0.$$

However, another maximal solution (that can be found using separation of variables) is

$$\psi :]-1,1[\longrightarrow \mathbb{R}, \quad \psi(x) := \begin{cases} -\big(1-\sqrt{1+x}\big)^2 & \text{for } -1 < x \leq 0, \\ \big(1-\sqrt{1-x}\big)^2 & \text{for } 0 \leq x < 1. \end{cases}$$

To confirm the maximality of the solution ψ, note $\lim_{x\downarrow -1}(x,\psi(x)) = (-1,-1) \in \partial G$ and $\lim_{x\uparrow 1}(x,\psi(x)) = (1,1) \in \partial G$.

3.5 Continuity in Initial Conditions

The goal of the present section is to show that, under suitable conditions, small changes in the initial condition for an ODE result in small changes in the solution. As, in situations of nonuniqueness, we can change the solution without having changed the initial condition at all, ensuring unique solutions to initial value problems is a minimal prerequisite for our considerations in this section.

Definition 3.32. Let $G \subseteq \mathbb{R} \times \mathbb{K}^{kn}$, $k,n \in \mathbb{N}$, and $f : G \longrightarrow \mathbb{K}^n$. We say that the explicit n-dimensional kth-order ODE (1.6), i.e.

$$y^{(k)} = f\big(x, y, y', \dots, y^{(k-1)}\big), \tag{3.79a}$$

admits unique maximal solutions if, and only if, f is such that every initial value problem consisting of (3.79a) and

$$\underset{j\in\{0,\dots,k-1\}}{\forall} \quad y^{(j)}(\xi) = \eta_j \in \mathbb{K}^n, \tag{3.79b}$$

with $(\xi,\eta) \in G$, has a unique maximal solution $\phi_{(\xi,\eta)} : I_{(\xi,\eta)} \longrightarrow \mathbb{K}^n$ (combining Cor. 3.16 with Th. 3.22 yields that G being open and f being continuous and locally Lipschitz with respect to y is *sufficient* for (3.79a) to admit unique maximal solutions, but we know from Rem. 3.17 that this condition is not *necessary*). If f is such that (3.79a) admits unique maximal solutions, then

$$Y : D_f \longrightarrow \mathbb{K}^n, \quad Y(x,\xi,\eta) := \phi_{(\xi,\eta)}(x), \tag{3.80}$$

defined on

$$D_f := \{(x,\xi,\eta) \in \mathbb{R} \times G : x \in I_{(\xi,\eta)}\}, \tag{3.81}$$

is called the *global* or *general* solution to (3.79a). Note that the domain D_f of Y is determined entirely by f, which is notationally emphasized by its lower index f.

Lemma 3.33. *In the situation of Def.* 3.32*, the following holds:*

(a) $Y(\xi,\xi,\eta) = \eta_0$ *for each* $(\xi,\eta) \in G$.

(b) *If* $k = 1$*, then* $\eta = \eta_0$ *and* $Y\big(x,\tilde{x},Y(\tilde{x},\xi,\eta)\big) = Y(x,\xi,\eta)$ *for each* $(x,\xi,\eta), (\tilde{x},\xi,\eta) \in D_f$.

(c) *If* $k = 1$*, then* $Y\big(\xi,x,Y(x,\xi,\eta)\big) = \eta$ *for each* $(x,\xi,\eta) \in D_f$.

Proof. (a) holds as $Y(\cdot,\xi,\eta)$ is a solution to (3.79b). For (b) note $(\tilde{x},\xi,\eta) \in D_f$ implying $\big(\tilde{x},Y(\tilde{x},\xi,\eta)\big) \in G$, i.e. $\big(\tilde{x},Y(\tilde{x},\xi,\eta)\big)$ are admissible initial data. Moreover, $Y\big(\cdot,\tilde{x},Y(\tilde{x},\xi,\eta)\big)$ and $Y(\cdot,\xi,\eta)$ are both maximal solutions for some intial value problem for (3.79a). Since both solutions agree at $x = \tilde{x}$, both functions must be identical by the assumed uniqueness of solutions. In particular, they are defined for the same x and yield the same value at each x. Setting $x := \xi$ in (b) yields (c). ∎

The core of the proof of continuity in initial conditions as stated in Cor. 3.36 below is the following Th. 3.34(a), which provides continuity in initial conditions locally. As a byproduct, we will also obtain a version of the Picard-Lindelöf theorem in Th. 3.34(b), which states the local uniform convergence of the so-called Picard iteration, a method for obtaining approximate solutions that is quite different from the Euler method considered above.

Theorem 3.34. *Consider the situation of Def.* 3.32 *for first-order problems, i.e. with* $k = 1$*, and with* f *being continuous and locally Lipschitz with respect to* y *on* G *open. Fix an arbitrary norm* $\|\cdot\|$ *on* $\mathbb{K}^n$.

(a) *For each* $(\sigma,\zeta) \in G \subseteq \mathbb{R} \times \mathbb{K}^n$ *and each* $-\infty < a < b < \infty$ *such that* $[a,b] \subseteq I_{(\sigma,\zeta)}$ *(i.e., using the notation introduced in Def.* 3.32*, the maximal solution* $\phi_{(\sigma,\zeta)} = Y(\cdot,\sigma,\zeta)$ *is defined on* $[a,b]$*), there exists* $\delta > 0$ *satisfying:*

(i) *For every point* (ξ,η) *in the open set*

$$U_\delta(\sigma,\zeta) := \big\{(\xi,\eta) \in G : \xi \in]a,b[, \ \big\|\eta - Y(\xi,\sigma,\zeta)\big\| < \delta\big\}, \tag{3.82}$$

the maximal solution $\phi_{(\xi,\eta)} = Y(\cdot,\xi,\eta)$ *is defined on* $]a,b[$ *(i.e.* $]a,b[\subseteq I_{(\xi,\eta)}$*).*

(ii) *The restriction of the global solution* $(x, \xi, \eta) \mapsto Y(x, \xi, \eta)$ *to the open set*

$$W :=]a, b[\times U_\delta(\sigma, \zeta) \tag{3.83}$$

is continuous.

(b) (Picard-Lindelöf) *For each* $(\sigma, \zeta) \in G$, *there exists* $\alpha > 0$ *such that the* Picard iteration, *i.e. the sequence of functions* $(\phi_m)_{m\in\mathbb{N}_0}$, $\phi_m :]\sigma-\alpha, \sigma+\alpha[\longrightarrow \mathbb{K}^n$, *defined recursively by*

$$\phi_0(x) := \zeta, \tag{3.84a}$$

$$\underset{m\in\mathbb{N}_0}{\forall} \quad \phi_{m+1}(x) := \zeta + \int_\sigma^x f\big(t, \phi_m(t)\big)\,\mathrm{d}t\,, \tag{3.84b}$$

converges uniformly to the solution of the initial value problem (3.79) *(with* $k = 1$ *and* $(\xi, \eta) := (\sigma, \zeta)$*) on* $]\sigma - \alpha, \sigma + \alpha[$.

Proof. We will obtain (b) as an aside while proving (a). To simplify notation, we introduce the function

$$\psi : [a, b] \longrightarrow \mathbb{K}^n, \quad \psi(x) := Y(x, \sigma, \zeta).$$

Since $[a, b]$ is compact and ψ is continuous,

$$\gamma := (\mathrm{Id}, \psi)[a, b] = \{(x, \psi(x)) \in \mathbb{R} \times \mathbb{K}^n : x \in [a, b]\}$$

is a compact subset of G (cf. C.14). Thus, γ has a positive distance from the closed set $(\mathbb{R} \times \mathbb{K}^n) \setminus G$, implying

$$\underset{\delta_1>0}{\exists} \quad C := \big\{(x, y) \in \mathbb{R} \times \mathbb{K}^n : x \in [a, b], \ \big\|y - \psi(x)\big\| \leq \delta_1\big\} \subseteq G. \tag{3.85}$$

Clearly, C is bounded and C is also closed (using the continuity of the distance function $d : \mathbb{R} \times \mathbb{K}^n \longrightarrow \mathbb{R}_0^+$, $d(\cdot) := \mathrm{dist}(\cdot, \gamma)$, the continuity of the projection to the first component $\pi_1 : \mathbb{R} \times \mathbb{K}^n \longrightarrow \mathbb{R}$, and noting $C = d^{-1}[0, \delta_1] \cap \pi_1^{-1}[a, b]$). Thus, C is compact, and the hypothesis of f being locally Lipschitz with respect to y implies f to be globally Lipschitz with some Lipschitz constant $L \geq 0$ on the compact set C by Prop. 3.13. We can now choose the number $\delta > 0$ claimed to exist in (a) to be any number

$$0 < \delta < e^{-L(b-a)}\,\delta_1. \tag{3.86}$$

Since $-L(b - a) \leq 0$, we have

$$\delta < \delta_1. \tag{3.87}$$

Moreover, with d and π_1 as above, $U_\delta(\sigma, \zeta)$ as defined in (3.82) can be written in the form

$$U_\delta(\sigma, \zeta) = d^{-1}[0, \delta[\cap \pi_1^{-1}]a, b[,$$

showing it is an open set ($[0, \delta[$ is, indeed, open in $\mathbb{R}_0^+$).

Even though we are mostly interested in what happens on the open set W, it will be convenient to define functions on the slightly larger compact set

$$\begin{aligned}&\overline{W} := [a,b]\times\overline{U},\\&\overline{U} := \big\{(x,y)\in\mathbb{R}\times\mathbb{K}^n:\ x\in[a,b],\ \big\|y-\psi(x)\big\|\le\delta\big\} = d^{-1}[0,\delta]\ \cap\ \pi_1^{-1}[a,b].\end{aligned}$$

To proceed with the proof, we now carry out a form of the Picard iteration, recursively defining a sequence of functions $(\psi_m)_{m\in\mathbb{N}_0}$, $\psi_m:\overline{W}\longrightarrow\mathbb{K}^n$, defined recursively by

$$\psi_0(x,\xi,\eta) := \psi(x)+\eta-\psi(\xi), \tag{3.88a}$$

$$\underset{m\in\mathbb{N}_0}{\forall}\quad \psi_{m+1}(x,\xi,\eta) := \eta+\int_\xi^x f\big(t,\psi_m(t,\xi,\eta)\big)\,\mathrm{d}t\,. \tag{3.88b}$$

The proof will be concluded if we can show the $(\psi_m)_{m\in\mathbb{N}_0}$ constitute a sequence of continuous functions converging uniformly on W to $Y\restriction_W$. As an intermediate step, we establish the following properties of the ψ_m (simultaneously) by induction on $m\in\mathbb{N}_0$:

(1) ψ_m is continuous for each $m\in\mathbb{N}_0$.

(2) One has

$$\underset{\substack{m\in\mathbb{N}_0,\\(x,\xi,\eta)\in\overline{W}}}{\forall}\quad \big\|\psi_m(x,\xi,\eta)-\psi(x)\big\|<\delta_1\quad\big(\ \Rightarrow\quad (x,\psi_m(x,\xi,\eta))\in C\big).$$

In particular, since $C\subseteq G$, this shows the ψ_m are well-defined by (3.88b).

(3) One has

$$\underset{\substack{m\in\mathbb{N}_0,\\(x,\xi,\eta)\in\overline{W}}}{\forall}\quad \big\|\psi_{m+1}(x,\xi,\eta)-\psi_m(x,\xi,\eta)\big\|\le\frac{L^{m+1}\,|x-\xi|^{m+1}\,\delta}{(m+1)!}.$$

To start the induction proof, notice that the continuity of ψ implies the continuity of ψ_0. Moreover, if $(x,\xi,\eta)\in\overline{W}$, then

$$\big\|\psi_0(x,\xi,\eta)-\psi(x)\big\|\overset{(3.88a)}{=}\big\|\eta-\psi(\xi)\big\|=\big\|\eta-Y(\xi,\sigma,\zeta)\big\|\le\delta<\delta_1. \tag{3.89}$$

Also, from $\psi=Y(\cdot,\sigma,\zeta)=\phi_{(\sigma,\zeta)}$, we know, for each $x,\xi\in[a,b]$,

$$\psi(x)-\psi(\xi)=\zeta+\int_\sigma^x f(t,\psi(t))\,\mathrm{d}t\,-\zeta-\int_\sigma^\xi f(t,\psi(t))\,\mathrm{d}t\,=\int_\xi^x f(t,\psi(t))\,\mathrm{d}t$$

and, for each $(x,\xi,\eta)\in\overline{W}$,

$$\begin{aligned}\big\|\psi_1(x,\xi,\eta)-\psi_0(x,\xi,\eta)\big\| &= \left\|\int_\xi^x\Big(f\big(t,\psi_0(t,\xi,\eta)\big)-f\big(t,\psi(t)\big)\Big)\,\mathrm{d}t\right\|\\ &\overset{f\ L\text{-Lip.}}{\le} L\left|\int_\xi^x\big\|\psi_0(t,\xi,\eta)-\psi(t)\big\|\,\mathrm{d}t\right|\\ &= L\left|\int_\xi^x\big\|\eta-\psi(\xi)\big\|\,\mathrm{d}t\right|\le L\,|x-\xi|\,\delta,\end{aligned}$$

completing the proof of (1) – (3) for $m = 0$. For the induction step, let $m \in \mathbb{N}_0$.

It is left as an exercise to prove the continuity of ψ_{m+1}.

Using the triangle inequality, we estimate, for each $(x, \xi, \eta) \in \overline{W}$,

$$\begin{aligned}
&\big\|\psi_{m+1}(x,\xi,\eta) - \psi(x)\big\| \\
&\quad\le \sum_{j=0}^{m} \big\|\psi_{j+1}(x,\xi,\eta) - \psi_j(x,\xi,\eta)\big\| + \big\|\psi_0(x,\xi,\eta) - \psi(x)\big\| \\
&\quad\overset{(3.89),\text{ ind.hyp. for }(3)}{\le} \sum_{j=0}^{m} \frac{L^{j+1}\,|x-\xi|^{j+1}\,\delta}{(j+1)!} + \delta \le e^{L|x-\xi|}\,\delta \overset{(3.86)}{<} e^{L(b-a)}\,e^{-L(b-a)}\,\delta_1 = \delta_1,
\end{aligned}$$

establishing the estimate of (2) for $m+1$. To prove the estimate in (3) for m replaced by $m+1$, one estimates, for each $(x, \xi, \eta) \in \overline{W}$,

$$\begin{aligned}
\big\|\psi_{m+2}(x,\xi,\eta) - \psi_{m+1}(x,\xi,\eta)\big\| &\le \left|\int_\xi^x \Big\| f\big(t,\psi_{m+1}(t,\xi,\eta)\big) - f\big(t,\psi_m(t,\xi,\eta)\big)\Big\|\,\mathrm{d}t\right| \\
&\le L\left|\int_\xi^x \big\|\psi_{m+1}(t,\xi,\eta) - \psi_m(t,\xi,\eta)\big\|\,\mathrm{d}t\right| \\
&\overset{\text{ind.hyp.}}{\le} L\left|\int_\xi^x \frac{L^{m+1}\,|t-\xi|^{m+1}\,\delta}{(m+1)!}\,\mathrm{d}t\right| \\
&= \frac{L^{m+2}\,|x-\xi|^{m+2}\,\delta}{(m+2)!},
\end{aligned}$$

completing the induction proof of (1) – (3).

As a consequence of (3), for each $l, m \in \mathbb{N}_0$ such that $m > l$:

$$\underset{(x,\xi,\eta)\in\overline{W}}{\forall} \quad \big\|\psi_m(x,\xi,\eta) - \psi_l(x,\xi,\eta)\big\| \le \delta \sum_{j=l+1}^{m} \frac{L^j\,(b-a)^j}{j!}. \tag{3.90}$$

The convergence of the exponential series, thus, implies that $(\psi_m(x,\xi,\eta))_{m\in\mathbb{N}_0}$ is a Cauchy sequence for each $(x, \xi, \eta) \in \overline{W}$, yielding pointwise convergence of the ψ_m to some function $\tilde\psi : \overline{W} \longrightarrow \mathbb{K}^n$. Letting m tend to infinity in (3.90) then shows

$$\underset{(x,\xi,\eta)\in\overline{W}}{\forall} \quad \big\|\tilde\psi(x,\xi,\eta) - \psi_l(x,\xi,\eta)\big\| \le \delta \sum_{j=l+1}^{\infty} \frac{L^j\,(b-a)^j}{j!},$$

where the independence of the right-hand side with respect to $(x, \xi, \eta) \in \overline{W}$ proves $\psi_m \to \tilde\psi$ uniformly on $\overline{W}$. The uniform convergence together with (1) then implies $\tilde\psi$ to be continuous.

In the final step of the proof, we show $\tilde\psi = Y$ on W, i.e. $\tilde\psi(\cdot, \xi, \eta)$ solves (3.79) (with $k = 1$). By Th. 1.5, we need to show

$$\underset{(x,\xi,\eta)\in W}{\forall} \quad \tilde\psi(x,\xi,\eta) = \eta + \int_\xi^x f\big(t,\tilde\psi(t,\xi,\eta)\big)\,\mathrm{d}t \tag{3.91}$$

(then uniqueness of solutions implies $\tilde{\psi}(\cdot,\xi,\eta) = Y(\cdot,\xi,\eta)$). To verify (3.91), given $\epsilon > 0$, by the uniform convergence $\psi_m \to \tilde{\psi}$, choose $m \in \mathbb{N}$ sufficiently large such that

$$\underset{k\in\{m-1,m\}}{\forall} \quad \underset{(x,\xi,\eta)\in W}{\forall} \quad \left\|\tilde{\psi}(x,\xi,\eta) - \psi_k(x,\xi,\eta)\right\| < \epsilon$$

and estimate, for each $(x,\xi,\eta) \in W$,

$$\begin{aligned}
&\left\|\tilde{\psi}(x,\xi,\eta) - \eta - \int_\xi^x f\big(t,\tilde{\psi}(t,\xi,\eta)\big)\,\mathrm{d}t\right\| \\
&\leq \left\|\tilde{\psi}(x,\xi,\eta) - \psi_m(x,\xi,\eta)\right\| + \left\|\int_\xi^x \Big(f\big(t,\psi_{m-1}(t,\xi,\eta)\big) - f\big(t,\tilde{\psi}(t,\xi,\eta)\big)\Big)\,\mathrm{d}t\right\| \\
&< \epsilon + L\left|\int_\xi^x \left\|\psi_{m-1}(t,\xi,\eta) - \tilde{\psi}(t,\xi,\eta)\right\|\,\mathrm{d}t\right| \leq \epsilon + L\,\epsilon\,(b-a). \qquad (3.92)
\end{aligned}$$

As (3.92) holds for every $\epsilon > 0$, (3.91) must be true as well.

It is noted that we have, indeed, proved (b) as a byproduct, since we know (for example from the Peano Th. 3.8) that ψ must be defined on $[\sigma-\alpha,\sigma+\alpha]$ for some $\alpha > 0$ and then $\phi_m = \psi_m(\cdot,\sigma,\zeta)$ on $[\sigma-\alpha,\sigma+\alpha]$ for each $m \in \mathbb{N}_0$. ■

Theorem 3.35. *As in Th. 3.34, consider the situation of Def. 3.32 for first-order problems, i.e. with $k = 1$, and with f being continuous and locally Lipschitz with respect to y on G open. Then the global solution $(x,\xi,\eta) \mapsto Y(x,\xi,\eta)$ as defined in Def. 3.32 is continuous. Moreover, its domain D_f is open.*

Proof. Let $(x,\sigma,\zeta) \in D_f$. Then, using the notation from Def. 3.32, x is in the domain of the maximal solution $\phi_{(\sigma,\zeta)}$, i.e. $x \in I_{(\sigma,\zeta)}$. Since $I_{(\sigma,\zeta)}$ is open, there must be $-\infty < a < x < b < \infty$ such that $[a,b] \subseteq I_{(\sigma,\zeta)}$ and then Th. 3.34(a) implies the global solution Y to be continuous on W, where W as defined in (3.83) is an open neighborhood of (x,σ,ζ). In particular, (x,σ,ζ) is an interior point of D_f and Y is continuous at (x,σ,ζ). As (x,σ,ζ) was arbitrary, D_f must be open and Y must be continuous. ■

Corollary 3.36. *Consider the situation of Def. 3.32 with f being continuous and locally Lipschitz with respect to y on G open. Then the global solution $(x,\xi,\eta) \mapsto Y(x,\xi,\eta)$ as defined in Def. 3.32 is continuous. Moreover, its domain D_f is open.*

Proof. It was part of the exercise that proved Cor. 3.16 to show that the right-hand side F of the first-order problem equivalent to (3.79) in the sense of Th. 3.1 is continuous and locally Lipschitz with respect to y, provided f is continuous and locally Lipschitz with respect to y. Thus, according to Th. 3.35, the equivalent first-order problem has a continuous global solution $\Upsilon : D_F \longrightarrow \mathbb{K}^{kn}$, defined on some open set D_F. As a consequence of Th. 3.1(b), $Y = \Upsilon_1 : D_F \longrightarrow \mathbb{K}^n$ is the global solution to (3.79a). So we have $D_f = D_F$ and, as Υ is continuous, so is Y. ■

It is sometimes interesting to consider situations where the right-hand side f depends on some (vector of) parameters μ in addition to depending on x and y:

Definition 3.37. If $G \subseteq \mathbb{R} \times \mathbb{K}^{kn} \times \mathbb{K}^l$ with $k, n, l \in \mathbb{N}$, and $f : G \longrightarrow \mathbb{K}^n$ is such that, for each $(\xi, \eta, \mu) \in G$, the explicit n-dimensional kth-order initial value problem

$$y^{(k)} = f\big(x, y, y', \dots, y^{(k-1)}, \mu\big), \tag{3.93a}$$

$$\underset{j \in \{0,\dots,k-1\}}{\forall} \quad y^{(j)}(\xi) = \eta_j \in \mathbb{K}^n, \tag{3.93b}$$

has a unique maximal solution $\phi_{(\xi,\eta,\mu)} : I_{(\xi,\eta,\mu)} \longrightarrow \mathbb{K}^n$, then

$$Y : D_f \longrightarrow \mathbb{K}^n, \quad Y(x, \xi, \eta, \mu) := \phi_{(\xi,\eta,\mu)}(x), \tag{3.94}$$

defined on

$$D_f := \big\{(x, \xi, \eta, \mu) \in \mathbb{R} \times G : x \in I_{(\xi,\eta,\mu)}\big\}, \tag{3.95}$$

is called the *global* or *general* solution to (3.93a).

Corollary 3.38. *Consider the situation of Def.* 3.37 *with f being continuous and locally Lipschitz with respect to (y, μ) on G open. Then the global solution Y as defined in Def.* 3.94 *is continuous. Moreover, its domain D_f is open.*

Proof. We consider $k = 1$ (i.e. (3.93a) is of first order) – the case $k > 1$ can then, in the usual way, be obtained by applying Th. 3.1. To apply Th. 3.35 to the present situation, define the auxiliary function

$$F : G \longrightarrow \mathbb{K}^{n+l}, \quad F_j(x, y) := \begin{cases} f_j(x, y) & \text{for } j = 1, \dots, n, \\ 0 & \text{for } j = n + 1, \dots, n + l. \end{cases} \tag{3.96}$$

Then, since f is continuous and locally Lipschitz with respect to (y, μ), F is continuous and locally Lipschitz with respect to y, and we can apply Th. 3.35 to

$$y' = F(x, y), \tag{3.97a}$$

$$y(\xi) = (\eta, \mu), \tag{3.97b}$$

where $(\xi, \eta, \mu) \in G$. According to Th. 3.35, the global solution $\tilde{Y} : D_F \longrightarrow \mathbb{K}^{n+l}$ of (3.97a) is continuous on the open set D_F. Moreover, by the definition of F in (3.96), we have

$$\underset{(x,\xi,\eta,\mu) \in D_F}{\forall} \quad \tilde{Y}(x, \xi, \eta, \mu) = \begin{pmatrix} Y(x, \xi, \eta, \mu) \\ \mu \end{pmatrix},$$

where Y is as defined in (3.94). In particular, $D_f = D_F$ and the continuity of $\tilde{Y}$ implies the continuity of Y. ■

Example 3.39. As a simple example of a parametrized ODE, consider $f : \mathbb{R} \times \mathbb{K}^2 \longrightarrow \mathbb{K}$, $f(x, y, \mu) := \mu y$,

$$y' = f(x, y, \mu) = \mu\, y,$$

$$y(\xi) = \eta,$$

with the global solution

$$Y : \mathbb{R} \times \mathbb{R} \times \mathbb{K}^2 \longrightarrow \mathbb{K}, \quad Y(x, \xi, \eta, \mu) = \eta\, e^{\mu\,(x-\xi)}.$$

4 Linear ODE

4.1 Definition, Setting

In Sec. 2.2, we saw that the solution of one-dimensional first-order linear ODE was particularly simple. One can now combine the general theory of ODE with some linear algebra to obtain results for n-dimensional linear ODE and, equivalently, for linear ODE of higher order.

Notation 4.1. For $n \in \mathbb{N}$, let $\mathcal{M}(n, \mathbb{K})$ denote the set of all $n \times n$ matrices over $\mathbb{K}$.

Definition 4.2. Let $I \subseteq \mathbb{R}$ be a nontrivial interval, $n \in \mathbb{N}$, and let $A : I \longrightarrow \mathcal{M}(n, \mathbb{K})$ and $b : I \longrightarrow \mathbb{K}^n$ be continuous. An ODE of the form

$$y' = A(x)y + b(x) \tag{4.1}$$

is called an n-dimensional *linear ODE* of first order. It is called *homogeneous* if, and only if, $b \equiv 0$; it is called *inhomogeneous* if, and only if, it is not homogeneous.

—

Using the notion of matrix norm (cf. Sec. G), it is not hard to show the right-hand side of (4.1) is continuous and locally Lipschitz with respect to y and, thus, every initial value problem for (4.1) has a unique maximal solution (exercise). However, to show the maximal solution is always defined on all of I, we need some additional machinery, which is developed in the next section.

4.2 Gronwall's Inequality

In the current section, we will provide Gronwall's inequality, which is also of interest outside the field of ODE. Here, Gronwall's inequality will allow us to prove the global existence of maximal solutions for ODE with linearly bounded right-hand side – a corollary being that maximal solutions of (4.1) are always defined on all of I.

As an auxiliary tool on our way to Gronwall's inequality, we will now briefly study (one-dimensional) differential *inequalities*:

Definition 4.3. Given $G \subseteq \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, and $f : G \longrightarrow \mathbb{R}$, call

$$y' \le f(x, y) \tag{4.2}$$

a (one-dimensional) *differential inequality* (of first order). A *solution* to (4.2) is a differentiable function $w : I \longrightarrow \mathbb{R}$ defined on a nontrivial interval $I \subseteq \mathbb{R}$ satisfying the two conditions

(i) $\big\{\big(x, w(x)\big) \in I \times \mathbb{R} : x \in I\big\} \subseteq G$,

(ii) $w'(x) \le f\big(x, w(x)\big)$ for each $x \in I$.

Proposition 4.4. *Let $G \subseteq \mathbb{R}^2$ be open, let $f : G \longrightarrow \mathbb{R}$ be continuous and locally Lipschitz with respect to y, and let $-\infty < a < b \leq \infty$. If $w : [a, b[\longrightarrow \mathbb{R}$ is a solution to the differential inequality* (4.2) *and $\phi : [a, b[\longrightarrow \mathbb{R}$ is a solution to the corresponding ODE, then*

$$w(a) \leq \phi(a) \quad \Rightarrow \quad \underset{x \in [a,b[}{\forall} \quad w(x) \leq \phi(x). \tag{4.3}$$

Proof. Consider the auxiliary function

$$g : G \times \mathbb{R} \longrightarrow \mathbb{R}, \quad g(x, y, \mu) := f(x, y) + \mu \tag{4.4}$$

and the (parametrized) ODE

$$y' = g(x, y, \mu) = f(x, y) + \mu. \tag{4.5}$$

Since f is continuous and locally Lipschitz with respect to y, g is continuous and locally Lipschitz with respect to (y, μ). Thus, continuity in initial conditions as given by Cor. 3.38 applies, yielding the global solution $Y : D_g \longrightarrow \mathbb{R}$, $(x, \xi, \eta, \mu) \mapsto Y(x, \xi, \eta, \mu)$, to be continuous on the open set D_g.

We now consider an arbitrary compact subinterval $[a, c] \subseteq [a, b[$ with $a < c < b$, noting that it suffices to prove $w \leq \phi$ on every such interval $[a, c]$. The set

$$\gamma := (\mathrm{Id}, a, \phi(a), 0)[a, c] = \big\{(x, a, \phi(a), 0) : x \in [a, c]\big\} \tag{4.6}$$

is a compact subset of D_g and, thus,

$$\underset{\epsilon > 0}{\exists} \quad \gamma_\epsilon := \Big\{(x, \xi, \eta, \mu) \in \mathbb{R}^4 : \operatorname{dist}\big((x, \xi, \eta, \mu), \gamma\big) < \epsilon\Big\} \subseteq D_g. \tag{4.7}$$

If we choose the distance in (4.7) to be meant with respect to the max-norm on $\mathbb{R}^4$ and if $0 < \mu < \epsilon$, then $(x, a, \phi(a), \mu) \in \gamma_\epsilon$ for each $x \in [a, c]$, such that $\phi_\mu := Y(\cdot, a, \phi(a), \mu)$ is defined on (a superset of) $[a, c]$. We proceed to prove $w \leq \phi_\mu$ on $[a, c]$: Seeking a contradiction, assume there exists $x_0 \in [a, c]$ such that $w(x_0) > \phi_\mu(x_0)$. Due to the continuity of w and ϕ_μ, $w > \phi_\mu$ must then hold in an entire neighborhood of x_0. On the other hand, $w(a) \leq \phi(a) = \phi_\mu(a)$, such that, for

$$x_1 := \inf\big\{x < x_0 : w(t) > \phi_\mu(t) \text{ for each } t \in]x, x_0]\big\},$$

$a \leq x_1 < x_0$ and $w(x_1) = \phi_\mu(x_1)$. But then, for each sufficiently small $h > 0$,

$$w(x_1 + h) - w(x_1) > \phi_\mu(x_1 + h) - \phi_\mu(x_1),$$

implying

$$\begin{aligned} w'(x_1) &= \lim_{h \to 0} \frac{w(x_1 + h) - w(x_1)}{h} \geq \lim_{h \to 0} \frac{\phi_\mu(x_1 + h) - \phi_\mu(x_1)}{h} = \phi'_\mu(x_1) \\ &= g\big(x_1, \phi_\mu(x_1), \mu\big) = f\big(x_1, \phi_\mu(x_1)\big) + \mu > f\big(x_1, \phi_\mu(x_1)\big) = f\big(x_1, w(x_1)\big), \end{aligned} \tag{4.8}$$

in contradiction to w being a solution to (4.2).

Thus, $w \le \phi_\mu$ on $[a,c]$ holds for every $0 < \mu < \epsilon$, and continuity of Y on D_g yields,

$$\underset{x\in[a,c]}{\forall} \quad w(x) \le \lim_{\mu\to 0} \phi_\mu(x) = \lim_{\mu\to 0} Y\big(x,a,\phi(a),\mu\big) = Y\big(x,a,\phi(a),0\big) = \phi(x), \tag{4.9}$$

concluding the proof. ∎

Theorem 4.5 (Gronwall's Inequality)**.** *Let $I := [a,b[$, where $-\infty < a < b \le \infty$. If $\alpha,\beta,\gamma : I \longrightarrow \mathbb{R}$ are continuous and $\beta(x) \ge 0$ for each $x \in I$, then*

$$\underset{x\in I}{\forall} \quad \gamma(x) \le \alpha(x) + \int_a^x \beta(t)\,\gamma(t)\,\mathrm{d}t \tag{4.10}$$

implies

$$\underset{x\in I}{\forall} \quad \gamma(x) \le \alpha(x) + \int_a^x \alpha(t)\,\beta(t)\,\exp\left(\int_t^x \beta(s)\,\mathrm{d}s\right)\,\mathrm{d}t\,. \tag{4.11}$$

Proof. Defining the auxiliary functions $\psi, w : I \longrightarrow \mathbb{R}$,

$$\psi(x) := \gamma(x) - \alpha(x), \tag{4.12a}$$

$$w(x) := \int_a^x \beta(t)\gamma(t)\,\mathrm{d}t\,, \tag{4.12b}$$

(4.10) can be written as

$$\underset{x\in I}{\forall} \quad \psi(x) \le w(x).$$

Moreover, this implies

$$\underset{x\in I}{\forall} \quad w'(x) = \beta(x)\gamma(x) = \beta(x)\big(\alpha(x)+\psi(x)\big) \le \beta(x)w(x) + \alpha(x)\beta(x),$$

showing w satisfies the (linear) differential inequality

$$y' \le \beta(x)\,y + \alpha(x)\,\beta(x). \tag{4.13}$$

Continuously extending α and β to $x < a$ (e.g. using the constant extensions $\alpha(x) = \alpha(a)$ and $\beta(x) := \beta(a)$ for $x < a$), we can consider the linear ODE corresponding to (4.13) on all of $]-\infty, b[$. Using the initial condition $y(a) = w(a) = 0$, yields the unique solution (employing the variation of constants Th. 2.3)

$$\begin{aligned} &\phi :]-\infty, b[\longrightarrow \mathbb{R},\\ &\phi(x) := \exp\left(\int_a^x \beta(s)\,\mathrm{d}s\right) \int_a^x \exp\left(-\int_a^t \beta(s)\,\mathrm{d}s\right)\alpha(t)\,\beta(t)\,\mathrm{d}t\\ &\quad = \int_a^x \alpha(t)\,\beta(t)\,\exp\left(\int_t^x \beta(s)\,\mathrm{d}s\right)\,\mathrm{d}t\,. \end{aligned} \tag{4.14}$$

Finally, we apply Prop. 4.4 to conclude

$$\underset{x\in I}{\forall} \quad \psi(x) \le w(x) \overset{(4.3)}{\le} \phi(x) \overset{(4.14)}{=} \int_a^x \alpha(t)\,\beta(t)\,\exp\left(\int_t^x \beta(s)\,\mathrm{d}s\right)\,\mathrm{d}t\,, \tag{4.15}$$

which, taking into account $\psi = \gamma - \alpha$, establishes (4.11). ∎

Example 4.6. Let $I := [a, b[$, where $-\infty < a < b \leq \infty$. If $\beta, \gamma : I \longrightarrow \mathbb{R}$ are continuous, $\beta(x) \geq 0$ for each $x \in I$, and $C \in \mathbb{R}$, then

$$\underset{x \in I}{\forall} \quad \gamma(x) \leq C + \int_a^x \beta(t)\,\gamma(t)\,\mathrm{d}t \tag{4.16}$$

implies

$$\underset{x \in I}{\forall} \quad \gamma(x) \leq C \exp\left(\int_a^x \beta(t)\,\mathrm{d}t\right) : \tag{4.17}$$

We apply Gronwall's inequality of Th. 4.5 with $\alpha \equiv C$ together with the fundamental theorem of calculus to obtain the estimate

$$\begin{aligned}
\gamma(x) &\leq C + \int_a^x C\,\beta(t) \exp\left(\int_t^x \beta(s)\,\mathrm{d}s\right)\,\mathrm{d}t \\
&= C - C \int_a^x \left(-\beta(t) \exp\left(\int_x^t -\beta(s)\,\mathrm{d}s\right)\right)\,\mathrm{d}t = C - C\left[\exp\left(\int_x^t -\beta(s)\,\mathrm{d}s\right)\right]_a^x \\
&= C \exp\left(\int_a^x \beta(t)\,\mathrm{d}t\right)
\end{aligned} \tag{4.18}$$

for each $x \in I$, proving (4.17).

—

The following Th. 4.7 will be applied to show maximal solutions to linear ODE are always defined on all of I (with I as in Def. 4.2). However, Th. 4.7 is often also useful to obtain the domains of maximal solutions for nonlinear ODE.

Theorem 4.7. *Let $n \in \mathbb{N}$, let $I \subseteq \mathbb{R}$ be an open interval, and let $f : I \times \mathbb{K}^n \longrightarrow \mathbb{K}^n$ be continuous. If there exist nonnegative continuous functions $\gamma, \beta : I \longrightarrow \mathbb{R}_0^+$ such that*

$$\underset{(x,y) \in I \times \mathbb{K}^n}{\forall} \quad \|f(x,y)\| \leq \gamma(x) + \beta(x)\,\|y\|, \tag{4.19}$$

where $\|\cdot\|$ denotes some arbitrary norm on $\mathbb{K}^n$, then every maximal solution to

$$y' = f(x,y)$$

is defined on all of I.

Proof. Let $c < d$ and $\phi :]c, d[\longrightarrow \mathbb{K}^n$ be a solution to $y' = f(x,y)$. We prove that $d < b := \sup I$ implies ϕ can be extended to the right and $a := \inf I < c$ implies ϕ can be extended to the left. First, assume $d < b$ and let $x_0 \in]c, d[$. The idea is to apply Example 4.6 on the interval $[x_0, d[$. To this end, we estimate, for each $x \in [x_0, d[$:

$$\begin{aligned}
\|\phi(x)\| \quad &= \quad \left\|\phi(x_0) + \int_{x_0}^x f\big(t, \phi(t)\big)\,\mathrm{d}t\right\| \leq \|\phi(x_0)\| + \int_{x_0}^x \Big\|f\big(t, \phi(t)\big)\Big\|\,\mathrm{d}t \\
&\overset{(4.19)}{\leq} \quad \|\phi(x_0)\| + \int_{x_0}^x \gamma(t)\,\mathrm{d}t + \int_{x_0}^x \beta(t)\,\|\phi(t)\|\,\mathrm{d}t\,.
\end{aligned} \tag{4.20}$$

Since the continuous function γ is uniformly bounded on the compact interval $[x_0, d]$,

$$\underset{C\geq 0}{\exists} \quad \underset{x\in[x_0,d[}{\forall} \quad \|\phi(x)\| \leq C + \int_{x_0}^{x} \beta(t)\,\|\phi(t)\|\,\mathrm{d}t\,.$$

Thus, Example 4.6 applies, providing

$$\underset{x\in[x_0,d[}{\forall} \quad \|\phi(x)\| \leq C\,\exp\left(\int_{x_0}^{x} \beta(t)\,\mathrm{d}t\right) \leq C\,e^{M(d-x_0)}, \tag{4.21}$$

where $M \geq 0$ is a uniform bound for the continuous function β on the compact interval $[x_0, d]$. As (4.21) states that the graph

$$\mathrm{gr}_+(\phi) = \left\{\big(x, \phi(x)\big) \in G :\, x \in [x_0, d[\right\}$$

is contained in the compact set

$$K := [x_0, d] \times \left\{y \in \mathbb{K}^n :\, \|y\| \leq C\,e^{M(d-x_0)}\right\},$$

Prop. 3.24 implies ϕ has an extension to the right.

Now assume $a < c$. The idea is to apply the time reversion Lem. 1.9(b): According to Lem. 1.9(b), $\psi :]-d, -c[\longrightarrow \mathbb{K}^n$, $\psi(x) = \phi(-x)$, is a solution to $y' = -f(-x, y)$ and the first part of the prove above shows ψ to have an extension to the right. However, then Rem. 3.21 tells us ϕ has an extension to the left. ■

4.3 Existence, Uniqueness, Vector Space of Solutions

Theorem 4.8. *Consider the setting of Def.* 4.2 *with an open interval I. Then every initial value problem consisting of the linear ODE* (4.1) *and $y(x_0) = y_0$, $x_0 \in I$, $y_0 \in \mathbb{K}^n$, has a unique maximal solution $\phi : I \longrightarrow \mathbb{K}^n$ (note that ϕ is defined on all of I).*

Proof. It is an exercise to show the right-hand side of (4.1) is continuous and locally Lipschitz with respect to y. Thus, every initial value problem has a unique maximal solution by using Cor. 3.16 and Th. 3.22. That each maximal solution is defined on I follows from Th. 4.7, as

$$\underset{x\in I}{\forall} \quad \big\|A(x)y + b(x)\big\| \leq \big\|b(x)\big\| + \big\|A(x)\big\|\,\|y\|,$$

where $\big\|A(x)\big\|$ denotes the matrix norm of $A(x)$ induced by the norm $\|\cdot\|$ on $\mathbb{K}^n$ (cf. Appendix G). ■

We will now proceed to study the solution spaces of linear ODE – as it turns out, these solution spaces inherit the linear structure of the ODE.

Notation 4.9. Again, we consider the setting of Def. 4.2. Define $\mathcal{L}_\mathrm{i}$ and $\mathcal{L}_\mathrm{h}$ to be the respective sets of solutions to (4.1) and its homogeneous version, i.e.

$$\mathcal{L}_\mathrm{i} := \Big\{(\phi :\, I \longrightarrow \mathbb{K}^n) :\, \phi' = A\phi + b\Big\}, \tag{4.22a}$$

$$\mathcal{L}_\mathrm{h} := \Big\{(\phi :\, I \longrightarrow \mathbb{K}^n) :\, \phi' = A\phi\Big\}. \tag{4.22b}$$

Lemma 4.10. *Using Not.* 4.9*, we have*

$$\underset{\phi \in \mathcal{L}_\mathrm{i}}{\forall} \quad \mathcal{L}_\mathrm{i} = \phi + \mathcal{L}_\mathrm{h} = \{\phi + \psi : \psi \in \mathcal{L}_\mathrm{h}\}, \tag{4.23}$$

i.e. one obtains all solutions to the inhomogeneous equation (4.1) *by adding solutions of the homogeneous equation to a particular solution to the inhomogeneous equation (note that this is completely analogous to what occurs for solutions to linear systems of equations in linear algebra).*

Proof. Exercise. ∎

Theorem 4.11. *Let $I \subseteq \mathbb{R}$ be a nontrivial interval, $n \in \mathbb{N}$, and let $A : I \longrightarrow \mathcal{M}(n, \mathbb{K})$ be continuous. Then the following holds:*

(a) *The set $\mathcal{L}_\mathrm{h}$ defined in* (4.22b) *constitutes a vector space over $\mathbb{K}$.*

(b) *For each $k \in \mathbb{N}$ and $\phi_1, \dots, \phi_k \in \mathcal{L}_\mathrm{h}$, the following statements are equivalent:*

- (i) *The k functions $\phi_1, \dots, \phi_k$ are linearly independent over $\mathbb{K}$.*
- (ii) *There exists $x_0 \in I$ such that the k vectors $\phi_1(x_0), \dots, \phi_k(x_0) \in \mathbb{K}^n$ are linearly independent over $\mathbb{K}$.*
- (iii) *The k vectors $\phi_1(x), \dots, \phi_k(x) \in \mathbb{K}^n$ are linearly independent over $\mathbb{K}$ for every $x \in I$.*

(c) *The dimension of $\mathcal{L}_\mathrm{h}$ is n.*

Proof. (a): Exercise.

(b): (iii) trivially implies (ii). That (ii) implies (i) can easily be shown by contraposition: If (i) does not hold, then there is $(\lambda_1, \dots, \lambda_k) \in \mathbb{K}^k \setminus \{0\}$ such that $\sum_{j=1}^k \lambda_j \phi_j = 0$, i.e. $\sum_{j=1}^k \lambda_j \phi_j(x) = 0$ holds for each $x \in I$, i.e. (ii) does not hold. It remains to show (i) implies (iii). Once again, we accomplish this via contraposition: If (iii) does not hold, then there are $(\lambda_1, \dots, \lambda_k) \in \mathbb{K}^k \setminus \{0\}$ and $x \in I$ such that $\sum_{j=1}^k \lambda_j \phi_j(x) = 0$. But then, since $\sum_{j=1}^k \lambda_j \phi_j \in \mathcal{L}_\mathrm{h}$ by (a), $\sum_{j=1}^k \lambda_j \phi_j = 0$ (using uniqueness of solutions). In consequence, $\phi_1, \dots, \phi_k$ are linearly dependent and (i) does not hold.

(c): Let $(b_1, \dots, b_n)$ be a basis of $\mathbb{K}^n$ and $x_0 \in I$. Let $\phi_1, \dots, \phi_n \in \mathcal{L}_\mathrm{h}$ be the solutions to the initial conditions $y(x_0) = b_1, \dots, y(x_0) = b_n$, respectively. Then the $\phi_1, \dots, \phi_n$ must be linearly independent by (b) (as they are linearly independent at x_0), proving $\dim \mathcal{L}_\mathrm{h} \geq n$. On the other hand, if $\phi_1, \dots, \phi_k \in \mathcal{L}_\mathrm{h}$ are linearly independent, $k \in \mathbb{N}$, then, once more by (b), $\phi_1(x), \dots, \phi_k(x) \in \mathbb{K}^n$ are linearly independent for each $x \in \mathbb{K}^n$, showing $k \leq n$ and $\dim \mathcal{L}_\mathrm{h} \leq n$. ∎

Example 4.12. In Example 1.4(e), we had claimed that the second-order ODE (1.16) on $[a, b]$, $a < b$, namely

$$y'' = -y$$

had the set of solutions $\mathcal{L}$ as in (1.17), namely

$$\mathcal{L} = \Big\{ \big((c_1 \sin + c_2 \cos) : [a,b] \longrightarrow \mathbb{K} \big) : c_1, c_2 \in \mathbb{K} \Big\}.$$

We are now in a position to verify this claim: The second-order ODE (1.16) is equivalent to the homogeneous linear first-order ODE

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} y_2 \\ -y_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \tag{4.24}$$

with the vector space of solutions $\mathcal{L}_{\mathrm{h}}$ of dimension 2 over $\mathbb{K}$. Clearly, $\phi_1, \phi_2 \in \mathcal{L}_{\mathrm{h}}$, where $\phi_1, \phi_2 : [a,b] \longrightarrow \mathbb{K}^2$ with

$$\phi_1(x) := \begin{pmatrix} \sin x \\ \cos x \end{pmatrix}, \quad \phi_2(x) := \begin{pmatrix} \cos x \\ -\sin x \end{pmatrix}. \tag{4.25}$$

Moreover, ϕ_1 and ϕ_2 are linearly independent (e.g. since $\phi_1(0) = \binom{0}{1}$ and $\phi_2(0) = \binom{1}{0}$ are linearly independent, so are $\phi_1, \phi_2 : \mathbb{R} \longrightarrow \mathbb{K}^2$ by Th. 4.11(b), implying, again by Th. 4.11(b), the linear independence of $\phi_1(a), \phi_2(a)$, finally implying the linear independence of $\phi_1, \phi_2 : [a,b] \longrightarrow \mathbb{K}^2$). Thus,

$$\mathcal{L}_{\mathrm{h}} = \Big\{ \big((c_1\phi_1 + c_2\phi_2) : [a,b] \longrightarrow \mathbb{K}^2 \big) : c_1, c_2 \in \mathbb{K} \Big\} \tag{4.26}$$

and, since, according to Th. 3.1 the solutions to (1.16) are precisely the first components of solutions to (4.24), the representation (1.17) is verified.

4.4 Fundamental Matrix Solutions and Variation of Constants

Definition and Remark 4.13. A basis $(\phi_1, \dots, \phi_n)$, $n \in \mathbb{N}$, of the n-dimensional vector space $\mathcal{L}_{\mathrm{h}}$ over $\mathbb{K}$ is called a *fundamental system* for the linear ODE (4.1). One then also calls the matrix

$$\Phi := \begin{pmatrix} \phi_{11} & \dots & \phi_{1n} \\ \vdots & & \vdots \\ \phi_{n1} & \dots & \phi_{nn} \end{pmatrix}, \tag{4.27}$$

where the kth column of the matrix consists of the component functions $\phi_{1k}, \dots, \phi_{nk}$ of ϕ_k, $k \in \{1, \dots, n\}$, a *fundamental system* or a *fundamental matrix solution* for (4.1). The latter term is justified by the observation that $\Phi : I \longrightarrow \mathcal{M}(n, \mathbb{K})$ can be interpreted as a solution to the matrix-valued ODE

$$Y' = A(x)\, Y : \tag{4.28}$$

Indeed,

$$\Phi' = \big(\phi_1', \dots, \phi_n'\big) = \big(A(x)\,\phi_1, \dots, A(x)\,\phi_n\big) = A(x)\,\Phi.$$

Corollary 4.14. *Let $\phi_1, \dots, \phi_n \in \mathcal{L}_{\mathrm{h}}$, $n \in \mathbb{N}$, and let Φ be defined as in* (4.27)*. Then the following statements are equivalent:*

(i) Φ *is a fundamental system for* (4.1).

(ii) *There exists* $x_0 \in I$ *such that* $\det \Phi(x_0) \neq 0$.

(iii) $\det \Phi(x) \neq 0$ *for every* $x \in I$.

Proof. The equivalences are a direct consequence of the equivalences in Th. 4.11(b). ■

Theorem 4.15 (Variation of Constants)**.** *Consider the setting of Def.* 4.2*. If* $\Phi : I \longrightarrow \mathcal{M}(n, \mathbb{K})$ *is a fundamental system for* (4.1)*, then the unique solution* $\psi : I \longrightarrow \mathbb{K}^n$ *of the initial value problem consisting of* (4.1) *and* $y(x_0) = y_0$*,* $(x_0, y_0) \in I \times \mathbb{K}^n$*, is given by*

$$\psi : I \longrightarrow \mathbb{K}^n, \quad \psi(x) = \Phi(x)\Phi^{-1}(x_0)\, y_0 + \Phi(x) \int_{x_0}^{x} \Phi^{-1}(t)\, b(t)\, \mathrm{d}t \,. \tag{4.29}$$

Proof. The initial condition is easily verified:

$$\psi(x_0) = \Phi(x_0)\Phi^{-1}(x_0)\, y_0 + 0 = \mathrm{Id}\; y_0 = y_0.$$

To check that ψ satisfies (4.1), one computes, for each $x \in I$,

$$\begin{aligned} \psi'(x) &\overset{\text{(I.3)}}{=} \Phi'(x)\Phi^{-1}(x_0)\, y_0 + \Phi'(x) \int_{x_0}^{x} \Phi^{-1}(t)\, b(t)\, \mathrm{d}t \, + \Phi(x)\Phi^{-1}(x)\, b(x) \\ &\overset{(4.28)}{=} A(x)\, \Phi(x)\Phi^{-1}(x_0)\, y_0 + A(x)\, \Phi(x) \int_{x_0}^{x} \Phi^{-1}(t)\, b(t)\, \mathrm{d}t \, + b(x) \\ &= A(x)\, \psi(x) + b(x), \end{aligned} \tag{4.30}$$

thereby establishing the case. ■

Remark 4.16. The 1-dimensional variation of constants formula (2.2) is actually a special case of (4.29): We note that, for $n = 1$ and $A(x) = a(x)$, the solution $\Phi := \phi_0$ to the 1-dimensional homogeneous equation as defined in (2.2b), i.e.

$$\phi_0 : I \longrightarrow \mathbb{K}, \quad \phi_0(x) = \exp\left(\int_{x_0}^{x} a(t)\, \mathrm{d}t \right) = e^{\int_{x_0}^{x} a(t)\, \mathrm{d}t}$$

constitutes a fundamental matrix solution in the sense of Def. and Rem. 4.13 (since $1/\phi_0$ exists). Taking into account $\Phi(x_0) = \phi_0(x_0) = 1$, we obtain, for each $x \in I$,

$$\begin{aligned} \phi(x) &\overset{(4.29)}{=} \Phi(x)\Phi^{-1}(x_0)\, y_0 + \Phi(x) \int_{x_0}^{x} \Phi^{-1}(t)\, b(t)\, \mathrm{d}t \\ &= \phi_0(x) \left(y_0 + \int_{x_0}^{x} \phi_0(t)^{-1}\, b(t)\, \mathrm{d}t \right), \end{aligned} \tag{4.31}$$

which is (2.2a).

—

In Sec. 4.6, we will study methods for actually finding fundamental matrix solutions in cases where A is constant. However, in general, fundamental matrix solutions are often not explicitly available. In such situations, the following Th. 4.17 can sometimes help to extract information about solutions.

Theorem 4.17 (Liouville's Formula)**.** *Consider the setting of Def.* 4.2 *and recall the* trace *of an $n \times n$ matrix $A = (a_{kl})$ is defined by*

$$\operatorname{tr} A := \sum_{k=1}^{n} a_{kk}.$$

If $\Phi : I \longrightarrow \mathcal{M}(n, \mathbb{K})$ is a fundamental system for (4.1), *then*

$$\forall_{x_0, x \in I} \quad \det\big(\Phi(x)\big) = \det\big(\Phi(x_0)\big) \exp\left(\int_{x_0}^{x} \operatorname{tr}\big(A(t)\big)\,\mathrm{d}t\right). \tag{4.32}$$

Proof. Exercise. ∎

4.5 Higher-Order, Wronskian

In Th. 3.1, we saw that higher-order ODE are equivalent to systems of first-order ODE. We can now combine Th. 3.1 with our findings regarding first-order linear ODE to help with the solution of higher-order linear ODE.

Definition 4.18. Let $I \subseteq \mathbb{R}$ be a nontrivial interval, $n \in \mathbb{N}$. Let $b : I \longrightarrow \mathbb{K}$ and $a_0, \dots, a_{n-1} : I \longrightarrow \mathbb{K}$ be continuous functions. Then a (1-dimensional) *linear ODE* of *nth order* is an equation of the form

$$y^{(n)} = a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y + b(x). \tag{4.33}$$

It is called *homogeneous* if, and only if, $b \equiv 0$; it is called *inhomogeneous* if, and only if, it is not homogeneous. Analogous to (4.22), define the respective sets of solutions

$$\mathcal{H}_{\mathrm{i}} := \left\{(\phi : I \longrightarrow \mathbb{K}) : \phi^{(n)} = b + \sum_{k=0}^{n-1} a_k \phi^{(k)}\right\}, \tag{4.34a}$$

$$\mathcal{H}_{\mathrm{h}} := \left\{(\phi : I \longrightarrow \mathbb{K}) : \phi^{(n)} = \sum_{k=0}^{n-1} a_k \phi^{(k)}\right\}. \tag{4.34b}$$

Definition 4.19. Let $I \subseteq \mathbb{R}$ be a nontrivial interval, $n \in \mathbb{N}$. For each n-tuple of $(n-1)$ times differentiable functions $\phi_1, \dots, \phi_n : I \longrightarrow \mathbb{K}$, define the *Wronskian*

$$W(\phi_1, \dots, \phi_n) : I \longrightarrow \mathbb{K}, \quad W(\phi_1, \dots, \phi_n)(x) := \det \begin{pmatrix} \phi_1(x) & \dots & \phi_n(x) \\ \phi_1'(x) & \dots & \phi_n'(x) \\ \vdots & & \vdots \\ \phi_1^{(n-1)}(x) & \dots & \phi_n^{(n-1)}(x) \end{pmatrix}. \tag{4.35}$$

Theorem 4.20. *Consider the setting of Def. 4.18.*

(a) *If $\mathcal{H}_\mathrm{i}$ and $\mathcal{H}_\mathrm{h}$ are the sets defined in (4.34), then $\mathcal{H}_\mathrm{h}$ is an n-dimensional vector space over $\mathbb{K}$ and, if $\phi \in \mathcal{H}_\mathrm{i}$ is arbitrary, then*

$$\mathcal{H}_\mathrm{i} = \phi + \mathcal{H}_\mathrm{h}. \tag{4.36}$$

(b) *Let $\phi_1, \dots, \phi_n \in \mathcal{H}_\mathrm{h}$. Then the following statements are equivalent:*

(i) *$\phi_1, \dots, \phi_n$ are linearly independent over $\mathbb{K}$ (i.e. $(\phi_1, \dots, \phi_n)$ forms a basis of $\mathcal{H}_\mathrm{h}$).*

(ii) *There exists $x_0 \in I$ such that the Wronskian does not vanish:*

$$W(\phi_1, \dots, \phi_n)(x_0) \neq 0.$$

(iii) *The Wronskian never vanishes, i.e. $W(\phi_1, \dots, \phi_n)(x) \neq 0$ for every $x \in I$.*

Proof. According to Th. 3.1, (4.33) is equivalent to the first-order linear ODE

$$\begin{aligned} y' &= \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ a_0(x) & a_1(x) & a_2(x) & \dots & a_{n-2}(x) & a_{n-1}(x) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-2} \\ y_{n-1} \\ y_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ b(x) \end{pmatrix} \\ &=: \tilde{A}(x)\, y + \tilde{b}(x). \end{aligned} \tag{4.37}$$

Define

$$\begin{aligned} \mathcal{L}_\mathrm{i} &:= \Big\{ (\Phi : I \longrightarrow \mathbb{K}^n) : \Phi' = \tilde{A}\Phi + \tilde{b} \Big\}, \\ \mathcal{L}_\mathrm{h} &:= \Big\{ (\Phi : I \longrightarrow \mathbb{K}^n) : \Phi' = \tilde{A}\phi \Big\}. \end{aligned}$$

(a): Let $\phi \in \mathcal{H}_\mathrm{i}$ and define

$$\Phi := \begin{pmatrix} \phi \\ \phi' \\ \vdots \\ \phi^{(n-1)} \end{pmatrix}.$$

Then

$$\mathcal{H}_\mathrm{h} \overset{\text{Th. 3.1(a),(b)}}{=} \{\Psi_1 : \Psi \in \mathcal{L}_\mathrm{h}\}$$

and

$$\begin{aligned} \mathcal{H}_\mathrm{i} &\overset{\text{Th. 3.1(a),(b)}}{=} \{\tilde{\Phi}_1 : \tilde{\Phi} \in \mathcal{L}_\mathrm{i}\} \overset{(4.23)}{=} \{\tilde{\Phi}_1 : \tilde{\Phi} \in \Phi + \mathcal{L}_\mathrm{h}\} \\ &= \{(\Phi + \Psi)_1 : \Psi \in \mathcal{L}_\mathrm{h}\} = \phi + \mathcal{H}_\mathrm{h}. \end{aligned}$$

As a consequence of Th. 3.1, the map $J : \mathcal{L}_\text{h} \longrightarrow \mathcal{H}_\text{h}$, $J(\Phi) := \Phi_1$, is a linear isomorphism, implying that $\mathcal{H}_\text{h}$, like $\mathcal{L}_\text{h}$, is an n-dimensional vector space over $\mathbb{K}$.

(b): For $\phi_1, \dots, \phi_n \in \mathcal{H}_\text{h}$, define $\Phi_{kl} := \phi_k^{(l-1)}$ for each $k, l \in \{1, \dots, n\}$ and

$$\underset{x \in I}{\forall} \quad \Phi(x) := (\Phi_1(x), \dots, \Phi_n(x)) = (\Phi_{kl}(x)) = \begin{pmatrix} \phi_1(x) & \dots & \phi_n(x) \\ \phi_1'(x) & \dots & \phi_n'(x) \\ \vdots & & \vdots \\ \phi_1^{(n-1)}(x) & \dots & \phi_n^{(n-1)}(x) \end{pmatrix}$$

such that $\det \Phi(x) = W(\phi_1, \dots, \phi_n)(x)$ for each $x \in I$. Since Th. 3.1 yields $\Phi_1, \dots, \Phi_n \in \mathcal{L}_\text{h}$ if, and only if, $\phi_1, \dots, \phi_n \in \mathcal{H}_\text{h}$, the equivalences of (b) follow from the equivalences of Cor. 4.14. ■

Example 4.21. Consider $a_0, a_1 : \mathbb{R}^+ \longrightarrow \mathbb{K}$, $a_1(x) := 1/(2x)$, $a_0(x) := -1/(2x^2)$, and

$$y'' = a_1(x)\, y' + a_0(x)\, y = \frac{y'}{2x} - \frac{y}{2x^2}.$$

One might be able to guess the solutions

$$\phi_1, \phi_2 : \mathbb{R}^+ \longrightarrow \mathbb{K}, \quad \phi_1(x) := x, \quad \phi_2(x) := \sqrt{x}.$$

The Wronskian is

$$\begin{aligned} &W(\phi_1, \phi_2) : \mathbb{R}^+ \longrightarrow \mathbb{K}, \\ &W(\phi_1, \phi_2)(x) = \det \begin{pmatrix} x & \sqrt{x} \\ 1 & 1/(2\sqrt{x}) \end{pmatrix} = \frac{\sqrt{x}}{2} - \sqrt{x} = -\frac{\sqrt{x}}{2} < 0, \end{aligned}$$

i.e. ϕ_1 and ϕ_2 span $\mathcal{H}_\text{h}$ according to Th. 4.20(b):

$$\mathcal{H}_\text{h} = \{c_1\phi_1 + c_2\phi_2 : c_1, c_2 \in \mathbb{K}\}.$$

4.6 Constant Coefficients

For 1-dimensional first-order linear ODE, we obtained a solution formula in Th. 2.3 in terms of integrals (of course, in general, evaluating integrals can still be very difficult, and one might need effective and efficient numerical methods). In the previous sections, we have studied systems of first-order linear ODE as well as linear ODE of higher order. Unfortunately, there are no general solution formulas for these situations (one can use (4.29) if one knows a fundamental system, but the problem is the absence of a general procedure to obtain such a fundamental system). However, there is a more satisfying solution theory for the situation of so-called *constant coefficients*, i.e. if A in (4.1) and the $a_0, \dots, a_{n-1}$ in (4.33) do not depend on x.

4.6.1 Linear ODE of Higher Order

Definition 4.22. Let $I \subseteq \mathbb{R}$ be a nontrivial interval, $n \in \mathbb{N}$. Let $b : I \longrightarrow \mathbb{K}$ be continuous and $a_0, \dots, a_{n-1} \in \mathbb{K}$. Then a (1-dimensional) *linear ODE* of *nth order* with *constant coefficients* is an equation of the form

$$y^{(n)} = a_{n-1}\, y^{(n-1)} + \dots + a_1\, y' + a_0\, y + b(x). \tag{4.38}$$

—

In the present context, it is useful to introduce the following notation:

Notation 4.23. Let $n \in \mathbb{N}_0$.

(a) Let $\mathcal{P}$ denote the set of all polynomials over $\mathbb{K}$, $\mathcal{P}_n := \{P \in \mathcal{P} : \deg P \leq n\}$. We will also write $\mathcal{P}[\mathbb{R}]$, $\mathcal{P}[\mathbb{C}]$, $\mathcal{P}_n[\mathbb{R}]$, $\mathcal{P}_n[\mathbb{C}]$ if we need to be specific about the field of coefficients.

(b) Let $I \subseteq \mathbb{R}$ be a nontrivial interval. Let $D^n(I) := D^n(I, \mathbb{K})$ denote the set of all n times differentiable functions $f : I \longrightarrow \mathbb{K}$, and let

$$\partial_x : D^1(I) \longrightarrow \mathcal{F}(I, \mathbb{K}) := \{f : I \longrightarrow \mathbb{K}\}, \quad \partial_x f := f',$$

and, for each $P \in \mathcal{P}_n$ with $P(x) = \sum_{j=0}^{n} a_j x^j$ $(a_0, \dots, a_n \in \mathbb{K})$ define the *differential operator*

$$P(\partial_x) : D^n(I) \longrightarrow \mathcal{F}(I, \mathbb{K}), \quad P(\partial_x) f := \sum_{j=0}^{n} a_j \partial_x^j f = \sum_{j=0}^{n} a_j f^{(j)}. \tag{4.39}$$

Remark 4.24. Using Not. 4.23(b), the ODE (4.38) can be written concisely as

$$P(\partial_x) y = b(x), \quad \text{where} \quad P(x) := x^n - \sum_{j=0}^{n-1} a_j x^j. \tag{4.40}$$

—

The following Prop. 4.25 implies that the differential operator $P(\partial_x)$ does not, actually, depend on the representation of the polynomial P.

Proposition 4.25. *Let $P, P_1, P_2 \in \mathcal{P}$.*

(a) *If $P = P_1 + P_2$ and $n := \max\{\deg P_1, \deg P_2\}$, then*

$$\underset{f \in D^n(I)}{\forall} \quad P(\partial_x) f = P_1(\partial_x) f + P_2(\partial_x) f.$$

(b) *If $P = P_1 P_2$ and $n := \max\{\deg P, \deg P_1, \deg P_2\}$, then*

$$\underset{f \in D^n(I)}{\forall} \quad P(\partial_x) f = P_1(\partial_x)\big(P_2(\partial_x) f\big).$$

Proof. Exercise. ■

Lemma 4.26. *Let* $\lambda \in \mathbb{K}$ *and*

$$f : \mathbb{R} \longrightarrow \mathbb{K}, \quad f(x) := e^{\lambda x}. \tag{4.41}$$

Then, for each $P \in \mathcal{P}$,

$$P(\partial_x)f : \mathbb{R} \longrightarrow \mathbb{K}, \quad P(\partial_x)f(x) = P(\lambda)\, e^{\lambda x}. \tag{4.42}$$

Proof. There exists $n \in \mathbb{N}_0$ and $a_0, \dots, a_n \in \mathbb{K}$ such that $P(x) = \sum_{j=0}^{n} a_j x^j$. One computes

$$P(\partial_x)f(x) = \sum_{j=0}^{n} a_j\, \partial_x^j\, e^{\lambda x} = \sum_{j=0}^{n} a_j \lambda^j e^{\lambda x} = e^{\lambda x}\, P(\lambda),$$

proving (4.42). ■

Theorem 4.27. *If* $a_0, \dots, a_{n-1} \in \mathbb{K}$, $n \in \mathbb{N}$, *and* $P(x) = x^n - \sum_{j=0}^{n-1} a_j x^j$ *has the distinct zeros* $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ *(i.e.* $P(\lambda_1) = \dots = P(\lambda_n) = 0$*), then* $(\phi_1, \dots, \phi_n)$*, where*

$$\underset{j \in \{1,\dots,n\}}{\forall} \quad \phi_j : I \longrightarrow \mathbb{K}, \quad \phi_j(x) := e^{\lambda_j x}, \tag{4.43}$$

is a basis of the homogeneous solution space

$$\mathcal{H}_{\mathrm{h}} = \Big\{ (\phi : I \longrightarrow \mathbb{K}) : P(\partial_x)\phi = 0 \Big\} \tag{4.44}$$

to (4.38) *(i.e. to* (4.40)*).*

Proof. It is immediate from (4.42) and $P(\lambda_j) = 0$ that each ϕ_j satisfies $P(\partial_x)\phi_j = 0$. From Th. 4.20(a), we already know $\mathcal{H}_{\mathrm{h}}$ is an n-dimensional vector space over $\mathbb{K}$. Thus, it merely remains to compute the Wronskian. One obtains (cf. (4.35)):

$$W(\phi_1, \dots, \phi_n)(0) = \begin{vmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_n \\ \vdots & & \vdots \\ \lambda_1^{n-1} & \dots & \lambda_n^{n-1} \end{vmatrix} \overset{\text{(H.2)}}{=} \prod_{\substack{k,l=0 \\ k>l}}^{n-1} (\lambda_k - \lambda_l) \neq 0,$$

since the λ_j are all distinct. We have used that the Wronskian, in the present case, turns out to be a Vandermonde determinant. The formula (H.2) for this type of determinant is provided and proved in Appendix H. We also used that the determinant of a matrix is the same as the determinant of its transpose: $\det A = \det A^{\mathrm{t}}$. From $W(\phi_1, \dots, \phi_n)(0) \neq 0$ and Th. 4.20(b), we conclude that $(\phi_1, \dots, \phi_n)$ is a basis of $\mathcal{H}_{\mathrm{h}}$. ■

Example 4.28. We consider the third-order linear ODE

$$y''' = 2y'' - y' + 2y, \tag{4.45}$$

which can be written as $P(\partial_x)y = 0$ with

$$P(x) := x^3 - 2x^2 + x - 2 = (x^2+1)(x-2) = (x-i)(x+i)(x-2), \tag{4.46}$$

i.e. P has the distinct zeros $\lambda_1 = i$, $\lambda_2 = -i$, $\lambda_3 = 2$. Thus, according to Th. 4.27, the three functions

$$\phi_1, \phi_2, \phi_3 : \mathbb{R} \longrightarrow \mathbb{C}, \quad \phi_1(x) = e^{ix}, \quad \phi_2(x) = e^{-ix}, \quad \phi_3(x) = e^{2x}, \tag{4.47}$$

form a basis of the $\mathbb{C}$-vector space $\mathcal{H}_\mathrm{h}$. If we consider (4.45) as an ODE over $\mathbb{R}$, then we are interested in a basis of the $\mathbb{R}$-vector space $\mathcal{H}_\mathrm{h}$. We can use linear combinations of ϕ_1 and ϕ_2 to obtain such a basis (cf. Rem. 4.33(b) below):

$$\psi_1, \psi_2 : \mathbb{R} \longrightarrow \mathbb{R}, \quad \psi_1(x) = \frac{e^{ix} + e^{-ix}}{2} = \cos x, \quad \psi_2(x) = \frac{e^{ix} - e^{-ix}}{2i} = \sin x. \tag{4.48}$$

As explained in Rem. 4.33(b) below, as (ϕ_1, ϕ_2, ϕ_3) are a basis of $\mathcal{H}_\mathrm{h}$ over $\mathbb{C}$, (ψ_1, ψ_2, ϕ_3) are a basis of $\mathcal{H}_\mathrm{h}$ over $\mathbb{R}$.

—

By working a bit harder, one can generalize Th. 4.27 to the case where P has zeros of higher multiplicity. We provide this generalization in Th. 4.32 below after recalling the notion of zeros of higher multiplicity in Rem. and Def. 4.29, and after providing two preparatory lemmas.

Remark and Definition 4.29. According to the fundamental theorem of algebra (cf. [Phi13a, Th. 8.32, Cor. 8.33]), for every polynomial $P \in \mathcal{P}_n$ with $\deg P = n$, $n \in \mathbb{N}$, there exists $r \in \mathbb{N}$ with $r \leq n$, $k_1, \dots, k_r \in \mathbb{N}$ with $k_1 + \dots + k_r = n$, and distinct numbers $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ such that

$$P(x) = (x - \lambda_1)^{k_1} \cdots (x - \lambda_r)^{k_r}. \tag{4.49}$$

Clearly, $\lambda_1, \dots, \lambda_r$ are precisely the distinct zeros of P and k_j is referred to as the *multiplicity* of the zero λ_j, $j = 1, \dots, r$.

Lemma 4.30. *Let $I \subseteq \mathbb{R}$ be a nontrivial interval, $\lambda \in \mathbb{K}$, $k \in \mathbb{N}_0$, and $f \in D^k(I)$. Then we have*

$$\underset{x \in I}{\forall} \quad (\partial_x - \lambda)^k \big(f(x)\, e^{\lambda x}\big) = f^{(k)}(x)\, e^{\lambda x}. \tag{4.50}$$

Proof. The proof is carried out by induction. The case $k = 0$ is merely the identity $(\partial_x - \lambda)^0 \big(f(x)\, e^{\lambda x}\big) = f(x)\, e^{\lambda x}$. For the induction step, let $k \geq 1$ and compute, using the product rule,

$$\begin{aligned} (\partial_x - \lambda)^k \big(f(x)\, e^{\lambda x}\big) &\overset{\text{ind. hyp.}}{=} (\partial_x - \lambda)\big(f^{(k-1)}(x)\, e^{\lambda x}\big) \\ &= f^{(k)}(x)\, e^{\lambda x} + f^{(k-1)}(x)\, \lambda\, e^{\lambda x} - \lambda\, f^{(k-1)}(x)\, e^{\lambda x} \\ &= f^{(k)}(x)\, e^{\lambda x}, \end{aligned} \tag{4.51}$$

thereby establishing the case. ∎

Lemma 4.31. *Let $P \in \mathcal{P}$ and $\lambda \in \mathbb{K}$ such that $P(\lambda) \neq 0$. Then, for each $Q \in \mathcal{P}$ with $\deg Q = k$, $k \in \mathbb{N}_0$, it holds that*

$$\forall_{x \in \mathbb{R}} \quad P(\partial_x)\big(Q(x)\, e^{\lambda x}\big) = R(x)\, e^{\lambda x}, \tag{4.52}$$

where $R \in \mathcal{P}$ is still a polynomial of degree k.

Proof. We can rewrite P (cf. [Phi13a, Th. 6.5(a)]) in the form

$$P(x) = \sum_{j=0}^{n} b_j\, (x-\lambda)^j, \quad n \in \mathbb{N}_0, \tag{4.53}$$

where $b_0 = P(\lambda) \neq 0$ and the remaining $b_j \in \mathbb{K}$ can also be calculated from the coefficients of P according to [Phi13a, (6.6)]. We compute

$$P(\partial_x)\big(Q(x)\, e^{\lambda x}\big) \overset{(4.53)}{=} \sum_{j=0}^{n} b_j\, (\partial_x - \lambda)^j \big(Q(x)\, e^{\lambda x}\big) \overset{(4.50)}{=} \sum_{j=0}^{n} b_j\, Q^{(j)}(x)\, e^{\lambda x},$$

i.e. (4.52) holds with $R := \sum_{j=0}^{k} b_j\, Q^{(j)}$ and $b_0 \neq 0$ implies $\deg R = \deg Q = k$. ∎

Theorem 4.32. *If $a_0, \dots, a_{n-1} \in \mathbb{K}$, $n \in \mathbb{N}$, and $P(x) = x^n - \sum_{j=0}^{n-1} a_j x^j$ has the distinct zeros $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ with respective multiplicities $k_1, \dots, k_r \in \mathbb{N}$, then the set*

$$\mathcal{B} := \Big\{ (\phi_{jm} : I \longrightarrow \mathbb{K}) : j \in \{1, \dots, r\},\, m \in \{0, \dots, k_j - 1\} \Big\}, \tag{4.54a}$$

where

$$\forall_{j \in \{1,\dots,r\}} \quad \forall_{m \in \{0,\dots,k_j-1\}} \quad \phi_{jm} : I \longrightarrow \mathbb{K}, \quad \phi_{jm}(x) := x^m\, e^{\lambda_j x}, \tag{4.54b}$$

yields a basis of the homogeneous solution space

$$\mathcal{H}_{\mathrm{h}} = \Big\{ (\phi : I \longrightarrow \mathbb{K}) : P(\partial_x)\phi = 0 \Big\}.$$

Proof. Since $k_1 + \dots + k_r = n$ implies $\#\mathcal{B} = n$ and we know $\dim \mathcal{H}_{\mathrm{h}} = n$, it suffices to show that $\mathcal{B} \subseteq \mathcal{H}_{\mathrm{h}}$ and the elements of $\mathcal{B}$ are linearly independent. Let ϕ_{jm} be as in (4.54b). As λ_j is a zero of multiplicity k_j of P, we can write $P(x) = Q_j(x)(x - \lambda_j)^{k_j}$ with some $Q_j \in \mathcal{P}$. From the computation

$$P(\partial_x)\phi_{jm}(x) = Q_j(\partial_x)(\partial_x - \lambda_j)^{k_j} \big(x^m\, e^{\lambda_j x}\big) \overset{(4.50)}{=} Q_j(\partial_x)\big(\partial_x^{k_j} x^m\big)\, e^{\lambda_j x} \overset{k_j > m}{=} 0,$$

we gather $\mathcal{B} \subseteq \mathcal{H}_{\mathrm{h}}$. Linear independence of the ϕ_{jm} is verified by showing

$$\left(\sum_{j=1}^{r} Q_j(x)\, e^{\lambda_j x} = 0 \quad \wedge \quad \forall_{j=1,\dots,r} \quad Q_j \in \mathcal{P}_{k_j - 1} \right) \quad \Rightarrow \quad \forall_{j=1,\dots,r} \quad Q_j \equiv 0. \tag{4.55}$$

We prove (4.55) by induction on r. Since $e^{\lambda_j x} \neq 0$ for each $x \in \mathbb{R}$, the case $r = 1$ is immediate. For the induction step, let $r \geq 2$. If at least one $Q_j \equiv 0$, then the remaining $Q_j \equiv 0$ as well by the induction hypothesis. It only remains to consider the case that none of the Q_j vanishes identically. In that case, we apply $(\partial_x - \lambda_r)^{k_r}$ to $\sum_{j=1}^{r} Q_j(x)\, e^{\lambda_j x} = 0$, obtaining

$$\sum_{j=1}^{r-1} R_j(x)\, e^{\lambda_j x} = 0 \tag{4.56}$$

with suitable $R_j \in \mathcal{P}$, since Lem. 4.30 yields $(\partial_x - \lambda_r)^{k_r}\big(Q_r(x)\, e^{\lambda_r x}\big) = Q_r^{(k_r)}(x)\, e^{\lambda_r x} = 0$ and, for $j < r$, Lem. 4.31 applies due to $(\lambda_j - \lambda_r)^{k_r} \neq 0$, also providing $\deg R_j = \deg Q_j$. Thus, none of the R_j in (4.56) can vanish identically, violating the induction hypothesis. This finishes the proof of $Q_j \equiv 0$ for each $j = 1, \dots, r$ and the proof of the theorem. ∎

As it can occur in Th. 4.32 that $P \in \mathcal{P}[\mathbb{R}]$, but $\lambda_j \in \mathbb{C} \setminus \mathbb{R}$ for some or all of the zeros λ_j, the question arises of how to obtain a basis of the $\mathbb{R}$-vector space $\mathcal{H}_\mathrm{h}$ from the basis of the $\mathbb{C}$-vector space $\mathcal{H}_\mathrm{h}$ provided by Th. 4.32. The following Rem. 4.33(b) answers this question.

Remark 4.33. **(a)** If $\lambda_1, \lambda_2 \in \mathbb{C}$, then complex conjugation has the properties (cf. [Phi13a, Def. and Rem. 5.5])

$$\overline{\lambda_1 \pm \lambda_2} = \bar\lambda_1 \pm \bar\lambda_2, \quad \overline{\lambda_1 \lambda_2} = \bar\lambda_1\, \bar\lambda_2.$$

In consequence, if $P \in \mathcal{P}[\mathbb{R}]$, then $\overline{P(\lambda)} = P(\bar\lambda)$ for each $\lambda \in \mathbb{C}$. In particular, if $P \in \mathcal{P}[\mathbb{R}]$ and $\lambda \in \mathbb{C} \setminus \mathbb{R}$ is a nonreal zero of P, then $\bar\lambda \neq \lambda$ is also a zero of P.

(b) Consider the situation of Th. 4.32 with $P \in \mathcal{P}[\mathbb{R}]$. Using (a), if $\phi_{jm} : I \longrightarrow \mathbb{C}$, $\phi_{jm}(x) = x^m\, e^{\lambda_j x}$, $\lambda_j \in \mathbb{C} \setminus \mathbb{R}$, occurs in a basis for the $\mathbb{C}$-vector space $\mathcal{H}_\mathrm{h}$ (with $m = 0$ in the special case of Th. 4.27), then $\phi_{\tilde{j}m} : I \longrightarrow \mathbb{C}$, $\phi_{\tilde{j}m}(x) = x^m\, e^{\lambda_{\tilde{j}} x}$, with $\lambda_{\tilde{j}} = \bar\lambda_j$ will occur as well. Noting that, for each $x \in \mathbb{R}$ and each $\lambda \in \mathbb{C}$,

$$e^{\lambda x} = e^{x(\operatorname{Re}\lambda + i \operatorname{Im}\lambda)} = e^{x \operatorname{Re}\lambda}\big(\cos(x \operatorname{Im}\lambda) + i \sin(x \operatorname{Im}\lambda)\big), \tag{4.57a}$$

$$e^{\bar\lambda x} = e^{x(\operatorname{Re}\lambda - i \operatorname{Im}\lambda)} = e^{x \operatorname{Re}\lambda}\big(\cos(x \operatorname{Im}\lambda) - i \sin(x \operatorname{Im}\lambda)\big), \tag{4.57b}$$

$$\frac{1}{2}(e^{\lambda x} + e^{\bar\lambda x}) = e^{x \operatorname{Re}\lambda} \cos(x \operatorname{Im}\lambda), \tag{4.57c}$$

$$\frac{1}{2i}(e^{\lambda x} - e^{\bar\lambda x}) = e^{x \operatorname{Re}\lambda} \sin(x \operatorname{Im}\lambda), \tag{4.57d}$$

one can define

$$\psi_{jm} : I \longrightarrow \mathbb{R}, \quad \psi_{jm}(x) := \frac{1}{2}(\phi_{jm}(x) + \phi_{\tilde{j}m}(x)) = x^m\, e^{x \operatorname{Re}\lambda_j} \cos(x \operatorname{Im}\lambda_j), \tag{4.58a}$$

$$\psi_{\tilde{j}m} : I \longrightarrow \mathbb{R}, \quad \psi_{\tilde{j}m}(x) := \frac{1}{2i}(\phi_{jm}(x) - \phi_{\tilde{j}m}(x)) = x^m\, e^{x \operatorname{Re}\lambda_j} \sin(x \operatorname{Im}\lambda_j). \tag{4.58b}$$

If one replaces each pair $\phi_{jm}, \phi_{\tilde{j}m}$ in the basis for the $\mathbb{C}$-vector space $\mathcal{H}_\mathrm{h}$ with the corresponding pair $\psi_{jm}, \psi_{\tilde{j}m}$, then one obtains a basis for the $\mathbb{R}$-vector space $\mathcal{H}_\mathrm{h}$: This follows from

$$\begin{pmatrix} \psi_{jm} \\ \psi_{\tilde{j}m} \end{pmatrix} = A \begin{pmatrix} \phi_{jm} \\ \phi_{\tilde{j}m} \end{pmatrix} \quad \text{with} \quad A := \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2i} & -\frac{1}{2i} \end{pmatrix}, \quad \det A = -\frac{1}{2i} \neq 0. \tag{4.59}$$

Example 4.34. We consider the fourth-order linear ODE

$$y^{(4)} = -8y'' - 16y, \tag{4.60}$$

which can be written as $P(\partial_x)y = 0$ with

$$P(x) := x^4 + 8x^2 + 16 = (x^2+4)^2 = (x-2i)^2(x+2i)^2, \tag{4.61}$$

i.e. P has the zeros $\lambda_1 = 2i$, $\lambda_2 = -2i$, both with multiplicity 2. Thus, according to Th. 4.32, the four functions

$$\phi_{10}, \phi_{11}, \phi_{20}, \phi_{21} : \mathbb{R} \longrightarrow \mathbb{C},$$
$$\phi_{10}(x) = e^{2ix}, \quad \phi_{11}(x) = x\, e^{2ix}, \quad \phi_{20}(x) = e^{-2ix}, \quad \phi_{21}(x) = x\, e^{-2ix},$$

form a basis of the $\mathbb{C}$-vector space $\mathcal{H}_{\mathrm{h}}$. If we consider (4.60) as an ODE over $\mathbb{R}$, we can use (4.58) to obtain the basis $(\psi_{10}, \psi_{11}, \psi_{20}, \psi_{21})$ of the $\mathbb{R}$-vector space $\mathcal{H}_{\mathrm{h}}$, where

$$\psi_{10}, \psi_{11}, \psi_{20}, \psi_{21} : \mathbb{R} \longrightarrow \mathbb{R},$$
$$\psi_{10}(x) = \cos(2x), \quad \psi_{11}(x) = x \cos(2x), \quad \psi_{20}(x) = \sin(2x), \quad \psi_{21}(x) = x \sin(2x).$$

—

If (4.38) is inhomogeneous, then one can use Th. 4.32 and, if necessary, Rem. 4.33(b), to obtain a basis of the homogeneous solution space $\mathcal{H}_{\mathrm{h}}$, then using the equivalence with systems of first-order linear ODE and variation of constants according to Th. 4.15 to solve (4.38). However, if the function b in (4.38) is such that the following Th. 4.35 applies, then one can avoid using the above strategy to obtain a particular solution ϕ to (4.38) (and, thus, the entire solution space via $\mathcal{H}_{\mathrm{i}} = \phi + \mathcal{H}_{\mathrm{h}}$).

Theorem 4.35. *Let $a_0, \dots, a_{n-1} \in \mathbb{K}$, $n \in \mathbb{N}$, and $P(x) = x^n - \sum_{j=0}^{n-1} a_j x^j$. Consider*

$$P(\partial_x)y = Q(x)e^{\mu x}, \quad Q \in \mathcal{P}, \quad \mu \in \mathbb{K}. \tag{4.62}$$

(a) *(no resonance): If $P(\mu) \neq 0$ and $m := \deg(Q) \in \mathbb{N}_0$, then there exists a polynomial $R \in \mathcal{P}$ such that $\deg(R) = m$ and*

$$\phi : \mathbb{R} \longrightarrow \mathbb{K}, \quad \phi(x) := R(x)\, e^{\mu x}, \tag{4.63}$$

is a solution to (4.62). Moreover, if $Q \equiv 1$, then one can choose $R \equiv 1/P(\mu)$.

(b) *(resonance): If μ is a zero of P with multiplicity $k \in \mathbb{N}$ and $m := \deg(Q) \in \mathbb{N}_0$, then there exists a solution to (4.62) of the following form:*

$$\phi : \mathbb{R} \longrightarrow \mathbb{K}, \quad \phi(x) := R(x)\, e^{\mu x}, \quad R \in \mathcal{P}, \quad R(x) = \sum_{j=k}^{m+k} c_j x^j, \quad c_k, \dots, c_{m+k} \in \mathbb{K}. \tag{4.64}$$

The reason behind the terms no resonance *and* resonance *will be explained in the following Example 4.36.*

Proof. Exercise. ■

Example 4.36. Consider the second-order linear ODE

$$\frac{d^2x}{dt^2} + \omega_0^2\, x = a\, \cos(\omega t), \quad \omega_0, \omega \in \mathbb{R}^+, \quad a \in \mathbb{R} \setminus \{0\}, \tag{4.65}$$

which can be written as $P(\partial_t)x = a\,\cos(\omega t)$ with

$$P(t) := t^2 + \omega_0^2 = (t - i\omega_0)(t + i\omega_0). \tag{4.66}$$

Note that the unknown function is written as x depending on the variable t (instead of y depending on x). This is due to the physical interpretation of (4.65), where x represents the position of a so-called *harmonic oscillator* at time t, having angular frequency ω_0 and being subjected to a periodic external force of angular frequency ω and amplitude a. We can find a particular solution ϕ to (4.65) by applying Th. 4.35 to

$$P(\partial_t)x = a\, e^{i\omega t}. \tag{4.67}$$

We have to distinguish two cases:

(a) Case $\omega \neq \omega_0$: In this case, one says that the oscillator and the external force are not in resonance, which explains the term *no resonance* in Th. 4.35(a). In this case, we can apply Th. 4.35(a) with $\mu := i\omega$ and $Q \equiv a$, yielding $R \equiv a/P(i\omega) = a/(\omega_0^2 - \omega^2)$, i.e.

$$\phi_0 : \mathbb{R} \longrightarrow \mathbb{C}, \qquad \phi_0(t) := R(t)\, e^{\mu t} = \frac{a}{\omega_0^2 - \omega^2}\, e^{i\omega t}, \tag{4.68a}$$

is a solution to (4.67) and

$$\phi : \mathbb{R} \longrightarrow \mathbb{R}, \qquad \phi(t) := \operatorname{Re}\phi_0(t) = \frac{a}{\omega_0^2 - \omega^2}\, \cos(\omega t), \tag{4.68b}$$

is a solution to (4.65).

(b) Case $\omega = \omega_0$: In this case, one says that the oscillator and the external force are in resonance, which explains the term *resonance* in Th. 4.35(b). In this case, we can apply Th. 4.35(b) with $\mu := i\omega$ and $Q \equiv a$, i.e. $m = 0$, $k = 1$, yielding $R(t) = ct$ for some $c \in \mathbb{C}$. To determine c, we plug $x(t) = R(t)\, e^{\mu t}$ into (4.67):

$$\begin{aligned} P(\partial_t)\left(ct\, e^{i\omega t}\right) &= \partial_t(c\, e^{i\omega t} + ci\omega t\, e^{i\omega t}) + \omega_0^2 ct\, e^{i\omega t} \\ &= ci\omega\, e^{i\omega t} + ci\omega\, e^{i\omega t} - c\omega^2 t\, e^{i\omega t} + \omega_0^2 ct\, e^{i\omega t} \\ &= 2ci\omega\, e^{i\omega t} = a\, e^{i\omega t} \quad \Rightarrow \quad c = a/(2i\omega). \end{aligned} \tag{4.69}$$

Thus,

$$\phi_0 : \mathbb{R} \longrightarrow \mathbb{C}, \qquad \phi_0(t) := \frac{a}{2i\omega}\, t\, e^{i\omega t}, \tag{4.70a}$$

is a solution to (4.67) and

$$\phi : \mathbb{R} \longrightarrow \mathbb{R}, \qquad \phi(t) := \operatorname{Re}\phi_0(t) = \frac{a}{2\omega}\, t\, \sin(\omega t), \tag{4.70b}$$

is a solution to (4.65).

4.6.2 Systems of First-Order Linear ODE

Matrix Exponential Function

Definition 4.37. Let $I \subseteq \mathbb{R}$ be a nontrivial interval, $n \in \mathbb{N}$, $A \in \mathcal{M}(n, \mathbb{K})$ and $b : I \longrightarrow \mathbb{K}^n$ be continuous. Then a *linear ODE* with *constant coefficients* is an equation of the form

$$y' = A\,y + b(x), \tag{4.71}$$

i.e. a linear ODE, where the matrix A does not depend on x.

—

Recalling that the ordinary exponential function $\exp_a : \mathbb{R} \longrightarrow \mathbb{C}$, $x \mapsto e^{ax}$, $a \in \mathbb{C}$, is precisely the solution to the initial value problem $y' = a\,y$, $y(0) = 1$, the following definition constitutes a natural generalization:

Definition 4.38. Given $n \in \mathbb{N}$, $A \in \mathcal{M}(n, \mathbb{C})$, define the *matrix exponential function*

$$\exp_A : \mathbb{R} \longrightarrow \mathcal{M}(n, \mathbb{C}), \quad x \mapsto e^{Ax}, \tag{4.72a}$$

to be the solution to the matrix-valued initial value problem

$$Y' = AY, \quad Y(0) = \mathrm{Id}, \tag{4.72b}$$

i.e. as the fundamental matrix solution of $y' = A\,y$ that satisfies $Y(0) = \mathrm{Id}$ (sometimes called the *principal maxtrix solution* of $y' = A\,y$).

—

The previous definition of the matrix exponential function is further justified by the following result:

Theorem 4.39. *For each $A \in \mathcal{M}(n, \mathbb{C})$, $n \in \mathbb{N}$, it holds that*

$$\underset{x \in \mathbb{R}}{\forall} \quad e^{Ax} = \sum_{k=0}^{\infty} \frac{(Ax)^k}{k!} \tag{4.73}$$

in the sense that the partial sums on the right-hand side converge pointwise to e^{Ax} on $\mathbb{R}$, where the convergence is even uniform on every compact interval.

Proof. By the equivalence of all norms on $\mathbb{C}^{n^2} \cong \mathcal{M}(n, \mathbb{C})$, we may choose a convenient norm on $\mathcal{M}(n, \mathbb{C})$. So we let $\|\cdot\|$ denote an arbitrary operator norm on $\mathcal{M}(n, \mathbb{C})$, induced by some norm $\|\cdot\|$ on $\mathbb{C}^n$. We first show that the partial sums $(A_m(x))_{m \in \mathbb{N}}$, $A_m(x) := \sum_{k=0}^{m} \frac{(Ax)^k}{k!}$, in (4.73) form a Cauchy sequence in $\mathcal{M}(n, \mathbb{C})$: For $M, N \in \mathbb{N}$, $N > M$, one estimates, for each $x \in \mathbb{R}$,

$$\|A_N(x) - A_M(x)\| = \left\| \sum_{k=M+1}^{N} \frac{(Ax)^k}{k!} \right\| \overset{\text{(G.10)}}{\leq} \sum_{k=M+1}^{N} \frac{\|A\|^k\,|x|^k}{k!}. \tag{4.74}$$

Since the convergence $\lim_{m\to\infty}\sum_{k=0}^{m}\frac{\|A\|^k\,|x|^k}{k!}=e^{\|A\||x|}$ is pointwise for $x\in\mathbb{R}$ and uniform on every compact interval, (4.74) shows each $(A_m(x))_{m\in\mathbb{N}}$ is a Cauchy sequence that converges to some $\Phi(x)\in\mathcal{M}(n,\mathbb{C})$ (by the completeness of $\mathcal{M}(n,\mathbb{C})$) pointwise for $x\in\mathbb{R}$ and uniform on every compact interval. It remains to show Φ is the solution to (4.72b), i.e.

$$\underset{x\in\mathbb{R}}{\forall}\quad \Phi(x)=\mathrm{Id}+\int_0^x A\Phi(t)\,\mathrm{d}t\,. \tag{4.75}$$

Using the identity

$$A_m(x)=\mathrm{Id}+\sum_{k=1}^{m}\frac{(Ax)^k}{k!}=\mathrm{Id}+A\sum_{k=0}^{m-1}\frac{A^k x^{k+1}}{(k+1)!}=\mathrm{Id}+\int_0^x A\sum_{k=0}^{m-1}\frac{A^k t^k}{k!}\,\mathrm{d}t\,,$$

we estimate, for each $x\in\mathbb{R}$ and each $m\in\mathbb{N}$,

$$\begin{aligned}&\left\|\Phi(x)-\mathrm{Id}-\int_0^x A\Phi(t)\,\mathrm{d}t\right\|\le\left\|\Phi(x)-A_m(x)\right\|+\left\|A_m(x)-\mathrm{Id}-\int_0^x A\Phi(t)\,\mathrm{d}t\right\|\\&=\left\|\Phi(x)-A_m(x)\right\|+\left\|\int_0^x\left(A\sum_{k=0}^{m-1}\frac{A^k t^k}{k!}-A\Phi(t)\right)\mathrm{d}t\right\|\\&\le\left\|\Phi(x)-A_m(x)\right\|+\left|\int_0^x\|A\|\,\left\|A_{m-1}(t)-\Phi(t)\right\|\,\mathrm{d}t\right|\to 0\quad\text{for } m\to\infty,\end{aligned}$$

which proves (4.75) and establishes the case. ■

The matrix exponential function has some properties that are familiar from the case $n=1$ (see Prop. 4.40(a),(b)), but also some properties that are, perhaps, unexpected (see Prop. 4.42(a),(b)).

Proposition 4.40. *Let $A\in\mathcal{M}(n,\mathbb{C})$, $n\in\mathbb{N}$.*

(a) $e^{A(t+s)}=e^{At}e^{As}$ *holds for each $s,t\in\mathbb{R}$.*

(b) $(e^{Ax})^{-1}=e^{A(-x)}=e^{-Ax}$ *holds for each $x\in\mathbb{R}$.*

(c) *For the transpose A^{t}, one has $e^{A^{\mathrm{t}}x}=(e^{Ax})^{\mathrm{t}}$ for each $x\in\mathbb{R}$.*

Proof. (a): Fix $s\in\mathbb{R}$. The function $\Phi_s:\mathbb{R}\longrightarrow\mathcal{M}(n,\mathbb{C})$, $\Phi_s(t):=e^{A(t+s)}$ is a solution to $Y'=AY$ (namely the one for the initial condition $Y(-s)=\mathrm{Id}$). Moreover, the function $\Psi_s:\mathbb{R}\longrightarrow\mathcal{M}(n,\mathbb{C})$, $\Psi_s(t):=e^{At}e^{As}$, is also a solution to $Y'=AY$, since

$$\partial_t\Psi_s(t)=\big(\partial_t e^{At}\big)e^{As}=Ae^{At}e^{As}=A\Psi_s(t).$$

Finally, since $\Psi_s(0)=e^{A0}e^{As}=\mathrm{Id}\,e^{As}=e^{As}=\Phi_s(0)$, the claimed $\Phi_s=\Psi_s$ follows by uniqueness of solutions.

(b) is an easy consequence of (a), since

$$\mathrm{Id}=e^{A0}=e^{A(x-x)}\overset{\text{(a)}}{=}e^{Ax}e^{-Ax}.$$

(c): Clearly, the map $A \mapsto A^{\mathrm{t}}$ is continuous on $\mathcal{M}(n,\mathbb{C})$ (since $\lim_{k\to\infty} A_k = A$ implies $\lim_{k\to\infty} a_{k,\alpha\beta} = a_{\alpha\beta}$ for all components, which implies $\lim_{k\to\infty} A_k^{\mathrm{t}} = A^{\mathrm{t}}$), providing, for each $x \in \mathbb{R}$,

$$e^{A^{\mathrm{t}}x} = \lim_{m\to\infty} \sum_{k=0}^{m} \frac{(A^{\mathrm{t}}x)^k}{k!} = \lim_{m\to\infty} \left(\sum_{k=0}^{m} \frac{(Ax)^k}{k!} \right)^{\mathrm{t}} = \left(\lim_{m\to\infty} \sum_{k=0}^{m} \frac{(Ax)^k}{k!} \right)^{\mathrm{t}} = (e^{Ax})^{\mathrm{t}},$$

completing the proof. ∎

Proposition 4.41. *Let $A \in \mathcal{M}(n,\mathbb{C})$, $n \in \mathbb{N}$. Then*

$$\det e^{A} = e^{\operatorname{tr} A}.$$

Proof. Applying Liouville's formula (4.32) to $\Phi(x) := e^{Ax}$, $x \in \mathbb{R}$, yields

$$\det e^{Ax} = \det e^{A0} \exp\left(\int_0^x \operatorname{tr} A \,\mathrm{d}t \right) = 1 \cdot e^{x \operatorname{tr} A}, \tag{4.76}$$

and setting $x = 1$ in (4.76) proves the proposition. ∎

Proposition 4.42. *Let $A, B \in \mathcal{M}(n,\mathbb{C})$, $n \in \mathbb{N}$.*

(a) *$Be^{Ax} = e^{Ax}B$ holds for each $x \in \mathbb{R}$ if, and only if, $AB = BA$.*

(b) *$e^{(A+B)x} = e^{Ax}e^{Bx}$ holds for each $x \in \mathbb{R}$ if, and only if, $AB = BA$.*

Proof. (a): If $Be^{Ax} = e^{Ax}B$ holds for each $x \in \mathbb{R}$, then differentiation yields $BAe^{Ax} = Ae^{Ax}B$ for each $x \in \mathbb{R}$, and the case $x = 0$ provides $BA\,\mathrm{Id} = A\,\mathrm{Id}\,B$, i.e. $BA = AB$. For the converse, assume $BA = AB$ and define the auxiliary maps

$$\begin{aligned} f_B &: \mathcal{M}(n,\mathbb{C}) \longrightarrow \mathcal{M}(n,\mathbb{C}), \quad f_B(C) := BC, \\ g_B &: \mathcal{M}(n,\mathbb{C}) \longrightarrow \mathcal{M}(n,\mathbb{C}), \quad g_B(C) := CB. \end{aligned}$$

If $\|\cdot\|$ denotes an operator norm, then $\|BC_1 - BC_2\| \le \|B\|\|C_1 - C_2\|$ and $\|C_1B - C_2B\| \le \|B\|\|C_1 - C_2\|$, showing f_B and g_B to be (even Lipschitz) continuous. Thus,

$$\begin{aligned} Be^{Ax} &= f_B(e^{Ax}) = f_B\left(\lim_{m\to\infty} \sum_{k=0}^{m} \frac{(Ax)^k}{k!} \right) = \lim_{m\to\infty} f_B\left(\sum_{k=0}^{m} \frac{(Ax)^k}{k!} \right) \\ &= \lim_{m\to\infty} B \sum_{k=0}^{m} \frac{(Ax)^k}{k!} \overset{AB=BA}{=} \lim_{m\to\infty} \left(\sum_{k=0}^{m} \frac{(Ax)^k}{k!} \right) B = \lim_{m\to\infty} g_B\left(\sum_{k=0}^{m} \frac{(Ax)^k}{k!} \right) \\ &= g_B\left(\lim_{m\to\infty} \sum_{k=0}^{m} \frac{(Ax)^k}{k!} \right) = g_B(e^{Ax}) = e^{Ax}B, \end{aligned}$$

thereby establishing the case.

(b): Exercise (hint: use (a)). ∎

Eigenvalues and Jordan Normal Form

We will see that the solution theory of linear ODE with constant coefficients is related to the eigenvalues of A. We recall the definition of this notion:

Definition 4.43. Let $n \in \mathbb{N}$ and $A \in \mathcal{M}(n, \mathbb{C})$. Then $\lambda \in \mathbb{C}$ is called an *eigenvalue* of A if, and only if, there exists $0 \neq v \in \mathbb{C}^n$ such that

$$Av = \lambda v. \tag{4.77}$$

If (4.77) holds, then $v \neq 0$ is called an *eigenvector* for the eigenvalue λ.

Theorem 4.44. *Let $n \in \mathbb{N}$ and $A \in \mathcal{M}(n, \mathbb{C})$.*

(a) *For each eigenvalue $\lambda \in \mathbb{C}$ of A with eigenvector $v \in \mathbb{C}^n \setminus \{0\}$, the function*

$$\phi : I \longrightarrow \mathbb{C}^n, \quad \phi(x) := e^{\lambda x} v, \tag{4.78}$$

is a solution to the homogeneous version of (4.71)*.*

(b) *If $\{v_1, \dots, v_n\}$ is a basis of eigenvectors for $\mathbb{C}^n$, where v_j is an eigenvector with respect to the eigenvalue $\lambda_j \in \mathbb{C}$ of A for each $j \in \{1, \dots, n\}$, then $\phi_1, \dots, \phi_n$ with*

$$\underset{j \in \{1,\dots,n\}}{\forall} \quad \phi_j : I \longrightarrow \mathbb{C}^n, \quad \phi_j(x) := e^{\lambda_j x} v_j, \tag{4.79}$$

form a fundamental system for (4.71)*.*

Proof. (a): One computes, for each $x \in I$,

$$\phi'(x) = \lambda\, e^{\lambda x} v = e^{\lambda x}\, Av = A\phi(x),$$

proving that ϕ solves the homogeneous version of (4.71).

(b): Without loss of generality, we may consider $I = \mathbb{R}$. We already know from (a) that each ϕ_j is a solution to the homogeneous version of (4.71). Thus, it merely remains to check that $\phi_1, \dots, \phi_n$ are linearly independent. As $\phi_1(0) = v_1, \dots, \phi_n(0) = v_n$ are linearly independent by hypothesis, the linear independence of $\phi_1, \dots, \phi_n$ is provided by Th. 4.11(b). ■

To proceed, we need a few more notions and results from linear algebra:

Theorem 4.45. *Let $n \in \mathbb{N}$ and $A \in \mathcal{M}(n, \mathbb{C})$. Then the following statements* (i) *and* (ii) *are equivalent:*

(i) *There exists a basis $\mathcal{B}$ of eigenvectors for $\mathbb{C}^n$, i.e. there exist $v_1, \dots, v_n \in \mathbb{C}^n$ and $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ such that $\mathcal{B} = \{v_1, \dots, v_n\}$ is a basis of $\mathbb{C}^n$ and $Av_j = \lambda_j v_j$ for each $j = 1, \dots, n$ (note that the v_j must all be distinct, whereas some (or all) of the λ_j may be identical).*

(ii) *There exists an invertible matrix $W \in \mathcal{M}(n, \mathbb{C})$ such that*

$$W^{-1}AW = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \tag{4.80}$$

i.e. A is diagonalizable *(if (4.80) holds, then the columns $v_1, \dots, v_n$ of W must actually be the respective eigenvectors to the eigenvalues $\lambda_1, \dots, \lambda_n$).*

Proof. See, e.g., [Koe03, Th. 8.3.1]. ■

Unfortunately, not every matrix $A \in \mathcal{M}(n, \mathbb{C})$ is diagonalizable. However, every $A \in \mathcal{M}(n, \mathbb{C})$ can at least be transformed into *Jordan normal form*:

Theorem 4.46 (Jordan Normal Form)**.** *Let $n \in \mathbb{N}$ and $A \in \mathcal{M}(n, \mathbb{C})$. There exists an invertible matrix $W \in \mathcal{M}(n, \mathbb{C})$ such that*

$$B := W^{-1}AW \tag{4.81}$$

is in Jordan normal form, *i.e. B has block diagonal form*

$$B = \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_r \end{pmatrix}, \tag{4.82}$$

$1 \le r \le n$, where each block B_j is a so-called Jordan matrix *or* Jordan block*, i.e.*

$$B_j = (\lambda_j) \quad \text{or} \quad B_j = \begin{pmatrix} \lambda_j & 1 & 0 & \dots & 0 \\ & \lambda_j & 1 & & \\ & & \ddots & \ddots & 0 \\ & 0 & & \lambda_j & 1 \\ & & & & \lambda_j \end{pmatrix}, \tag{4.83}$$

where λ_j is an eigenvalue of A.

Proof. See, e.g., [Koe03, Th. 9.5.6] or [Str08, Th. 27.13]. ■

The reason Th. 4.46 regarding the Jordan normal form is useful for solving linear ODE with constant coefficients is the following theorem:

Theorem 4.47. *Let $n \in \mathbb{N}$ and $A, W \in \mathcal{M}(n, \mathbb{C})$, where W is assumed invertible.*

(a) *The following statements* (i) *and* (ii) *are equivalent:*

(i) *$\phi : I \longrightarrow \mathbb{C}^n$ is a solution to $y' = Ay$.*

(ii) *$\psi := W^{-1}\phi : I \longrightarrow \mathbb{C}^n$ is a solution to $y' = W^{-1}AWy$.*

(b) $e^{W^{-1}AWx} = W^{-1}e^{Ax}W$ *for each* $x \in \mathbb{R}$.

Proof. (a): The equivalences

$$\phi' = A\phi \quad \Leftrightarrow \quad W^{-1}\phi' = W^{-1}A\phi \quad \Leftrightarrow \quad \psi' = W^{-1}AW\psi$$

establish the case.

(b): By definition, $x \mapsto e^{W^{-1}AWx}$ is the solution to the initial value problem $Y' = W^{-1}AWY$, $Y(0) = \mathrm{Id}$. Thus, noting $W^{-1}e^{A0}W = \mathrm{Id}$ and

$$(W^{-1}e^{Ax}W)' = W^{-1}Ae^{Ax}W = W^{-1}AWW^{-1}e^{Ax}W$$

shows $x \mapsto W^{-1}e^{Ax}W$ is a solution to the same initial value problem, establishing (b). ■

Remark 4.48. To obtain a fundamental system for (4.71) with $A \in \mathcal{M}(n, \mathbb{C})$, it suffices to obtain a fundamental system for $y' = By$, where $B := W^{-1}AW$ is in Jordan normal form and $W \in \mathcal{M}(n, \mathbb{C})$ is invertible: If $\phi_1, \dots, \phi_n$ are linearly independent solutions to $y' = By$, then $A = WBW^{-1}$, Th. 4.47(a), and W being a linear isomorphism yield that $\psi_1 := W\phi_1, \dots, \psi_n := W\phi_n$ are linearly independent solutions to $y' = Ay$.

Moreover, since B is in block diagonal form with each block being a Jordan matrix according to (4.82) and (4.83), it actually suffices to solve $y' = By$ assuming that

$$B = \lambda \,\mathrm{Id} + N, \tag{4.84}$$

where $\lambda \in \mathbb{C}$ and N is a so-called *canonical nilpotent* matrix, i.e.

$$N = 0 \text{ (zero matrix)} \quad \text{or} \quad N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & 0 \\ & 0 & & 0 & 1 \\ & & & & 0 \end{pmatrix}, \tag{4.85}$$

where the case $N = 0$ is already covered by Th. 4.44. The remaining case is covered by the following Th. 4.49.

Theorem 4.49. *Let* $\lambda \in \mathbb{C}$, $k \in \mathbb{N}$, $k \geq 2$, *and assume* $0 \neq N \in \mathcal{M}(k, \mathbb{C})$ *is a canonical nilpotent matrix according to* (4.85)*. Then*

$$\Phi : \mathbb{R} \longrightarrow \mathcal{M}(k, \mathbb{C}), \quad \Phi(x) := e^{\lambda x} \begin{pmatrix} 1 & x & \frac{x^2}{2} & \dots & \frac{x^{k-2}}{(k-2)!} & \frac{x^{k-1}}{(k-1)!} \\ 0 & 1 & x & \dots & \frac{x^{k-3}}{(k-3)!} & \frac{x^{k-2}}{(k-2)!} \\ 0 & 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 & x \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}, \tag{4.86}$$

is a fundamental matrix solution to

$$Y' = (\lambda \operatorname{Id} + N)Y, \quad Y(0) = \operatorname{Id}, \tag{4.87}$$

i.e.

$$\forall_{x \in \mathbb{R}} \quad \Phi(x) = e^{(\lambda \operatorname{Id} + N)x}; \tag{4.88}$$

in particular, the columns of Φ *provide* k *solutions to* $y' = (\lambda \operatorname{Id} + N)y$ *that are linearly independent.*

Proof. $\Phi(0) = \operatorname{Id}$ is immediate from (4.86). Since $\Phi(x)$ has upper triangular form with all 1's on the diagonal, we obtain $\det \Phi(x) = e^{k\lambda x} \neq 0$ for each $x \in \mathbb{R}$, showing the columns of Φ are linearly independent. Let $\phi_{\alpha\beta} : \mathbb{R} \longrightarrow \mathbb{C}$ denote the αth component function of the βth column of Φ, i.e.

$$\forall_{\alpha,\beta \in \{1,\dots,k\}} \quad \phi_{\alpha\beta} : \mathbb{R} \longrightarrow \mathbb{C}, \quad \phi_{\alpha\beta}(x) := \begin{cases} e^{\lambda x} \frac{x^{\beta-\alpha}}{(\beta-\alpha)!} & \text{for } \alpha \leq \beta, \\ 0 & \text{for } \alpha > \beta. \end{cases}$$

It remains to show that

$$\forall_{\alpha,\beta \in \{1,\dots,k\}} \quad \phi'_{\alpha\beta} = \begin{cases} \lambda \phi_{\alpha\beta} + \phi_{\alpha+1,\beta} & \text{for } \alpha < \beta, \\ \lambda \phi_{\alpha\beta} & \text{for } \alpha = \beta, \\ 0 & \text{for } \alpha > \beta. \end{cases} \tag{4.89}$$

One computes,

$$\forall_{\alpha,\beta \in \{1,\dots,k\}} \quad \phi'_{\alpha\beta}(x) = \begin{cases} \lambda\, e^{\lambda x} \frac{x^{\beta-\alpha}}{(\beta-\alpha)!} + e^{\lambda x} \frac{x^{\beta-(\alpha+1)}}{(\beta-(\alpha+1))!} & \text{for } \alpha < \beta, \\ \lambda\, e^{\lambda x} \frac{x^{\beta-\alpha}}{(\beta-\alpha)!} + 0 & \text{for } \alpha = \beta, \\ 0 & \text{for } \alpha > \beta, \end{cases}$$

i.e. (4.89) holds, completing the proof. ∎

Example 4.50. For a 2-dimensional real system of linear ODE

$$y' = Ay, \quad A \in \mathcal{M}(2, \mathbb{R}), \tag{4.90}$$

there exist precisely the following three possibilities (i) – (iii):

(i) The matrix A is diagonalizable with real eigenvalues $\lambda_1, \lambda_2 \in \mathbb{R}$ ($\lambda_1 = \lambda_2$ is possible), i.e. there is a basis $\{v_1, v_2\}$ of $\mathbb{R}^2$ such that v_j is an eigenvector for λ_j, $j \in \{1, 2\}$. In this case, according to Th. 4.44(b), the two functions

$$\phi_1, \phi_2 : \mathbb{R} \longrightarrow \mathbb{K}^2, \quad \phi_1(x) := e^{\lambda_1 x} v_1, \quad \phi_2(x) := e^{\lambda_2 x} v_2, \tag{4.91}$$

form a fundamental system for (4.90) (over $\mathbb{K}$).

(ii) The matrix A is diagonalizable with two complex conjugate eigenvalues $\lambda_1, \lambda_2 \in \mathbb{C} \setminus \mathbb{R}$, $\lambda_2 = \bar{\lambda}_1$. Analogous to (i), one has a basis $\{v_1, v_2\}$ of $\mathbb{C}^2$ such that v_j is an eigenvector for λ_j, $j \in \{1, 2\}$, and the two functions in (4.91) still form a fundamental system for (4.90), but with $\mathbb{K}$ replaced by $\mathbb{C}$. However, one can still obtain a real-valued fundamental system as follows: We have

$$\lambda_1 = \mu + i\omega, \quad \lambda_2 = \mu - i\omega, \quad \text{where} \quad \mu \in \mathbb{R}, \quad \omega \in \mathbb{R} \setminus \{0\}. \tag{4.92}$$

Thus, if $Av_1 = \lambda_1 v_1$ with $0 \neq v_1 = \alpha + i\beta$, where $\alpha, \beta \in \mathbb{R}^2$, then, letting $v_2 := \bar{v}_1 = \alpha - i\beta$, and taking complex conjugates

$$Av_2 = A\bar{v}_1 = \overline{Av_1} = \overline{\lambda_1 v_1} = \bar{\lambda}_1 \bar{v}_1 = \lambda_2 v_2$$

shows v_2 is an eigenvector with respect to λ_2. Thus, $\phi_2 = \bar{\phi}_1$ and, similar to the approach described in Rem. 4.33(b) above, we can let

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2i} & -\frac{1}{2i} \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix},$$

to obtain a fundamental system $\{\psi_1, \psi_2\}$ for (4.90) over $\mathbb{R}$, where $\psi_1, \psi_2 : \mathbb{R} \longrightarrow \mathbb{R}^2$,

$$\begin{aligned} \psi_1(x) &= \mathrm{Re}(\phi_1(x)) = \mathrm{Re}\left(e^{(\mu+i\omega)x}(\alpha + i\beta)\right) \\ &= \mathrm{Re}\left(e^{\mu x}\big(\cos(\omega x) + i\sin(\omega x)\big)(\alpha + i\beta)\right) \\ &= e^{\mu x}\big(\alpha\cos(\omega x) - \beta\sin(\omega x)\big), &\text{(4.93a)} \\ \psi_2(x) &= \mathrm{Im}(\phi_1(x)) = e^{\mu x}\big(\alpha\sin(\omega x) + \beta\cos(\omega x)\big). &\text{(4.93b)} \end{aligned}$$

(iii) The matrix A has precisely one eigenvalue $\lambda \in \mathbb{R}$ and the corresponding eigenspace is 1-dimensional. Then there is an invertible matrix $W \in \mathcal{M}(2, \mathbb{R})$ such that $B := W^{-1}AW$ is in (nondiagonal) Jordan normal form, i.e.

$$B = W^{-1}AW = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

According to Th. 4.49, the two functions

$$\phi_1, \phi_2 : \mathbb{R} \longrightarrow \mathbb{K}^2, \quad \phi_1(x) := e^{\lambda x}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \phi_2(x) := e^{\lambda x}\begin{pmatrix} x \\ 1 \end{pmatrix}, \tag{4.94}$$

form a fundamental system for $y' = By$ (over $\mathbb{K}$). Thus, according to Th. 4.47, the two functions

$$\psi_1, \psi_2 : \mathbb{R} \longrightarrow \mathbb{K}^2, \quad \psi_1(x) := W\phi_1(x), \quad \psi_2(x) := W\phi_2(x), \tag{4.95}$$

form a fundamental system for (4.90) (over $\mathbb{K}$).

Remark 4.51. One way of finding a fundamental matrix solution for $y' = A\,y$, $A \in \mathcal{M}(n,\mathbb{C})$, is to obtain e^{Ax}, using the following strategy based on Jordan normal forms:

(i) Determine the distinct eigenvalues $\lambda_1, \dots, \lambda_s$, $1 \le s \le n$, of A, which are precisely the zeros of the characteristic polynomial $\chi_A(x) := \det(A - x\,\mathrm{Id})$ (the multiplicity of the zero λ_j is called its *algebraic multiplicity*, the dimension of the eigenspace $\ker(A - \lambda_j\,\mathrm{Id})$ its *geometric multiplicity*).

(ii) Determine the Jordan normal form B of A and W such that $B = W^{-1}AW$. In general, this means computing the (finitely many distinct) powers $(A-\lambda_j\,\mathrm{Id})^k$ and (suitable bases of) $\ker(A-\lambda_j\,\mathrm{Id})^k$ (in general, this is a somewhat involved procedure and it is referred to [Mar04, Sections 4.2,4.3] and [Str08, Sec. 27] for details – the lecture notes do not provide further details here, as they rather recommend using the Putzer algorithm as described below instead).

(iii) For each Jordan block B_j (as in (4.83)) of B compute $e^{B_j x}$ as in (4.86).

As step (ii) above tends to be complicated in practise, it is usually easier to obtain e^{Ax} using the Putzer algorithm described next.

Putzer Algorithm

The Putzer algorithm due to [Put66] is a procedure for computing e^{Ax} that avoids the difficulty of determining the Jordan normal form of A, and, thus, is often more efficient to employ in practise than the procedure described in Rem. 4.51 above. The Putzer algorithm is provided by the following theorem:

Theorem 4.52. *Let $A \in \mathcal{M}(n,\mathbb{C})$, $n \in \mathbb{N}$. If $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ are precisely the eigenvalues of A (not necessarily distinct, each eigenvalue occurring possibly repeatedly according to its multiplicity). Then*

$$\underset{x\in\mathbb{R}}{\forall} \quad e^{Ax} = \sum_{k=0}^{n-1} p_{k+1}(x)\,M_k, \tag{4.96}$$

where the functions $p_1, \dots, p_n : \mathbb{R} \longrightarrow \mathbb{C}$ and matrices $M_0, \dots, M_{n-1} \in \mathcal{M}(n,\mathbb{C})$ are defined recursively by

$$p_1' = \lambda_1\,p_1, \qquad p_1(0) = 1, \tag{4.97a}$$
$$p_k' = \lambda_k\,p_k + p_{k-1}, \qquad p_k(0) = 0 \qquad \textit{for} \quad k = 2, \dots, n \tag{4.97b}$$

(i.e. each p_k is a solution to a (typically nonhomogeneous) 1-dimensional first-order linear ODE that can be solved using (2.2)) and

$$M_0 := \mathrm{Id}, \tag{4.98a}$$
$$M_k = M_{k-1}\,(A - \lambda_k\,\mathrm{Id}) \quad \textit{for} \quad k = 1, \dots, n-1. \tag{4.98b}$$

Proof. Note that (4.98) can be extended to $k = n$, yielding

$$M_n = \prod_{k=1}^{n} (A - \lambda_k \operatorname{Id}) = \chi_A(A) = 0,$$

since each matrix annihilates its characteristic polynomial according to the Cayley-Hamilton theorem (cf. [Koe03, Th. 8.4.6] or [Str08, Th. 26.6]). Also note

$$\forall_{k=0,\dots,n-1} \quad AM_k = M_k(A - \lambda_{k+1} \operatorname{Id}) + \lambda_{k+1} M_k = M_{k+1} + \lambda_{k+1} M_k. \tag{4.99}$$

We have to show that $x \mapsto \Phi(x) := \sum_{k=0}^{n-1} p_{k+1}(x)\, M_k$ solves the initial value problem $Y' = AY$, $Y(0) = \operatorname{Id}$. The initial condition is satisfied, as $\Phi(0) = p_1(0)\, M_0 = \operatorname{Id}$, and the ODE is satisfied, as, for each $x \in \mathbb{R}$,

$$\begin{aligned}
\Phi'(x) - A\Phi(x) \quad &= \quad \sum_{k=0}^{n-1} p'_{k+1}(x)\, M_k - A \sum_{k=0}^{n-1} p_{k+1}(x)\, M_k \\
&\overset{(4.97),\,(4.99)}{=} \quad \lambda_1 p_1(x)\, M_0 + \sum_{k=1}^{n-1} \big(\lambda_{k+1} p_{k+1}(x) + p_k(x)\big) M_k \\
&\qquad - \sum_{k=0}^{n-1} p_{k+1}(x) \big(M_{k+1} + \lambda_{k+1} M_k\big) \\
&= \quad -p_n(x)\, M_n = 0,
\end{aligned}$$

completing the proof. ■

5 Stability

5.1 Qualitative Theory, Phase Portraits

In the *qualitative theory* of ODE, which can be seen as part of the field of *dynamical systems*, the idea is to understand the set of solutions to an ODE (or to a class of ODE), if possible, without making use of explicit solution formulas, which, in most situations, are not available anyway. Examples of qualitative questions are if, and under which conditions, solutions to an ODE are constant, periodic, are unbounded, approach some limit (more generally, the solutions' asymptotic behavior), etc. One often thinks of the solutions as depending on a time-like variable, and then qualitative theory typically means disregarding the speed of change, but rather focusing on the shape/geometry of the solution's image.

The topic of *stability* takes continuity in intial conditions further and investigates the behavior of solutions that are, at least initially, close to some given solution. Under which conditions do nearby solutions approach each other or diverge away from each other, show the same or different asymptotic behavior etc.

Even though the abovedescribed considerations are not limited to this situation, a natural starting point is to consider first-order ODE where the right-hand side does not depend on x. In the following, we will mostly be concerned with this type of ODE, which has a special name:

Definition 5.1. If $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{K}^n$, then the n-dimensional first-order ODE

$$y' = f(y) \tag{5.1}$$

is called *autonomous* and Ω is called the *phase space*.

Remark 5.2. In fact, nonautonomous ODE are not really more general than autonomous ODE, due to the, perhaps, surprising Th. J.1 of the Appendix, which states that every nonautonomous ODE is equivalent to an autonomous ODE. However, this fact is of little practical relevance, since the autonomous ODE arising via Th. J.1 from nonautonomous ODE can never have bounded solutions on unbounded intervals, whereas the theory of autonomous ODE is most powerful and useful for ODE that admit bounded solutions on unbounded intervals (such as constant or periodic solutions, or solutions approaching constant or periodic functions).

Lemma 5.3. *If, in the context of Def.* 5.1, $\phi : I \longrightarrow \mathbb{K}^n$ *is a solution to* (5.1)*, defined on the interval* $I \subseteq \mathbb{R}$*, then*

$$\underset{\xi\in\mathbb{R}}{\forall} \quad \phi_\xi : I - \xi \longrightarrow \mathbb{K}^n, \quad \phi_\xi(x) := \phi(x+\xi), \quad \textit{where} \quad I - \xi := \{x - \xi \in \mathbb{R} : x \in I\}, \tag{5.2}$$

is another solution to (5.1)*. In consequence, if* ϕ *is a maximal solution, then so is* ϕ_ξ*.*

Proof. Clearly, $I - \xi$ is an interval. Note $x \in I - \xi \;\Rightarrow\; x + \xi \in I$ and, since ϕ is a solution to (5.1), it is $\phi(I) \subseteq \Omega$, implying $\phi_\xi(I - \xi) \subseteq \Omega$. Finally,

$$\underset{x\in I-\xi}{\forall} \quad \phi_\xi'(x) = \phi'(x+\xi) = f\big(\phi(x+\xi)\big) = f\big(\phi_\xi(x)\big),$$

completing the proof that ϕ_ξ is a solution. Since each extension of ϕ yields an extension of ϕ_ξ and vice versa, ϕ is a maximal solution if, and only if, ϕ_ξ is a maximal solution. ■

Lemma 5.4. *If* $\Omega \subseteq \mathbb{K}^n$*,* $n \in \mathbb{N}$*, and* $f : \Omega \longrightarrow \mathbb{K}^n$ *is such that* (5.1) *admits unique maximal solutions (*f *being locally Lipschitz on* Ω *open is sufficient, but not necessary, cf. Def.* 3.32*), then the global solution* $Y : D_f \longrightarrow \mathbb{K}^n$ *of* (5.1) *satisfies*

(a) $Y(x,\xi,\eta) = Y(x-\xi,0,\eta)$ *for each* $(x,\xi,\eta) \in D_f$*.*

(b) $Y\big(x,0,Y(\tilde{x},0,\eta)\big) = Y(x+\tilde{x},0,\eta)$ *for each* $(x,\tilde{x},\eta) \in \mathbb{R}\times\mathbb{R}\times\mathbb{K}^n$ *such that* $(\tilde{x},0,\eta) \in D_f$ *and* $\big(x,0,Y(\tilde{x},0,\eta)\big) \in D_f$*.*

Proof. (a): If $\psi : I_{\xi,\eta} \longrightarrow \mathbb{K}^n$ and $\phi : I_{0,\eta} \longrightarrow \mathbb{K}^n$ denote the maximal solutions to the initial data $y(\xi) = \eta$ and $y(0) = \eta$, respectively, then (a) claims, using the notation from Lem. 5.3, $\psi = \phi_{-\xi}$. As a consequence of Lem. 5.3, $\phi_{-\xi} : I_{0,\eta} + \xi \longrightarrow \mathbb{K}^n$, is some maximal solution to (5.1) and, since $\phi_{-\xi}(\xi) = \phi(0) = \eta = \psi(\xi)$, the assumed uniqueness yields the claimed $\psi = \phi_{-\xi}$, in particular, $I_{\xi,\eta} = I_{0,\eta} + \xi$.

(b): Let $\tilde{\eta} := Y(\tilde{x}, 0, \eta)$. If $\psi : I_{0,\tilde{\eta}} \longrightarrow \mathbb{K}^n$ and $\phi : I_{0,\eta} \longrightarrow \mathbb{K}^n$ denote the maximal solutions to the initial data $y(0) = \tilde{\eta}$ and $y(0) = \eta$, respectively, then (b) claims $\psi = \phi_{\tilde{x}}$. As a consequence of Lem. 5.3, $\phi_{\tilde{x}} : I_{0,\eta} - \tilde{x} \longrightarrow \mathbb{K}^n$, is some maximal solution to (5.1) and, since $\phi_{\tilde{x}}(0) = \phi(\tilde{x}) = \tilde{\eta} = \psi(0)$, the assumed uniqueness yields the claimed $\psi = \phi_{\tilde{x}}$, in particular, $I_{0,\tilde{\eta}} = I_{0,\eta} - \tilde{x}$. ∎

Definition 5.5. Let $I \subseteq \mathbb{R}$ be an interval and $\phi : I \longrightarrow S$ (in principle, S can be arbitrary).

(a) The image of I under ϕ, i.e.

$$\mathcal{O}(\phi) := \phi(I) = \{\phi(x) : x \in I\} \subseteq S \tag{5.3}$$

is often referred to as the *orbit* of ϕ in the present context of qualitative ODE theory.

(b) $\phi : \mathbb{R} \longrightarrow S$ (note $I = \mathbb{R}$) is called *periodic* if, and only if, there exists a smallest $\omega > 0$ (called the *period* of ϕ) such that

$$\underset{x \in \mathbb{R}}{\forall} \quad \phi(x + \omega) = \phi(x). \tag{5.4}$$

The requirement $\omega > 0$ means constant functions are not periodic in the sense of this definition.

Lemma 5.6. *Let $\phi : \mathbb{R} \longrightarrow \mathbb{K}^n$, $n \in \mathbb{N}$.*

(a) *If ϕ is continuous and* (5.4) *holds for some $\omega > 0$, then ϕ is either constant or periodic in the sense of Def.* 5.5(b).

(b) (a) *is false without the assumption of ϕ being continuous.*

Proof. Exercise. ∎

Definition 5.7. Let $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{K}^n$. In the context of the autonomous ODE (5.1), the zeros of f are called the *fixed points* of the ODE (5.1) (cf. Lem. 5.8 below). One then sometimes uses the notation

$$\mathcal{F} := \mathcal{F}_f := \{\eta \in \Omega : F(\eta) = 0\} \tag{5.5}$$

for the set of fixed points.

Lemma 5.8. *Let $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, $f : \Omega \longrightarrow \mathbb{K}^n$, $\eta \in \Omega$. Then the following statements are equivalent:*

(i) *$f(\eta) = 0$, i.e. η is a fixed point of* (5.1).

(ii) *$\phi : \mathbb{R} \longrightarrow \mathbb{K}^n$, $\phi \equiv \eta$, is a solution to* (5.1).

Proof. If $f(\eta) = 0$ and $\phi \equiv \eta$, then $\phi'(x) = 0 = f(\phi(x))$ for each $x \in \mathbb{R}$, i.e. (i) implies (ii). Conversely, if $\phi \equiv \eta$ is a solution to (5.1), then $f(\eta) = f(\phi(x)) = \phi'(x) = 0$, i.e. (ii) implies (i). ■

Proposition 5.9. *If $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{K}^n$ is such that* (5.1) *admits unique maximal solutions (f being locally Lipschitz on Ω open is sufficient), then, for maximal solutions $\phi_1 : I_1 \longrightarrow \mathbb{K}^n$, $\phi_2 : I_2 \longrightarrow \mathbb{K}^n$ to* (5.1)*, defined on open intervals I_1, I_2, respectively, precisely one of the following two statements* (i) *and* (ii) *is true:*

(i) *$\mathcal{O}(\phi_1) \cap \mathcal{O}(\phi_2) = \emptyset$, i.e. the solutions have disjoint orbits.*

(ii) *There exists $\xi \in \mathbb{R}$ such that*

$$I_2 = I_1 - \xi \quad \text{and} \quad \underset{x \in I_2}{\forall} \quad \phi_2(x) = \phi_1(x + \xi). \tag{5.6}$$

In particular, it follows in this case that $\mathcal{O}(\phi_1) = \mathcal{O}(\phi_2)$, i.e. the solutions have the same orbit.

Proof. Suppose (i) does not hold. Then there are $x_1 \in I_1$ and $x_2 \in I_2$ such that $\phi_1(x_1) = \phi_2(x_2)$. Define $\xi := x_1 - x_2$ and consider

$$\phi : I_1 - \xi \longrightarrow \mathbb{K}^n, \quad \phi(x) := \phi_1(x + \xi). \tag{5.7}$$

Then ϕ is a maximal solution of (5.1) by Lem. 5.3 and $\phi(x_2) = \phi_1(x_1) = \phi_2(x_2)$. By uniqueness of maximal solutions, we obtain $\phi = \phi_2$, in particular, $I_2 = I_1 - \xi$, proving (5.6). Clearly, (5.6) implies $\mathcal{O}(\phi_1) = \mathcal{O}(\phi_2)$. ■

Proposition 5.10. *If $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{K}^n$ is such that* (5.1) *admits unique maximal solutions (f being locally Lipschitz on Ω open is sufficient), then, for each maximal solution $\phi : I \longrightarrow \mathbb{K}^n$ to* (5.1)*, defined on the open interval I, precisely one of the following three statements is true:*

(i) *ϕ is injective.*

(ii) *$I = \mathbb{R}$ and ϕ is periodic.*

(iii) *$I = \mathbb{R}$ and ϕ is constant (in this case $\eta := \phi(0)$ is a fixed point of* (5.1)*).*

Proof. Clearly, (i) – (iii) are mutually exclusive. Suppose (i) does not hold. Then there exist $x_1, x_2 \in I$, $x_1 < x_2$, such that $\phi(x_1) = \phi(x_2)$. Set $\omega := x_2 - x_1$. According to Lem. 5.3, $\psi : I - \omega \longrightarrow \mathbb{K}^n$, $\psi(x) := \phi(x + \omega)$, must also be a maximal solution to (5.1). Since $\psi(x_1) = \phi(x_1 + \omega) = \phi(x_2) = \phi(x_1)$, uniqueness implies $\psi = \phi$ and $I = I - \omega$. As $\omega > 0$, this means $I = \mathbb{R}$ and the validity of (5.4). As ϕ is also continuous, by Lem. 5.6(a), either (ii) or (iii) must hold. ■

Corollary 5.11. *If $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{K}^n$ is such that* (5.1) *admits unique maximal solutions (f being locally Lipschitz on Ω open is sufficient), then the orbits of maximal solutions to* (5.1) *partition the phase space Ω into disjoint sets. Moreover, every point $\eta \in \Omega$ is either a fixed point, or it belongs to some periodic orbit, or it belongs to the orbit of some injective solution.*

Proof. The corollary merely summarizes Prop. 5.9 and Prop. 5.10. ∎

Definition 5.12. In the situation of Cor. 5.11, a *phase portrait* for (5.1) is a sketch showing representative orbits. Thus, the sketch shows subsets of the phase space Ω, including fixed points (if any) and representative periodic solutions (if any). Usually, one also uses arrows to indicate the direction in which each drawn orbit is traced as the variable x increases.

Example 5.13. Even though it is a main goal of qualitative theory to obtain phase portraits without the need of explicit solution formulas, and we will study techniques for accomplishing this below, we will make use of explicit solution formulas for our first two examples of phase portraits.

(a) Consider the autonomous linear ODE

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix}. \tag{5.8}$$

Here we have $\Omega = \mathbb{R}^2$ and $f : \Omega \longrightarrow \Omega$, $f(y_1, y_2) = (-y_2, y_1)$. The only fixed point is $(0,0)$. Clearly, for each $r > 0$, $\phi : \mathbb{R} \longrightarrow \mathbb{R}^2$, $\phi(x) := (r\cos x, r\sin x)$ is a solution to (5.8) and its orbit is the circle with radius r around the origin. Since every point of Ω belongs to such a circle, every orbit is either the origin or a circle around the origin. Thus, the phase portrait consists of such circles plus the origin and arrows that indicate the circles are traversed counterclockwise.

(b) As compared to the previous one, the phase portrait of the autonomous linear ODE

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} y_2 \\ y_1 \end{pmatrix} \tag{5.9}$$

is more complicated: While $(0,0)$ is still the only fixed point, for each $r > 0$, all the following functions $\phi_1, \phi_2, \phi_3, \phi_4 : \mathbb{R} \longrightarrow \mathbb{R}^2$ are solutions:

$$\phi_1(x) := (r\cosh x, r\sinh x), \tag{5.10a}$$
$$\phi_2(x) := (-r\cosh x, -r\sinh x), \tag{5.10b}$$
$$\phi_3(x) := (r\sinh x, r\cosh x), \tag{5.10c}$$
$$\phi_4(x) := (-r\sinh x, -r\cosh x), \tag{5.10d}$$

each type describing a hyperbolic orbit in some section of the plane $\mathbb{R}^2$. These sections are separated by rays, forming the orbits of the solutions $\phi_5, \phi_6, \phi_7, \phi_8 : \mathbb{R} \longrightarrow \mathbb{R}^2$:

$$\phi_5(x) := (e^x, e^x), \tag{5.10e}$$
$$\phi_6(x) := (-e^x, -e^x), \tag{5.10f}$$
$$\phi_7(x) := (e^{-x}, -e^{-x}), \tag{5.10g}$$
$$\phi_8(x) := (-e^{-x}, e^{-x}). \tag{5.10h}$$

The two rays on $\{(y_1, y_1) : y_1 \neq 0\}$ move away from the origin, whereas the two rays on $\{(y_1, -y_1) : y_1 \neq 0\}$ move toward the origin. The hyperbolic orbits asymptotically approach the ray orbits and are traversed such that the flow direction agrees between approaching orbits.

—

The next results will be useful to obtain new phase portraits from previously known phase portraits in certain situations.

Proposition 5.14. *Let $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, let $I \subseteq \mathbb{R}$ be some nontrivial interval, let $f : \Omega \longrightarrow \mathbb{K}^n$, and let $\phi : I \longrightarrow \mathbb{K}^n$ be a solution to*

$$y' = \gamma(x)\, f(y), \tag{5.11}$$

where $\gamma : I \longrightarrow \mathbb{R}$ is continuous. If $\phi'(x) \neq 0$ for each $x \in I$ (if one thinks of x as time, then one can think of ϕ' as the velocity of ϕ), then there exists a continuously differentiable bijective map $\lambda : J \longrightarrow I$, defined on some nontrivial interval J, such that $(\phi \circ \lambda) : J \longrightarrow \mathbb{K}^n$ is a solution to $y' = f(y)$.

Proof. Since $\phi'(x) \neq 0$ for each $x \in I$, one has $\gamma(x) \neq 0$ for each $x \in I$. As γ is also continuous, it must be either always negative or always positive. In consequence, fixing $x_0 \in I$, $\Gamma : I \longrightarrow \mathbb{R}$, $\Gamma(x) := \int_{x_0}^{x} \gamma(t)\,\mathrm{d}t$, is continuous and either strictly increasing or strictly decreasing. In particular, Γ is injective, $J := \Gamma(I)$ is an interval, and $\Gamma : I \longrightarrow J$ is bijective. The desired function λ is $\lambda := \Gamma^{-1} : J \longrightarrow I$. Indeed, according to [Phi13a, Th. 9.8], λ is differentiable and its derivative is the continuous function

$$\lambda' : J \longrightarrow \mathbb{R}, \quad \lambda'(x) = \frac{1}{\gamma\big(\lambda(x)\big)},$$

implying, for each $x \in J$,

$$(\phi \circ \lambda)'(x) = \phi'\big(\lambda(x)\big)\,\lambda'(x) = \gamma\big(\lambda(x)\big)\, f\big(\phi(\lambda(x))\big)\,\frac{1}{\gamma\big(\lambda(x)\big)} = f\big(\phi(\lambda(x))\big),$$

showing $\phi \circ \lambda$ is a solution to $y' = f(y)$ as required. ■

Proposition 5.15. *Let $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{K}^n$. Moreover, consider a continuous function $h : \Omega \longrightarrow \mathbb{R}$ with the property that either $h > 0$ everywhere on Ω or $h < 0$ everywhere on Ω.*

(a) *If f has no zeros (i.e. $\mathcal{F} = \emptyset$), then the ODE*

$$y' = f(y), \tag{5.12a}$$
$$y' = h(y)\, f(y) \tag{5.12b}$$

have precisely the same orbits, i.e. every orbit of a solution to (5.12a) *is an orbit of a solution to* (5.12b) *and vice versa.*

(b) *If f and h are such that the ODE* (5.12) *admit unique maximal solutions, then the ODE* (5.12) *have precisely the same orbits (even if $\mathcal{F} \neq \emptyset$).*

Proof. (a): If $\phi : I \longrightarrow \mathbb{K}^n$ is a solution to (5.12b), then $\gamma := h \circ \phi$ is well-defined and continuous. Since $\mathcal{F} = \emptyset$ implies $\phi' \neq 0$, we can apply Prop. 5.14 to obtain the existence of a bijective $\lambda_1 : J_1 \longrightarrow I$ such that $\phi \circ \lambda_1$ is a solution to (5.12a). Thus, $\mathcal{O}(\phi) = \mathcal{O}(\phi \circ \lambda_1)$. Conversely, if $\psi : I \longrightarrow \mathbb{K}^n$ is a solution to (5.12a), i.e. to $y' = \frac{h(y)}{h(y)} f(y)$, then $\gamma := 1/(h \circ \psi)$ is well-defined and continuous. Since $\mathcal{F} = \emptyset$ implies $\psi' \neq 0$, we can apply Prop. 5.14 to obtain the existence of a bijective $\lambda_2 : J_2 \longrightarrow I$ such that $\psi \circ \lambda_2$ is a solution to (5.12b). Thus, $\mathcal{O}(\psi) = \mathcal{O}(\psi \circ \lambda_2)$.

(b): We are now in the situation of Prop. 5.10 and Cor. 5.11, and from (a) we know every nonconstant orbit of (5.12a) is a nonconstant orbit of (5.12b) and vice versa. However, since $h > 0$ or $h < 0$, both ODE in (5.12) have precisely the same constant solutions, concluding the proof. ■

Remark 5.16. We apply Prop. 5.15 to phase portraits (in particular, assume unique maximal solutions). Prop. 5.15 says that overall multiplication with a continuous positive function h does not change the phase portrait at all. Moreover, Prop. 5.15 also states that overall multiplication with a continuous negative function h does not change the partition of Ω into solution orbits. However, after multiplication with a negative h, the orbits are clearly traversed in the opposite direction, i.e., for negative h, the arrows in the phase portrait have to be reversed. For a general continuous h, this implies the phase portrait remains the same in each region of Ω, where $h > 0$; it remains the same, except for the arrows reversed, in each region of Ω, where $h < 0$; and the zeros of h add additional fixed points, cutting some of the previous orbits. We summarize how to obtain the phase portrait of (5.12b) from that of (5.12a):

(1) Start with the phase portrait of (5.12a).

(2) Add the zeros of h as additional fixed points (if any). Previous orbits are cut, where fixed points are added.

(3) Reverse the arrows where $h < 0$.

Example 5.17. (a) Consider the ODE

$$\begin{aligned} y_1' &= -y_2 \left((y_1 - 1)^2 + y_2^2\right), \\ y_2' &= y_1 \left((y_1 - 1)^2 + y_2^2\right), \end{aligned} \tag{5.13}$$

which comes from multiplying the right-hand side of (5.8) by $h(y) = (y_1 - 1)^2 + y_2^2$. The phase portrait is the same as the one for (5.8), except for the added fixed point at $\{(1, 0)\}$.

(b) Consider the ODE

$$\begin{aligned} y_1' &= -y_1\, y_2 + y_2^2, \\ y_2' &= -y_1\, y_2 + y_1^2, \end{aligned} \tag{5.14}$$

which comes from multiplying the right-hand side of (5.8) by $h(y) = y_1 - y_2$. The phase portrait is obtained from that of (5.8), where additional fixed points are on the line with $y_1 = y_2$. This line cuts each previously circular orbit into two segments. The arrows have to be reversed for $y_2 > y_1$, that means above the $y_1 = y_2$ line.

Definition 5.18. Let $\Omega \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{R}^n$. A function $E : \Omega \longrightarrow \mathbb{R}$ is called an *integral* for the autonomous ODE (5.1), i.e. for $y' = f(y)$, if, and only if, $E \circ \phi$ is constant for every solution ϕ of (5.1).

Lemma 5.19. *Let $\Omega \subseteq \mathbb{R}^n$ be open, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{R}^n$ such that each initial value problem for* (5.1) *has at least one solution (f continuous is sufficient by Th.* 3.8*). Then a differentiable function $E : \Omega \longrightarrow \mathbb{R}$ is an integral for* (5.1) *if, and only if,*

$$\underset{y \in \Omega}{\forall} \quad (\nabla E)(y) \bullet f(y) = \sum_{j=1}^{n} \partial_j E(y)\, f_j(y) = 0. \tag{5.15}$$

Proof. Let $\phi : I \longrightarrow \mathbb{R}^n$ be a solution to $y' = f(y)$. Then, by the chain rule,

$$\underset{x \in I}{\forall} \quad (E \circ \phi)'(x) = (\nabla E)(\phi(x)) \bullet \phi'(x) = (\nabla E)(\phi(x)) \bullet f(\phi(x)). \tag{5.16}$$

The differentiable function $E \circ \phi : I \longrightarrow \mathbb{R}$ is constant on the interval I if, and only if, $(E \circ \phi)' \equiv 0$. Thus, by (5.16), $E \circ \phi$ being constant for every solution ϕ is equivalent to $\big((\nabla E) \bullet f\big)(y) = 0$ for each $y \in \Omega$ such that at least one solution passes through y. ■

Example J.2 of the Appendix, pointed out by Anton Sporrer, shows the hypothesis of Lem. 5.19, that each initial value problem for (5.1) has at least one solution, can not be omitted. The following Prop. 5.20 makes use of integrals and applies to phase portraits of 2-dimensional real ODE:

Proposition 5.20. *Let $\Omega \subseteq \mathbb{R}^2$ be open, and let $f : \Omega \longrightarrow \mathbb{R}^2$ be continuous and such that* (5.1) *admits unique maximal solutions (f being locally Lipschitz is sufficient). Assume $E : \Omega \longrightarrow \mathbb{R}$ to be a continuously differentiable integral for* (5.1)*, i.e. for $y' = f(y)$, satisfying $\nabla E(y) \neq 0$ for each $y \in \Omega$. Then the following statements hold for each maximal solution $\phi : I \longrightarrow \mathbb{R}^2$ of* (5.1) *($I \subseteq \mathbb{R}$ some open interval):*

(a) *If $(x_m)_{m \in \mathbb{N}}$ is a sequence in I such that $\lim_{m\to\infty} \phi(x_m) = \eta \in \Omega$, then $\eta \in \mathcal{F}$ (i.e. η is a fixed point) or $\eta \in \mathcal{O}(\phi)$ (i.e. there exists $\xi \in I$ with $\phi(\xi) = \eta$).*

(b) *Let $C \in \mathbb{R}$ be such that $E \circ \phi \equiv C$ (such a C exists, as E is an integral). If $E^{-1}\{C\} = \{y \in \Omega : E(y) = C\}$ is compact and $E^{-1}\{C\} \cap \mathcal{F} = \emptyset$, then ϕ is periodic.*

Proof. Throughout the proof let C be as in (b), i.e. $E \circ \phi \equiv C$.

(a): The continuity of E yields $E(\eta) = \lim_{m\to\infty} E(\phi(x_m)) = C$. Moreover, by hypothesis, $(\epsilon_1, \epsilon_2) := \nabla E(\eta) \neq (0, 0)$. We proceed with the proof for $\epsilon_2 \neq 0$ – if $\epsilon_2 = 0$ and $\epsilon_1 \neq 0$, then the roles of the indices 1, 2 have to be switched in the following. We apply the implicit function theorem [Phi13b, Th. C.9] to the function $\tilde{f}: \Omega \longrightarrow \mathbb{R}$, $\tilde{f}(y) := E(y) - C$ at its zero $\eta = (\eta_1, \eta_2)$. By [Phi13b, Th. C.9], there exist $\epsilon, \delta > 0$ and a continuously differentiable map $g: I_g \longrightarrow \mathbb{R}$, $I_g :=]\eta_1 - \delta, \eta_1 + \delta[$, such that $g(\eta_1) = \eta_2$,

$$\forall_{s \in I_g} \quad E\big(s, g(s)\big) = C, \tag{5.17a}$$

and, having fixed some arbitrary norm $\|\cdot\|$ on $\mathbb{R}^2$,

$$\forall_{y \in \Omega} \quad \bigg(\Big(\|y - \eta\| < \epsilon \ \wedge \ E(y) = C \Big) \quad \Rightarrow \quad \exists_{s \in I_g} \quad y = \big(s, g(s)\big) \bigg). \tag{5.17b}$$

We now assume $\eta \notin \mathcal{F}$ and show $\eta \in \mathcal{O}(\phi)$. If $\eta \notin \mathcal{F}$, then $f(\eta) \neq 0$ and the continuity of f and g imply there is $\tilde{\delta} > 0$, $\tilde{\delta} \leq \delta$, such that, for each $s \in \tilde{I} :=]\eta_1 - \tilde{\delta}, \eta_1 + \tilde{\delta}[$, $f(s, g(s)) \neq 0$. Define the auxiliary function $\varphi: \tilde{I} \longrightarrow \Omega$, $\varphi(s) = (s, g(s))$. Since $E \circ \varphi \equiv C$, we can employ the chain rule to conclude

$$\forall_{s \in \tilde{I}} \quad 0 = (E \circ \varphi)'(s) = (\nabla E)(\varphi(s)) \bullet \varphi'(s), \tag{5.18}$$

i.e. the two-dimensional vectors $(\nabla E)(\varphi(s))$ and $\varphi'(s)$ are orthogonal with respect to the Euclidean scalar product. As E is an integral, using (5.15), $f(\varphi(s))$ is another vector orthogonal to $(\nabla E)(\varphi(s))$ and, since all vectors in $\mathbb{R}^2$ orthogonal to $(\nabla E)(\varphi(s))$ form a 1-dimensional subspace of $\mathbb{R}^2$ (recalling $(\nabla E)(\varphi(s)) \neq 0$), there exists $\gamma(s) \in \mathbb{R}$ such that

$$\varphi'(s) = \gamma(s) f(\varphi(s)) \tag{5.19}$$

(note $f(\varphi(s)) \neq 0$ as $s \in \tilde{I}$). We can now apply Prop. 5.14, since (5.19) says φ is a solution to (5.11), the function $\gamma: \tilde{I} \longrightarrow \mathbb{R}$, $s \mapsto \gamma(s) = \varphi'(s)/f(\varphi(s))$ is continuous, and $\varphi'(s) = (1, g'(s)) \neq (0,0)$ for each $s \in \tilde{I}$. Thus, Prop. 5.14 provides a bijective $\lambda: J \longrightarrow \tilde{I}$, such that $\varphi \circ \lambda$ is a solution to $y' = f(y)$.

As we assume $\lim_{m\to\infty} \phi(x_m) = \eta$, there exists $M \in \mathbb{N}$ such that $\|\phi(x_m) - \eta\| < \epsilon$ for each $m \geq M$. Since $E(\phi(x_m)) = C$ also holds, (5.17b) implies the existence of a sequence $(s_m)_{m\in\mathbb{N}}$ in $\tilde{I}$ such that $\phi(x_m) = (s_m, g(s_m))$ for each $m \geq M$. Then, for each $m \geq M$ and $\tau_m := \lambda^{-1}(s_m)$, $(\varphi \circ \lambda)(\tau_m) = \varphi(s_m) = \phi(x_m)$. On the other hand, for $\tau_0 := \lambda^{-1}(\eta_1)$, $(\varphi \circ \lambda)(\tau_0) = \varphi(\eta_1) = \eta$, showing $\phi(x_m), \eta \in \mathcal{O}(\varphi \circ \lambda)$. Since $\phi(x_m) \in \mathcal{O}(\phi)$ as well, Prop. 5.9 implies $\mathcal{O}(\varphi \circ \lambda) \subseteq \mathcal{O}(\phi)$, i.e. $\eta \in \mathcal{O}(\phi)$, which proves (a). In preparation for (b), we also observe that $\|\phi(x_m) - \eta\| < \epsilon$ for each $m \geq M$ implies the s_m for $m \geq M$ all are in some compact interval I_1 with $\eta_1 \in I_1$, implying the τ_m to be in the compact interval $J_1 := \lambda^{-1}[I_1]$ with $\tau_0 \in J_1$. We will use for (b) that J_1 is bounded.

(b): As we have $\mathcal{O}(\phi) \subseteq E^{-1}\{C\}$ according to the choice of C, the assumed compactness of $E^{-1}\{C\}$ and Prop. 3.24 show ϕ can only be maximal if it is defined on all of $\mathbb{R}$ (since $(x, \phi(x))$ must escape every compact $[-m, m] \times E^{-1}\{C\}$, $m \in \mathbb{N}$, on the left and on the right). Using the compactness of $E^{-1}\{C\}$ a second time, we obtain the existence of a sequence $(x_m)_{m\in\mathbb{N}}$ in $\mathbb{R}$ such that $\lim_{m\to\infty} x_m = \infty$ and $\lim_{m\to\infty} \phi(x_m) = \eta \in E^{-1}\{C\}$. So we see that we are in the situation of (a). Let ψ be the maximal extension of the solution $\varphi \circ \lambda$ constructed in the proof of (a). Then we know $\mathcal{O}(\psi) \cap \mathcal{O}(\phi) \neq \emptyset$ from the proof of (a) and, since ψ and ϕ both are maximal, Prop. 5.9 implies $\mathcal{O}(\psi) = \mathcal{O}(\phi)$ and, more importantly for us here, there exists $\xi \in \mathbb{R}$ such that $\psi(x) = \phi(x+\xi)$ for each $x \in \mathbb{R}$. Let $m \geq M$ with M from the proof of (a). If $\xi \neq 0$, then $\phi(x_m) = \psi(\tau_m) = \phi(x_m + \xi)$ shows ϕ is not injective. If $\xi = 0$, then $\phi = \psi$ and $\phi(x_m) = \phi(\tau_m)$. Since the τ_m are bounded, whereas the x_m are unbounded, $x_m = \tau_m$ cannot be true for all m, again showing ϕ is not injective. Since $E^{-1}\{C\} \cap \mathcal{F} = \emptyset$, ϕ cannot be constant, therefore it must be periodic by Prop. 5.10. ■

Example 5.21. Using condition (5.15), i.e. $\nabla E \bullet f \equiv 0$, one readily verifies that the functions

$$E : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad E(y_1, y_2) := y_1^2 + y_2^2, \tag{5.20a}$$
$$E : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad E(y_1, y_2) := y_1^2 - y_2^2, \tag{5.20b}$$

are integrals for (5.8) and (5.9), i.e. for

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix}$$

and

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} y_2 \\ y_1 \end{pmatrix},$$

respectively, and we recover the respective phase portraits via the respective level curves $E(y_1, y_2) = C$, $C \in \mathbb{R}$.

Example 5.22. Consider the autonomous ODE

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 2y_1 y_2 \\ 1 - 2y_1^2 \end{pmatrix}. \tag{5.21}$$

We claim that

$$E : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad E(y_1, y_2) := y_1 \, e^{-(y_1^2 + y_2^2)}, \tag{5.22}$$

is an integral for (5.21) and intend to use Prop. 5.20 to establish (5.21) has orbits that are fixed points, orbits that are periodic, and orbits that are neither. To verify E is an integral, one computes, for each $(y_1, y_2) \in \mathbb{R}^2$,

$$\begin{aligned}
&\nabla E(y_1, y_2) \bullet (2y_1 y_2,\ 1 - 2y_1^2) \\
&= \left(e^{-(y_1^2+y_2^2)} - 2y_1^2 e^{-(y_1^2+y_2^2)},\ -2y_1 y_2 e^{-(y_1^2+y_2^2)}\right) \bullet (2y_1 y_2,\ 1 - 2y_1^2) \\
&= e^{-(y_1^2+y_2^2)} \, (1 - 2y_1^2, -2y_1 y_2) \bullet (2y_1 y_2,\ 1 - 2y_1^2) = 0.
\end{aligned}$$

Clearly, the set of fixed points is

$$\mathcal{F} = \left\{ \left(-\frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{2}}, 0\right) \right\}.$$

The level set of 0 is $E^{-1}\{0\} = \{(0, y_2) : y_2 \in \mathbb{R}\}$, i.e. it is the y_2-axis. This is a nonperiodic orbit (actually, the orbit of solutions of the form $\phi : \mathbb{R} \longrightarrow \mathbb{R}^2$, $\phi(x) := (0, x + c)$, $c \in \mathbb{R}$). Now consider the level set

$$E^{-1}\{e^{-1}\} = \left\{(y_1, y_2) : y_1 > 0,\ y_2^2 = \ln y_1 - y_1^2 + 1\right\}.$$

Using $g : \mathbb{R}^+ \longrightarrow \mathbb{R}$, $g(y_1) = \ln y_1 - y_1^2 + 1$, and its derivative, it is not hard to show g has precisely two zeros, namely $\lambda_1 = 1$ and $0 < \lambda_2 < 1$, and $g \geq 0$ precisely on the compact interval $J := [\lambda_2, 1]$, implying $E^{-1}\{e^{-1}\} = \big\{\big(y_1, \pm\sqrt{g(y_1)}\big) : y_1 \in J\big\}$, showing $E^{-1}\{e^{-1}\}$ is compact. According to Prop. 5.20(b), $E^{-1}\{e^{-1}\}$ must consist of one or more periodic orbits.

5.2 Stability at Fixed Points

Given an autonomous ODE with a fixed point p, we will investigate the question under what conditions a solution $\phi(x)$ starting out near p will remain near p as x increases or decreases.

To simplify notation, we will restrict ourselves to initial data $y(0) = y_0$, which, in light of Lem. 5.4(b), is not an essential restriction.

Notation 5.23. Let $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{K}^n$ such that

$$y' = f(y) \tag{5.23}$$

admits unique maximal solutions (f being locally Lipschitz on Ω open is sufficient). Let $Y : D_f \longrightarrow \mathbb{K}^n$ denote the general solution to (5.23) and define

$$\begin{aligned} Y : D_{f,0} \longrightarrow \mathbb{K}^n, \quad Y(x,\eta) := Y(x,0,\eta), \\ D_{f,0} := \{(x,\eta) \in \mathbb{R} \times \mathbb{K}^n : (x,0,\eta) \in D_f\}. \end{aligned} \tag{5.24}$$

Definition 5.24. Let $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{K}^n$ such that (5.23) admits unique maximal solutions (f being locally Lipschitz on Ω open is sufficient). Moreover, assume the set of fixed points to be nonempty, $\mathcal{F} \neq \emptyset$, and let $p \in \mathcal{F}$. The fixed point p is said to be *positively* (resp. *negatively*) *stable* if, and only if, the following conditions (i) and (ii) hold:

(i) There exists $r > 0$ such that, for each $\eta \in \Omega$ with $\|\eta - p\| < r$, the maximal solution $x \mapsto Y(x,\eta)$ (cf. (5.24)) is defined on (a superset of) $\mathbb{R}_0^+$ (resp. $\mathbb{R}_0^-$).

(ii) For each $\epsilon > 0$, there exists $\delta > 0$ such that, for each $\eta \in \Omega$,

$$\|\eta - p\| < \delta \quad \Rightarrow \quad \underset{\substack{x \geq 0 \\ (\text{resp. } x \leq 0)}}{\forall} \quad \|Y(x,\eta) - p\| < \epsilon. \tag{5.25}$$

The fixed point p is said to be *positively* (resp. *negatively*) *asymptotically stable* if, and only if, (i) and (ii) hold plus the additional condition

(iii) There exists $\gamma > 0$ such that, for each $\eta \in \Omega$,

$$\|\eta - p\| < \gamma \quad \Rightarrow \quad \lim_{x \to \infty} Y(x,\eta) = p \quad \Big(\text{resp.} \quad \lim_{x \to -\infty} Y(x,\eta) = p\Big). \tag{5.26}$$

The norm $\|\cdot\|$ on $\mathbb{K}^n$ used in (i) – (iii) above is arbitrary. Due to the equivalence of norms on $\mathbb{K}^n$, changing the norm does not change the defined stability properties, even though, in general, it does change the sizes of r, δ, γ.

Remark 5.25. In the situation of Def. 5.24, consider the time-reversed version of (5.23), i.e.

$$y' = -f(y). \tag{5.27}$$

According to Lem. 1.9(b), (5.27) has the general solution

$$\begin{aligned} \tilde{Y} : D_{-f,0} \longrightarrow \mathbb{K}^n, \quad \tilde{Y}(x,\eta) := Y(-x,\eta), \\ D_{-f,0} = \{(x,\eta) \in \mathbb{R} \times \mathbb{K}^n : (-x,\eta) \in D_{f,0}\}. \end{aligned} \tag{5.28}$$

(a) Clearly, for a fixed point $p \in \mathcal{F}$, we have the following equivalences:

$$\begin{aligned} p \text{ is positively stable for (5.23)} \quad &\Leftrightarrow \quad p \text{ is negatively stable for (5.27)}, \\ p \text{ is negatively stable for (5.23)} \quad &\Leftrightarrow \quad p \text{ is positively stable for (5.27)}. \end{aligned}$$

(b) Clearly, for a fixed point $p \in \mathcal{F}$, we have the following equivalences:

$$\begin{aligned} p \text{ is pos. asympt. stable for (5.23)} \quad &\Leftrightarrow \quad p \text{ is neg. asympt. stable for (5.27)}, \\ p \text{ is neg. asympt. stable for (5.23)} \quad &\Leftrightarrow \quad p \text{ is pos. asympt. stable for (5.27)}. \end{aligned}$$

Lemma 5.26. *Consider the situation of Def. 5.24 with $f : \Omega \longrightarrow \mathbb{K}^n$ continuous on $\Omega \subseteq \mathbb{K}^n$ open. Then the fixed point p is positively (resp. negatively) stable if, and only if, for each $\epsilon > 0$, there exists $\delta > 0$ such that, for each $\eta \in \Omega$,*

$$\|\eta - p\| < \delta \quad \Rightarrow \quad \mathop{\forall}_{\substack{x \in I_{(0,\eta)} \cap \mathbb{R}_0^+ \\ (\text{resp. } x \in I_{(0,\eta)} \cap \mathbb{R}_0^-)}} \quad \|Y(x,\eta) - p\| < \epsilon, \tag{5.29}$$

where $I_{(0,\eta)}$ denotes the domain of the maximal solution $Y(\cdot,\eta)$.

Proof. Clearly, stability in the sense of Def. 5.24 implies (5.29), and it merely remains to show that (5.29) implies Def. 5.24(i). As (5.29) holds, we can consider $\epsilon := 1$ and obtain a corresponding $\delta =: r$. Then (5.29) states that, for each $\eta \in \Omega$ with $\|\eta - p\| < r$, for $x \geq 0$ (resp. for $x \leq 0$), the maximal solution $Y(x,\eta)$ remains in the compact set $\overline{B}_1(p)$. Since $f : \Omega \longrightarrow \mathbb{K}^n$ is continuous on $\Omega \subseteq \mathbb{K}^n$ open, Th. 3.28 implies $\mathbb{R}_0^+ \subseteq I_{(0,\eta)}$ (resp. $\mathbb{R}_0^- \subseteq I_{(0,\eta)}$), proving Def. 5.24(i). ■

It is an exercise to show Lem. 5.26 becomes false if the hypothesis that f be continuous is omitted.

Example 5.27. **(a)** Consider the 1-dimensional $\mathbb{R}$-valued ODE

$$y' = y(y-1). \tag{5.30}$$

The set of fixed points is $\mathcal{F} = \{0, 1\}$. Moreover, $Y'(\cdot,\eta) < 0$ for $0 < \eta < 1$ and $Y'(\cdot,\eta) > 0$ for $\eta \in]-\infty, 0[\cup]1,\infty[$. It follows that, for $p = 0$, the positive stability part of (5.29) holds (where, given $\epsilon > 0$, one can choose $\delta := \min\{1, \epsilon\}$). Moreover, for $\eta < 0$ and $0 < \eta < 1$, one has $\lim_{x\to\infty} Y(x, \eta) = 0$. Thus, all three conditions of Def. 5.24 are satisfied and 0 is positively asymptotically stable. Analogously, one sees that 1 is negatively asymptotically stable.

(b) For the $\mathbb{R}^2$-valued ODE of (5.8), $(0, 0)$ is a fixed point that is positively and negatively stable, but neither positively nor negatively asymptotically stable. For the $\mathbb{R}^2$-valued ODE of (5.9), $(0, 0)$ is a fixed point that is neither positively nor negatively stable.

(c) Consider the 1-dimensional $\mathbb{R}$-valued ODE

$$y' = y^2. \tag{5.31}$$

The only fixed point is 0, which is neither positively nor negatively stable. Indeed, not even Def. 5.24(i) is satisfied: One obtains

$$Y : D_{f,0} \longrightarrow \mathbb{R}, \quad Y(x, \eta) := \frac{\eta}{1 - \eta x},$$

where

$$\begin{aligned} D_{f,0} =& (\mathbb{R} \times \{0\}) \\ & \cup \big\{(x, \eta) \in \mathbb{R}^2 : \eta > 0,\ x \in]-\infty, 1/\eta[\big\} \\ & \cup \big\{(x, \eta) \in \mathbb{R}^2 : \eta < 0,\ x \in]1/\eta, \infty[\big\}, \end{aligned}$$

showing every neighborhood of 0 contains η such that $Y(\cdot, \eta)$ is not defined on all of $\mathbb{R}_0^+$ and η such that $Y(\cdot, \eta)$ is not defined on all of $\mathbb{R}_0^-$.

Remark 5.28. There exist examples of autonomous ODE that show fixed points can satisfy Def. 5.24(iii) without satisfying Def. 5.24(ii). For example, [Aul04, Ex. 7.4.16] provides the following ODE in *polar coordinates* (r, φ):

$$r' = r\,(1 - r), \tag{5.32a}$$

$$\varphi' = \frac{1 - \cos\varphi}{2} = \sin^2\frac{\varphi}{2}. \tag{5.32b}$$

Even though it is somewhat tedious, one can show that its fixed point $(1, 0)$ satisfies Def. 5.24(iii) without satisfying Def. 5.24(ii) (see Claim 4 of Example K.2 in the Appendix).

—

We will now study a method that allows, in certain cases, to determine the stability properties of a fixed point without having to know the solutions to an ODE. The method is known as *Lyapunov's method*. The key ingredient to this method is a test function V, known as a *Lyapunov function*. Once a Lyapunov function is known, stability is often easily tested. The catch, however, is that Lyapunov functions can be hard to find. From the literature, it appears there is no definition for an all-purpose Lyapunov function, as a suitable choice depends on the circumstances.

Definition 5.29. Let $\Omega_0 \subseteq \mathbb{R}^n$ be open, $n \in \mathbb{N}$. A function $V : \Omega_0 \longrightarrow \mathbb{R}$ is said to be *positive* (resp. *negative*) *definite* at $p \in \Omega_0$ if, and only if, the following conditions (i) and (ii) hold:

(i) $V(y) \geq 0$ (resp. $V(y) \leq 0$) for each $y \in \Omega_0$.

(ii) $V(y) = 0$ if, and only if, $y = p$.

Theorem 5.30 (Lyapunov)**.** *Consider the situation of Def.* 5.24 *with* $\mathbb{K} = \mathbb{R}$, $\Omega \subseteq \mathbb{R}^n$ *open, and* $f : \Omega \longrightarrow \mathbb{R}^n$ *continuous. Let* Ω_0 *be open with* $p \in \Omega_0 \subseteq \Omega \subseteq \mathbb{R}^n$. *Assume* $V : \Omega_0 \longrightarrow \mathbb{R}$ *to be continuously differentiable and define*

$$\dot{V} : \Omega_0 \longrightarrow \mathbb{R}, \quad \dot{V}(y) := (\nabla V)(y) \bullet f(y) = \sum_{j=1}^{n} \partial_j V(y)\, f_j(y). \tag{5.33}$$

If V *is positive definite at* p *and* $\dot{V} \leq 0$ *(resp.* $\dot{V} \geq 0$*) on* Ω_0*, then* p *is positively (resp. negatively) stable. If, in addition,* $\dot{V}$ *is negative (resp. positive) definite at* p*, then* p *is positively (resp. negatively) asymptotically stable.*

Proof. The proof is carried out for the case of postive (asymptotic) stability; the proof for the case of negative (asymptotic) stability is then easily obtained by reversing time, i.e. by using Rem. 5.25 together with noting $\dot{V}$ changing its sign when replacing f with $-f$. Fix your favorite norm $\|\cdot\|$ on $\mathbb{R}^n$. Let $r > 0$ such that $\overline{B}_r(p) = \{y \in \mathbb{R}^n : \|y-p\| \leq r\} \subseteq \Omega_0$ (such an $r > 0$ exists, as Ω_0 is open). Define

$$k :]0, r] \longrightarrow \mathbb{R}^+, \quad k(\epsilon) := \min\big\{V(y) : \|y - p\| = \epsilon\big\}, \tag{5.34}$$

where k is well-defined, since the continuous function V assumes its min on compact sets, and $k(\epsilon) > 0$ by the positive definiteness of V. Given $\epsilon \in]0, r]$, since $V(p) = 0$, $k(\epsilon) > 0$, and V continuous,

$$\underset{0<\delta(\epsilon)<\epsilon}{\exists} \quad \underset{y \in B_{\delta(\epsilon)}(p)}{\forall} \quad V(y) < k(\epsilon), \tag{5.35}$$

where we used Not. 3.3 to denote an open ball with center p with respect to $\|\cdot\|$.

We now claim that, for each $\eta \in B_{\delta(\epsilon)}(p)$, the maximal solution $x \mapsto \phi(x) := Y(x, \eta)$ must remain inside $B_\epsilon(p)$ for each $x \geq 0$ in its domain $I_{(0,\eta)}$ (implying p to be positively stable by Lem. 5.26). Seeking a contradiction, assume there exists $\xi \geq 0$ such that $\|\phi(\xi) - p\| \geq \epsilon$ and let

$$s := \sup\Big\{x \geq 0 : \phi(t) \in B_\epsilon(p) \text{ for each } t \in [0, x]\Big\} \leq \xi < \infty. \tag{5.36}$$

The continuity of ϕ then implies $\|\phi(s) - p\| = \epsilon$, i.e.

$$V(\phi(s)) \geq k(\epsilon) \tag{5.37}$$

by the definition of $k(\epsilon)$. On the other hand, by the chain rule $(V\circ\phi)'(x) = \dot V(\phi(x))$ (cf. (5.16)), such that $\dot V \le 0$ implies

$$V(\phi(s)) = V(\eta) + \int_0^s \dot V(\phi(x))\,\mathrm{d}x \le V(\eta) \overset{(5.35)}{<} k(\epsilon), \tag{5.38}$$

in contradiction to (5.37), proving $\phi(x) \in B_\epsilon(p)$ for each $x \in I_{(0,\eta)} \cap \mathbb{R}_0^+$ and the positive stability of p.

For the remaining part of the proof, we additionally assume $\dot V$ to be negative definite at p, while continuing to use the notation from above. Set $\gamma := \delta(r)$. We have to show $\lim_{x\to\infty} Y(x,\eta) = p$ for each $\eta \in B_\gamma(p)$, i.e.

$$\underset{\epsilon\in]0,r]}{\forall} \quad \underset{\xi_\epsilon\ge 0}{\exists} \quad \underset{x\ge\xi_\epsilon}{\forall} \quad \|Y(x,\eta) - p\| < \epsilon. \tag{5.39}$$

So fix $\eta \in B_\gamma(p)$ and, as above, let $\phi(x) := Y(x,\eta)$. Given $\epsilon \in]0,r]$, we first claim that there exists $\xi_\epsilon \ge 0$ such that $\phi(\xi_\epsilon) \in B_{\delta(\epsilon)}(p)$, where $\delta(\epsilon)$ is as in the first part of the proof above. Indeed, seeking a contradiction, assume $\|\phi(x) - p\| \ge \delta(\epsilon)$ for all $x \ge 0$, and set

$$\alpha := \max\big\{\dot V(y) : \delta(\epsilon) \le \|y - p\| \le r\big\}. \tag{5.40}$$

Then $\alpha < 0$ due to the negative definiteness of $\dot V$ at p. Moreover, due to the choice of γ, we have $\delta(\epsilon) \le \|\phi(x) - p\| \le r$ for each $x \ge 0$, implying

$$\underset{x\ge 0}{\forall} \quad 0 \le V(\phi(x)) = V(\eta) + \int_0^x \dot V(\phi(t))\,\mathrm{d}t \le V(\eta) + \alpha x, \tag{5.41}$$

which is the desired contradiction, as $\alpha < 0$ implies the right-hand side to go to $-\infty$ for $x \to \infty$. Thus, we know the existence of ξ_ϵ such that $\eta_\epsilon := \phi(\xi_\epsilon) \in B_{\delta(\epsilon)}(p)$.

To finish the proof, we recall from the first part of the proof that $\|Y(x,\eta_\epsilon) - p\| < \epsilon$ for each $x \ge 0$. Using Lem. 5.4(a), we obtain

$$\underset{x\ge 0}{\forall} \quad \phi(\xi_\epsilon + x) = Y(\xi_\epsilon + x, \xi_\epsilon, \eta_\epsilon) = Y(\xi_\epsilon + x - \xi_\epsilon, \eta_\epsilon) = Y(x, \eta_\epsilon) \in B_\epsilon(p), \tag{5.42}$$

showing $\|\phi(x) - p\| < \epsilon$ for each $x \ge \xi_\epsilon$ as needed. ■

Example 5.31. Let $k, m \in \mathbb{N}$ and $\alpha, \beta > 0$. We claim that $(0,0)$ is a positively asymptotically stable fixed point for each $\mathbb{R}^2$-valued ODE of the form

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} -y_1^{2k-1} + \alpha y_1 y_2^2 \\ -y_2^{2m-1} - \beta y_1^2 y_2 \end{pmatrix}. \tag{5.43}$$

Indeed, $(0,0)$ is clearly a fixed point, and we consider the Lyapunov function

$$V : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad V(y_1, y_2) := \frac{y_1^2}{\alpha} + \frac{y_2^2}{\beta}, \tag{5.44a}$$

which is clearly positive definite at $(0,0)$. Since $\dot V : \mathbb{R}^2 \longrightarrow \mathbb{R}$,

$$\begin{aligned}\dot V(y_1,y_2) &= \nabla V(y_1,y_2) \bullet \big(-y_1^{2k-1} + \alpha y_1 y_2^2,\ -y_2^{2m-1} - \beta y_1^2 y_2\big)\\ &= (2y_1/\alpha,\ 2y_2/\beta) \bullet \big(-y_1^{2k-1} + \alpha y_1 y_2^2,\ -y_2^{2m-1} - \beta y_1^2 y_2\big)\\ &= -2(y_1^{2k}/\alpha + y_2^{2m}/\beta), \qquad (5.44b)\end{aligned}$$

is clearly negative definite at $(0,0)$, Th. 5.30 proves $(0,0)$ to be a positively asymptotically stable fixed point.

Theorem 5.32. *Consider the situation of Def.* 5.24 *with* $\mathbb{K} = \mathbb{R}$. *Let* Ω_0 *be open with* $p \in \Omega_0 \subseteq \Omega \subseteq \mathbb{R}^n$. *Assume* $V : \Omega_0 \longrightarrow \mathbb{R}$ *to be continuously differentiable and assume there is an open set* $U \subseteq \Omega_0$ *such that the following conditions* (i) – (iii) *are satisfied:*

(i) $p \in \partial U$, *i.e.* p *is in the boundary of* U.

(ii) $V > 0$ *and* $\dot V > 0$ *(resp.* $\dot V < 0$*) on* U *(where* $\dot V$ *is defined as in* (5.33)*).*

(iii) $V(y) = 0$ *for each* $y \in \Omega_0 \cap \partial U$.

Then the fixed point p *is not positively (resp. negatively) stable.*

Proof. We assume V and $\dot V$ are positive, proving p not to be positively stable; the corresponding statement regarding p not to be negatively stable is then, once again, easily obtained by reversing time, i.e. by using Lem. 5.25 together with noting $\dot V$ changing its sign when replacing f with $-f$.

Seeking a contradiction, assume p to be positively stable. Then there exists $r > 0$ such that $\overline{B}_r(p) = \{y \in \mathbb{R}^n : \|y - p\| \le r\} \subseteq \Omega_0$ and $\eta \in B_r(p)$ implies $Y(x,\eta)$ is defined for each $x \ge 0$. Moreover, positive stability and $p \in \partial U$ also imply the existence of $\eta \in U \cap B_r(p)$ such that $\phi(x) := Y(x,\eta) \in B_r(p)$ for all $x \ge 0$ (note $p \ne \eta$ as $p \in \partial U$). Set

$$s := \sup\big\{x \ge 0 : \phi(t) \in U \text{ for each } t \in [0,x]\big\}. \qquad (5.45)$$

If $s < \infty$, then the maximality of ϕ implies $\phi(s)$ to be defined. Moreover, $\phi(s) \in \partial U$ by the definition of s, and $\phi(s) \in B_r(p) \subseteq \Omega_0$ by the choice of η. Thus, $\phi(s) \in \Omega_0 \cap \partial U$ and $V(\phi(s)) = 0$. On the other hand, as V and $\dot V$ are positive on U, we have

$$V(\phi(s)) = V(\eta) + \int_0^s \dot V(\phi(t))\,\mathrm{d}t > V(\eta) > 0, \qquad (5.46)$$

which is a contradiction to $V(\phi(s)) = 0$, implying $s = \infty$ and $\phi(x) \in U$ as well as $V(\phi(x)) > V(\eta) > 0$ hold for each $x > 0$.

To conclude the proof, consider the compact set

$$C := \overline{B}_r(p) \cap \overline{U} \cap V^{-1}[V(\eta), \infty[\subseteq \Omega_0. \qquad (5.47)$$

Then the choice of η guarantees $\phi(x) \in C$ for all $x \ge 0$. If $y \in C$, then $V(y) \ge V(\eta) > 0$. If $y \in \Omega_0 \cap \partial U$, then $V(y) = 0$, showing $C \cap \partial U = \emptyset$, i.e. $C \subseteq U$ and

$$\alpha := \min\{\dot V(y) : y \in C\} > 0. \qquad (5.48)$$

Thus,

$$\underset{x\geq 0}{\forall} \quad V(\phi(x)) = V(\eta) + \int_0^x \dot{V}(\phi(t))\,\mathrm{d}t \geq V(\eta) + \alpha x. \tag{5.49}$$

But this means that the continuous function V is unbounded on the compact set C and this contradiction proves p is not positively stable. ■

Example 5.33. Let $h_1, h_2 : \Omega \longrightarrow \mathbb{R}$ be continuously differentiable functions defined on some open set $\Omega \subseteq \mathbb{R}^2$ with $(0,0) \in \Omega$ and $h_1(0,0) > 0$, $h_2(0,0) > 0$. We claim that $(0,0)$ is not a positively stable fixed point for each $\mathbb{R}^2$-valued ODE of the form

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} y_2\, h_1(y_1,y_2) \\ y_1\, h_2(y_1,y_2) \end{pmatrix}. \tag{5.50}$$

Indeed, $(0,0)$ is clearly a fixed point, and we let Ω_0 be some open neighborhood of $(0,0)$, where both h_1 and h_2 are positive (such an Ω_0 exists by continuity of h_1, h_2 and h_1, h_2 being positive at $(0,0)$), and consider the Lyapunov function

$$V : \Omega_0 \longrightarrow \mathbb{R}, \quad V(y_1,y_2) := y_1 y_2, \tag{5.51a}$$

with $\dot{V} : \Omega_0 \longrightarrow \mathbb{R}$,

$$\begin{aligned} \dot{V}(y_1,y_2) &= \nabla V(y_1,y_2) \bullet \big(y_2\, h_1(y_1,y_2),\ y_1\, h_2(y_1,y_2)\big) \\ &= (y_2,\ y_1) \bullet \big(y_2\, h_1(y_1,y_2),\ y_1\, h_2(y_1,y_2)\big) \\ &= y_2^2\, h_1(y_1,y_2) + y_1^2\, h_2(y_1,y_2) > 0 \quad \text{on } \Omega_0 \setminus \{(0,0)\}. \end{aligned} \tag{5.51b}$$

Letting $U := \Omega_0 \cap (\mathbb{R}^+ \times \mathbb{R}^+)$, one has $(0,0) \in \partial U$, both V and $\dot{V}$ are positive on U, and $V = 0$ on $\Omega_0 \cap \partial U \subseteq (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})$. Thus, Th. 5.32 applies, yielding that $(0,0)$ is not positively stable.

Theorem 5.34. *Let $\Omega \subseteq \mathbb{R}^n$ be open, $n \in \mathbb{N}$. Let $F : \Omega \longrightarrow \mathbb{R}$ be C^2 and consider*

$$y' = -\nabla F(y). \tag{5.52}$$

If $p \in \Omega$ is an isolated critical point of F (i.e. $\nabla F(p) = 0$ and there exists an open set O with $p \in O \subseteq \Omega$ and $\nabla F \neq 0$ on $O \setminus \{p\}$), then p is a fixed point of (5.52) that is positively asymptotically stable, negatively asymptotically stable, neither positively nor negatively stable as p is a local minimum for F, local maximum for F, neither.

Proof. Note that F being C^2 implies ∇F to be C^1 and, in particular, locally Lipschitz, such that (5.52) admits unique maximal solutions. Suppose F has a local min at p. As p is an isolated critical point, the local min at p must be strict, i.e. there exists an open neighborhood Ω_0 of p such that $F(p) < F(y)$ for each $y \in \Omega_0 \setminus \{p\}$. Then the Lyapunov function $V : \Omega_0 \longrightarrow \mathbb{R}$, $V(y) := F(y) - F(p)$, is clearly positive definite at p and $\dot{V} : \Omega_0 \longrightarrow \mathbb{R}$,

$$\dot{V}(y) = \nabla V(y) \bullet (-\nabla F(y)) = -\nabla F(y) \bullet \nabla F(y) = -\|\nabla F(y)\|_2^2, \tag{5.53}$$

is clearly negative definite at p. Thus, p is a positively asymptotically stable fixed point by Th. 5.30. If F has a local max at p, then the proof is conducted analogously, using $V : \Omega_0 \longrightarrow \mathbb{R}$, $V(y) := F(p) - F(y)$, or, alternatively, by using time reversion (if F has a local max at p, then $-F$ has a local min at p, i.e. p is a positively asymptotically stable for $y' = \nabla F(y)$, i.e. p is a negatively asymptotically stable for $y' = -\nabla F(y)$ by Rem. 5.25(b)).

If p is neither a local min nor max for F, then let $\Omega_0 := O$, and $V : \Omega_0 \longrightarrow \mathbb{R}$, $V(y) := F(p) - F(y)$, where O was chosen such that $\nabla F \neq 0$ on $O \setminus \{p\}$, i.e. $\dot{V} : \Omega_0 \longrightarrow \mathbb{R}$, $\dot{V}(y) = \|\nabla F(y)\|_2^2$, is positive definite at p. Let $U := \{y \in \Omega_0 : F(y) < F(p)\}$. Then U is open by the continuity of F, and $p \in \partial U$, as p is neither a local min nor max for F. By the continuity of F, $F(y) = F(p)$ for each $y \in \Omega_0 \cap \partial U$, i.e. $V = 0$ on $\Omega_0 \cap \partial U$. Thus, Th. 5.32 applies, showing p is not positively stable. Analogously, using $U := \{y \in \Omega_0 : F(y) > F(p)\}$ and $V(y) := F(y) - F(p)$ shows p is not negatively stable. ■

Example 5.35. **(a)** The function $F : \mathbb{R}^n \longrightarrow \mathbb{R}$, $F(y) = \|y\|_2^2$, has an isolated critical point at 0, which is also a min for F. Thus, by Th. 5.34,

$$y' = -\nabla F(y) = (-2y_1, \dots, -2y_n) \tag{5.54}$$

has 0 as a fixed point that is positively asymptotically stable.

(b) The function $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $F(y) = e^{y_1 y_2}$, has an isolated critical point at 0, which is neither a local min nor local max for F. Thus, by Th. 5.34,

$$y' = -\nabla F(y) = (-y_2 e^{y_1 y_2}, -y_1 e^{y_1 y_2}) \tag{5.55}$$

has 0 as a fixed point that is neither positively nor negatively stable.

5.3 Constant Coefficients

The stability properties of systems of first-order linear ODE (cf. Sec. 4.6.2) are closely related to the eigenvalues of the matrix A. As it turns out, the stability of the origin is essentially determined by the sign of the real part of the eigenvalues of A (cf. Th. 5.38 below). We start with a preparatory lemma:

Lemma 5.36. *Let $n \in \mathbb{N}$ and $W \in \mathcal{M}(n, \mathbb{K})$ be invertible. Moreover, let $\|\cdot\|$ be some norm on $\mathcal{M}(n, \mathbb{K})$. Then*

$$\|\cdot\|^W : \mathcal{M}(n, \mathbb{K}) \longrightarrow \mathbb{R}_0^+, \quad \|A\|^W := \|W^{-1} A W\|, \tag{5.56}$$

also constitutes a norm on $\mathcal{M}(n, \mathbb{K})$.

Proof. If $A = 0$, then $\|A\|^W = \|W^{-1} 0 W\| = \|0\| = 0$. If $\|A\|^W = 0$, then $W^{-1} A W = 0$, i.e. $A = W 0 W^{-1} = 0$, showing $\|\cdot\|^W$ is positive definite. Next,

$$\forall_{\lambda \in \mathbb{K}} \quad \|\lambda A\|^W = |\lambda| \, \|W^{-1} A W\| = |\lambda| \, \|A\|^W,$$

showing $\|\cdot\|^W$ is homogeneous of degree 1. Finally,

$$\forall_{A,B\in\mathcal{M}(n,\mathbb{K})} \quad \|A+B\|^W = \|W^{-1}(A+B)W\| = \|W^{-1}AW + W^{-1}BW\| \leq \|A\|^W + \|B\|^W,$$

showing $\|\cdot\|^W$ satisfies the triangle inequality. ∎

Remark and Definition 5.37. Let $n \in \mathbb{N}$, $A \in \mathcal{M}(n,\mathbb{C})$, and let $\lambda \in \mathbb{C}$ be an eigenvalue of A.

(a) Clearly, one has

$$\{0\} \subseteq \ker(A - \lambda\,\mathrm{Id}) \subseteq \ker(A - \lambda\,\mathrm{Id})^2 \subseteq \dots$$

and the inclusion can be strict for at most n times. Let

$$r(\lambda) := \min\big\{k \in \mathbb{N}_0 : \ker(A - \lambda\,\mathrm{Id})^k = \ker(A - \lambda\,\mathrm{Id})^{k+1}\big\}.$$

Then

$$\forall_{k\in\mathbb{N}} \quad \ker(A - \lambda\,\mathrm{Id})^{r(\lambda)} = \ker(A - \lambda\,\mathrm{Id})^{r(\lambda)+k} :$$

Indeed, otherwise, let $k_0 := \min\{k \in \mathbb{N} : \ker(A - \lambda\,\mathrm{Id})^{r(\lambda)} \subsetneq \ker(A - \lambda\,\mathrm{Id})^{r(\lambda)+k}\}$. Then there exists $v \in \mathbb{C}^n$ such that $(A-\lambda\,\mathrm{Id})^{r(\lambda)+k_0}v = 0$, but $(A-\lambda\,\mathrm{Id})^{r(\lambda)+k_0-1}v \neq 0$. However, that means $w := (A - \lambda\,\mathrm{Id})^{k_0-1}v \in \ker(A - \lambda\,\mathrm{Id})^{r(\lambda)+1}$, but $w \notin \ker(A - \lambda\,\mathrm{Id})^{r(\lambda)}$, in contradiction to the definition of $r(\lambda)$. The space

$$M(\lambda) := \ker(A - \lambda\,\mathrm{Id})^{r(\lambda)}$$

is called the *generalized eigenspace* corresponding to the eigenvalue λ.

(b) Due to $A(A - \lambda\,\mathrm{Id}) = (A - \lambda\,\mathrm{Id})A$, one has

$$\forall_{k\in\mathbb{N}_0} \quad A\big(\ker(A - \lambda\,\mathrm{Id})^k\big) \subseteq \ker(A - \lambda\,\mathrm{Id})^k,$$

i.e. all the kernels (in particular, the generalized eigenspace $M(\lambda)$) are invariant subspaces for A.

(c) As already mentioned in Rem. 4.51 the *algebraic multiplicity* of λ, denoted $m_{\mathrm{a}}(\lambda)$, is its multiplicity as a zero of the characteristic polynomial $\chi_A(x) = \det(A - x\,\mathrm{Id})$, and the *geometric multiplicity* of λ is $m_{\mathrm{g}}(\lambda) := \dim\ker(A - \lambda\,\mathrm{Id})$. We call the eigenvalue λ *semisimple* if, and only if, its algebraic and geometric multiplicities are equal. We then have the equivalence of the following statements (i) – (iv):

(i) λ is semisimple.

(ii) $M(\lambda) = \ker(A - \lambda\,\mathrm{Id})$.

(iii) $A{\restriction}_{M(\lambda)}$ is diagonalizable.

(iv) All the Jordan blocks corresponding to λ are trivial, i.e. they all have size 1 (i.e. there are $\dim\ker(A - \lambda\,\mathrm{Id})$ such blocks).

Indeed, note that $m_a(\lambda) = \dim\ker(A - \lambda\,\mathrm{Id})^{m_a(\lambda)}$ (e.g., since, if A is in Jordan normal form, then $m_a(\lambda)$ provides the size of the λ-block and, for $A - \lambda\,\mathrm{Id}$, this block is canonically nilpotent). This shows the equivalence between (i) and (ii). Moreover, $m_g(\lambda) = m_a(\lambda)$ means $\ker(A - \lambda\,\mathrm{Id})$ has a basis of $m_a(\lambda)$ eigenvectors $v_1, \dots, v_{m_a(\lambda)}$ for the eigenvalue λ. The equivalence of (i),(ii) with (iii) and with (iv) is then given by Th. 4.45 and Th. 4.46, respectively.

Theorem 5.38. *Let $n \in \mathbb{N}$ and $A \in \mathcal{M}(n, \mathbb{C})$. Moreover, let $\|\cdot\|$ be some norm on $\mathcal{M}(n, \mathbb{C})$ and let $\lambda_1, \dots, \lambda_s \in \mathbb{C}$, $1 \le s \le n$, be the distinct eigenvalues of A.*

(a) *The following statements* (i) – (iii) *are equivalent:*

(i) *There exists $K > 0$ such that $\|e^{Ax}\| \le K$ holds for each $x \ge 0$ (resp. $x \le 0$).*

(ii) $\operatorname{Re}\lambda_j \le 0$ *(resp.* $\operatorname{Re}\lambda_j \ge 0$*) for every $j = 1, \dots, s$ and if* $\operatorname{Re}\lambda_j = 0$ *occurs, then λ_j is a semisimple eigenvalue (i.e. its algebraic and geometric multiplicities are equal).*

(iii) *The fixed point* 0 *of $y' = Ay$ is positively (resp. negatively) stable.*

(b) *The following statements* (i) – (iii) *are equivalent:*

(i) *There exist $K, \alpha > 0$ such that $\|e^{Ax}\| \le K e^{-\alpha|x|}$ holds for each $x \ge 0$ (resp. $x \le 0$).*

(ii) $\operatorname{Re}\lambda_j < 0$ *(resp.* $\operatorname{Re}\lambda_j > 0$*) for every $j = 1, \dots, s$.*

(iii) *The fixed point* 0 *of $y' = Ay$ is positively (resp. negatively) asymptotically stable.*

Proof. Let $\|\cdot\|_{\max}$ denote the max-norm on $\mathbb{C}^{n^2} \cong \mathcal{M}(n, \mathbb{C})$, i.e.

$$\|(m_{kl})\|_{\max} := \max\big\{|m_{kl}| : k, l \in \{1, \dots, n\}\big\}$$

(caveat: for $n > 1$, this is *not* the operator norm induced by the max-norm on $\mathbb{C}^n$). Moreover, using Th. 4.46, let $W \in \mathcal{M}(n, \mathbb{C})$ be invertible and such that $B := W^{-1}AW$ is in Jordan normal form. Then, according to Lem. 5.36, $\|M\|^W_{\max} := \|W^{-1}MW\|_{\max}$ also defines a norm on $\mathcal{M}(n, \mathbb{C})$. According to Th. 4.47(b),

$$\forall_{x \in \mathbb{R}} \quad \|e^{Ax}\|^W_{\max} = \|W^{-1}e^{Ax}W\|_{\max} = \|e^{W^{-1}AWx}\|_{\max} = \|e^{Bx}\|_{\max}. \tag{5.57}$$

According to Th. 4.44 and Th. 4.49, the entries $\beta_{kl}(x)$ of $(\beta_{kl}(x)) := e^{Bx}$ enjoy the following property:

$$\forall_{k,l \in \{1,\dots,n\}} \quad \exists_{j \in \{1,\dots,s\}} \quad \exists_{C > 0} \quad \exists_{m \in \mathbb{N}_0} \quad \forall_{x \in \mathbb{R}} \quad |\beta_{kl}(x)| = C\, e^{\operatorname{Re}\lambda_j x}\, |x|^m. \tag{5.58}$$

Moreover,

$$\big(|\beta_{kl}(x)| = C\, e^{\operatorname{Re}\lambda_j x}\, |x|^m \ \wedge\ \operatorname{Re}\lambda_j < 0\big) \quad\Rightarrow\quad \lim_{x\to\infty} |\beta_{kl}(x)| = 0, \tag{5.59a}$$

$$\big(|\beta_{kl}(x)| = C\, e^{\operatorname{Re}\lambda_j x}\, |x|^m \ \wedge\ \operatorname{Re}\lambda_j > 0\big) \quad\Rightarrow\quad \lim_{x\to\infty} |\beta_{kl}(x)| = \infty, \tag{5.59b}$$

$$\big(|\beta_{kl}(x)| = C\, e^{\operatorname{Re}\lambda_j x}\, |x|^m \ \wedge\ \operatorname{Re}\lambda_j = 0 \ \wedge\ m = 0\big) \quad\Rightarrow\quad |\beta_{kl}| \equiv C, \tag{5.59c}$$

$$\big(|\beta_{kl}(x)| = C\, e^{\operatorname{Re}\lambda_j x}\, |x|^m \ \wedge\ \operatorname{Re}\lambda_j = 0 \ \wedge\ m > 0\big) \quad\Rightarrow\quad \lim_{x\to\infty} |\beta_{kl}(x)| = \infty. \tag{5.59d}$$

(a): We start with the equivalence between (i) and (ii): Suppose, $\operatorname{Re}\lambda_j \leq 0$ for every $j = 1, \dots, s$ and if $\operatorname{Re}\lambda_j = 0$ occurs, then λ_j is a semisimple eigenvalue. Then, using Rem. and Def. 5.37(c) and (5.58), we are either in situation (5.59a) or in situation (5.59c). Thus, there exists $K_0 > 0$ such that $|\beta_{kl}(x)| \leq K_0$ for each $x \geq 0$ and each $k, l = 1, \dots, n$. Then there exists $K_1 > 0$ such that

$$\underset{x\geq 0}{\forall} \quad \|e^{Ax}\| \leq K_1 \, \|e^{Ax}\|_{\max}^W = K_1 \, \|e^{Bx}\|_{\max} \leq K_1 \, K_0, \tag{5.60}$$

showing (i) holds with $K := K_1 \, K_0$. Conversely, if there is $j \in \{1, \dots, s\}$ such that $\operatorname{Re}\lambda_j > 0$, then there is β_{kl} such that (5.59b) occurs; if there is $j \in \{1, \dots, s\}$ such that $\operatorname{Re}\lambda_j = 0$ and λ_j is not semisimple, then, using Rem. and Def. 5.37(c), there is β_{kl} such that (5.59d) occurs. In both cases,

$$\lim_{x\to\infty} \|e^{Ax}\| = \lim_{x\to\infty} \|e^{Ax}\|_{\max}^W = \lim_{x\to\infty} \|e^{Bx}\|_{\max} = \infty, \tag{5.61}$$

i.e., the corresponding statement of (i) can not be true. The remaining case is handled via time reversion: $\|e^{Ax}\| \leq K$ holds for each $x \leq 0$ if, and only if, $\|e^{-Ax}\| \leq K$ holds for each $x \geq 0$, which holds if, and only if, $\operatorname{Re}(-\lambda_j) \leq 0$ for every $j = 1, \dots, s$ with λ_j semisimple for $\operatorname{Re}(-\lambda_j) = 0$, which is equivalent to $\operatorname{Re}\lambda_j \geq 0$ for every $j = 1, \dots, s$ with λ_j semisimple for $\operatorname{Re}\lambda_j = 0$.

We proceed to the equivalence between (i) and (iii): Fix some arbitary norm $\|\cdot\|$ on $\mathbb{C}^n$, and let $\|\cdot\|_{\text{op}}$ denote the induced operator norm on $\mathcal{M}(n, \mathbb{C})$. Let $C_1, C_2 > 0$ be such that $\|M\|_{\text{op}} \leq C_1 \|M\|$ and $\|M\| \leq C_2 \|M\|_{\text{op}}$ for each $M \in \mathcal{M}(n, \mathbb{C})$. Suppose there exists $K > 0$ such that $\|e^{Ax}\| \leq K$ holds for each $x \geq 0$. Given $\epsilon > 0$, choose $\delta := \epsilon/(C_1 K)$. Then

$$\underset{\eta\in B_\delta(0)}{\forall} \quad \underset{x\geq 0}{\forall} \quad \|Y(x,\eta) - 0\| = \|e^{Ax}\,\eta\| \leq \|e^{Ax}\|_{\text{op}} \, \|\eta\| < C_1 K \, \frac{\epsilon}{C_1 K} = \epsilon, \tag{5.62}$$

proving 0 is positively stable. Conversely, assume 0 to be positively stable. Then there exists $\delta > 0$ such that $\|Y(x,\eta)\| = \|e^{Ax}\,\eta\| < 1$ for each $\eta \in B_\delta(0)$ and each $x \geq 0$. Thus,

$$\underset{x\geq 0}{\forall} \quad \begin{aligned} \|e^{Ax}\|_{\text{op}} &\overset{\text{(G.1)}}{=} \sup\big\{\|e^{Ax}\,\eta\| : \eta \in \mathbb{C}^n, \ \|\eta\| = 1\big\} \\ &= \frac{1}{\delta} \sup\big\{\|e^{Ax}\,\eta\| : \eta \in \mathbb{C}^n, \ \|\eta\| = \delta\big\} \leq \frac{1}{\delta}, \end{aligned} \tag{5.63}$$

showing (i) holds with $K := C_2/\delta$. The remaining case is handled via time reversion: $\|e^{Ax}\| \leq K$ holds for each $x \leq 0$ if, and only if, $\|e^{-Ax}\| \leq K$ holds for each $x \geq 0$, which holds if, and only if, 0 is positively stable for $y' = -Ay$, which, by Rem. 5.25(a), holds if, and only if, 0 is negatively stable for $y' = Ay$.

(b): As in (a), we start with the equivalence between (i) and (ii): Suppose, $\operatorname{Re}\lambda_j < 0$ for every $j = 1, \dots, s$. We first show, using (5.58),

$$\underset{k,l=1,\dots,n}{\forall} \quad \underset{K_{kl},\alpha_{kl}>0}{\exists} \quad \underset{x\geq 0}{\forall} \quad |\beta_{kl}(x)| \leq K_{kl} \, e^{-\alpha_{kl} x} : \tag{5.64}$$

According to (5.58),

$$\underset{C_{kl}>0}{\exists} \quad \underset{x\geq 0}{\forall} \quad |\beta_{kl}(x)| = C_{kl} \, e^{\operatorname{Re}\lambda_j x/2} \, e^{\operatorname{Re}\lambda_j x/2} \, x^m.$$

Since $\operatorname{Re}\lambda_j < 0$, one has $\lim_{x\to\infty} e^{\operatorname{Re}\lambda_j x/2}\, x^m = 0$, i.e. $e^{\operatorname{Re}\lambda_j x/2}\, x^m$ is uniformly bounded on $[0,\infty[$ by some $M_{kl} > 0$. Thus, (5.64) holds with $K_{kl} := C_{kl} M_{kl}$ and $\alpha_{kl} := -\operatorname{Re}\lambda_j/2$. In consequence, if K_1 is chosen as in (5.60), then $\|e^{Ax}\| \le Ke^{-\alpha|x|} \le K$ for each $x \ge 0$ holds with $K := K_1 \max\{K_{kl} : k,l = 1,\dots,n\}$ and $\alpha := \min\{\alpha_{kl} : k,l = 1,\dots,n\}$. Conversely, if there is $j \in \{1,\dots,s\}$ such that $\operatorname{Re}\lambda_j \ge 0$, then there is β_{kl} such that (5.59b) or (5.59c) or (5.59d) occurs. In each case,

$$\lim_{x\to\infty} \|e^{Ax}\| = \lim_{x\to\infty} \|e^{Ax}\|^W_{\max} = \lim_{x\to\infty} \|e^{Bx}\|_{\max} \in\,]0,\infty], \tag{5.65}$$

i.e., the corresponding statement of (i) can not be true. The remaining case is handled via time reversion: $\|e^{Ax}\| \le Ke^{-\alpha|x|}$ holds for each $x \le 0$ if, and only if, $\|e^{-Ax}\| \le Ke^{-\alpha|x|}$ holds for each $x \ge 0$, which holds if, and only if, $\operatorname{Re}(-\lambda_j) < 0$ for every $j = 1,\dots,s$, which is equivalent to $\operatorname{Re}\lambda_j > 0$ for every $j = 1,\dots,s$.

It remains to consider the equivalence between (i) and (iii): Let $\|\cdot\|_{\mathrm{op}}$ and $C_1, C_2 > 0$ be as in the proof of the equivalence between (i) and (iii) in (a). Suppose, there exist $K,\alpha > 0$ such that $\|e^{Ax}\| \le Ke^{-\alpha|x|}$ holds for each $x \ge 0$. Since $\|e^{Ax}\| \le Ke^{-\alpha|x|} \le K$ for each $x \ge 0$, 0 is positively stable by (a). Moreover,

$$\underset{\eta\in\mathbb{C}^n}{\forall}\ \underset{x\ge 0}{\forall}\quad \|Y(x,\eta)\| = \|e^{Ax}\,\eta\| \le \|e^{Ax}\|_{\mathrm{op}}\,\|\eta\| \le C_1 K\, e^{-\alpha|x|}\,\|\eta\| \to 0 \quad \text{for } x\to\infty, \tag{5.66}$$

showing 0 to be positively asymptotically stable. For the converse, we will actually show (iii) implies (ii). If 0 is positively asymptotically stable, then, in particular, it is positively stable, such that (ii) of (a) must hold. It merely remains to exclude the possibility of a semisimple eigenvalue λ with $\operatorname{Re}\lambda = 0$. If there were a semisimple eigenvalue λ with $\operatorname{Re}\lambda = 0$, then e^{Bx} had a Jordan block of size 1 with entry $e^{\lambda x}$, i.e. $\beta_{kk}(x) = e^{\lambda x}$ for some $k \in \{1,\dots,n\}$. Let e_k be the corresponding standard unit vector of $\mathbb{C}^n$ (all entries 0, except the kth entry, which is 1). Then, for $\eta := We_k$,

$$\underset{x\in\mathbb{R}}{\forall}\quad \begin{aligned} \|W^{-1}\,e^{Ax}\,\eta\| &= \|W^{-1}\,e^{Ax}\,We_k\| = \|e^{Bx}\,e_k\| = \|e^{\lambda x}\,e_k\| \\ &= |e^{\lambda x}|\,\|e_k\| = 1\cdot\|e_k\| > 0, \end{aligned} \tag{5.67}$$

showing 0 were not positively asymptotically stable (e.g., since $y \mapsto \|W^{-1}y\|$ defines a norm on $\mathbb{C}^n$). The remaining case is, once again, handled via time reversion: $\|e^{Ax}\| \le Ke^{-\alpha|x|}$ holds for each $x \le 0$ if, and only if, $\|e^{-Ax}\| \le Ke^{-\alpha|x|}$ holds for each $x \ge 0$, which holds if, and only if, 0 is positively asymptotically stable for $y' = -Ay$, which, by Rem. 5.25(b), holds if, and only if, 0 is negatively asymptotically stable for $y' = Ay$. ■

Example 5.39. (a) The matrix

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

has eigenvalues 1 and 3 and, thus, the fixed point 0 of $y' = Ay$ is negatively asymptotically stable, but not positively stable.

(b) The matrix

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

has eigenvalue 0, which is not semisimple, i.e. the fixed point 0 of $y' = Ay$ is neither negatively nor positively stable.

(c) The matrix

$$A = \begin{pmatrix} i & 1 & 2 & 2-3i \\ 0 & -i & 5 & -17 \\ 0 & 0 & -1+3i & 0 \\ 0 & 0 & 0 & -5 \end{pmatrix}$$

has simple eigenvalues i, $-i$, $-1+3i$, -5, i.e. the fixed point 0 of $y' = Ay$ is positively stable (since all real parts are ≤ 0), but neither negatively stable nor positively asymptotically stable (since there are eigenvalues with 0 real part).

5.4 Linearization

If the right-hand side f of an autonomous ODE is differentiable and p is a fixed point (i.e. $f(p) = 0$), then one can sometimes use its linearization, i.e. its derivative $A := Df(p)$ (which is an $n \times n$ matrix), to infer stability properties of $y' = f(y)$ at p from those of $y' = Ay$ at 0 (see Th. 5.44 below). We start with some preparatory results:

Lemma 5.40. *Let $n \in \mathbb{N}$ and consider the bilinear function*

$$\beta : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \quad \beta(y,z) := y \bullet (Bz) = y^{\mathrm{t}} Bz = \sum_{k,l=1}^{n} y_k\, b_{kl}\, z_l, \tag{5.68}$$

where $B = (b_{kl}) \in \mathcal{M}(n, \mathbb{R})$, "$\bullet$" denotes the Euclidean scalar product, and elements of $\mathbb{R}^n$ are interpreted as column vectors when involved in matrix multiplications.

(a) *The function β is differentiable (it is even a polynomial,* $\deg(\beta) \leq 2$*, and, thus, C^∞) and*

$$\forall_{k\in\{1,\dots,n\}} \quad \forall_{(y,z)\in\mathbb{R}^n\times\mathbb{R}^n} \quad \begin{aligned} &\partial_{y_k}\beta : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \\ &\partial_{y_k}\beta(y,z) = \sum_{l=1}^{n} b_{kl}\, z_l = (Bz)_k, \end{aligned} \tag{5.69a}$$

$$\forall_{l\in\{1,\dots,n\}} \quad \forall_{(y,z)\in\mathbb{R}^n\times\mathbb{R}^n} \quad \begin{aligned} &\partial_{z_l}\beta : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \\ &\partial_{z_l}\beta(y,z) = \sum_{k=1}^{n} y_k\, b_{kl} = (y^{\mathrm{t}}B)_l, \end{aligned} \tag{5.69b}$$

$$\forall_{(y,z)\in\mathbb{R}^n\times\mathbb{R}^n} \quad \begin{aligned} D\beta(y,z) = \nabla\, \beta(y,z) &: \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \\ \nabla\, \beta(y,z)(u,v) &= \beta(y,v) + \beta(u,z) = y^{\mathrm{t}}Bv + u^{\mathrm{t}}Bz. \end{aligned} \tag{5.69c}$$

(b) *The function*

$$V : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad V(y) := \beta(y,y) = y \bullet (By) = y^{\mathrm{t}}By = \sum_{k,l=1}^{n} y_k\, b_{kl}\, y_l, \tag{5.70}$$

is differentiable (it is also even a polynomial, $\deg(\beta) \leq 2$*, and, thus,* C^∞*) and*

$$\forall_{k \in \{1,\dots,n\}} \quad \begin{aligned} &\partial_{y_k} V : \mathbb{R}^n \longrightarrow \mathbb{R}, \\ \forall_{y \in \mathbb{R}^n} \quad &\partial_{y_k} V(y) = \sum_{l=1}^{n} y_l (b_{kl} + b_{lk}) = y^{\mathrm{t}} (B + B^{\mathrm{t}})_k, \end{aligned} \tag{5.71a}$$

$$\forall_{y \in \mathbb{R}^n} \quad \begin{aligned} DV(y) = \nabla V(y) &: \mathbb{R}^n \longrightarrow \mathbb{R}, \\ \nabla V(y)(u) = \beta(y, u) + \beta(u, y) &= y^{\mathrm{t}} B u + u^{\mathrm{t}} B y = y^{\mathrm{t}} (B + B^{\mathrm{t}}) u. \end{aligned} \tag{5.71b}$$

Proof. (a): (5.69a) and (5.69b) are immediate from (5.68) and, then, imply (5.69c).
(b): (5.71a) is immediate from (5.70) and, then, implies (5.71b). ■

Lemma 5.41. *Let* $A, B \in \mathcal{M}(n, \mathbb{R})$*,* $n \in \mathbb{N}$*, and* $V : \mathbb{R}^n \longrightarrow \mathbb{R}$ *as in* (5.70)*. Then*

$$\forall_{y \in \mathbb{R}^n} \quad \nabla V(y) \bullet (Ay) = y^{\mathrm{t}} (BA + A^{\mathrm{t}} B) y. \tag{5.72}$$

Proof. We note

$$\forall_{y \in \mathbb{R}^n} \quad (y^{\mathrm{t}} B^{\mathrm{t}}) \bullet (Ay) = y^{\mathrm{t}} B^{\mathrm{t}} A y = \left(y^{\mathrm{t}} B^{\mathrm{t}} A y\right)^{\mathrm{t}} = y^{\mathrm{t}} A^{\mathrm{t}} B y \tag{5.73}$$

and, thus, obtain

$$\forall_{y \in \mathbb{R}^n} \quad \nabla V(y) \bullet (Ay) \overset{(5.71b)}{=} y^{\mathrm{t}} (B + B^{\mathrm{t}}) \bullet (Ay) \overset{(5.73)}{=} y^{\mathrm{t}} (BA + A^{\mathrm{t}} B) y, \tag{5.74}$$

proving (5.72). ■

Definition 5.42. A matrix $B \in \mathcal{M}(n, \mathbb{R})$, $n \in \mathbb{N}$, is called *positive definite* if, and only if, the function V of (5.70) is positive definite at $p = 0$ in the sense of Def. 5.29.

Proposition 5.43. *Let* $A \in \mathcal{M}(n, \mathbb{R})$*,* $n \in \mathbb{N}$*. Then the following statements* (i) – (iii) *are equivalent:*

(i) *There exist positive definite matrices* $B, C \in \mathcal{M}(n, \mathbb{R})$*, satisfying*

$$BA + A^{\mathrm{t}} B = -C. \tag{5.75}$$

(ii) $\operatorname{Re} \lambda < 0$ *holds for each eigenvalue* $\lambda \in \mathbb{C}$ *of* A*.*

(iii) *For each given positive definite (symmetric)* $C \in \mathcal{M}(n, \mathbb{R})$*, there exists a positive definite (symmetric)* $B \in \mathcal{M}(n, \mathbb{R})$*, satisfying* (5.75)*.*

Proof. (iii) immediately implies (i) (e.g. by applying (iii) with $C := \mathrm{Id}$).

For the proof that (i) implies (ii), let $B, C \in \mathcal{M}(n, \mathbb{R})$ be positive definite matrices, satisfying (5.75). By Th. 5.38(b), it suffices to show 0 is a positively asymptotically stable fixed point for $y' = Ay$. To this end, we apply Th. 5.30, using $V : \mathbb{R}^n \longrightarrow \mathbb{R}$ of

(5.70) as the Lyapunov function. Then, by Def. 5.42, B being positive definite means V being positive definite at 0. Since

$$\dot{V} : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad \dot{V}(y) = \nabla V(y) \bullet (Ay) \overset{(5.72)}{=} y^{\mathrm{t}}(BA + A^{\mathrm{t}}B)y \overset{(5.75)}{=} -y^{\mathrm{t}}Cy \tag{5.76}$$

and C is positive definite, $\dot{V}$ is negative definite at 0, i.e. Th. 5.30 yields 0 to be a positively asymptotically stable fixed point for $y' = Ay$ as desired.

It remains to show that (ii) implies (iii). If all eigenvalues of A have negative real part, then, as A and A^{t} have the same eigenvalues, all eigenvalues of A^{t} have negative real part as well. Thus, according to Th. 5.38(b),

$$\underset{K,\alpha>0}{\exists} \quad \underset{x\geq 0}{\forall} \quad \Big(\|e^{Ax}\|_{\max} \leq K e^{-\alpha x} \quad \wedge \quad \|e^{A^{\mathrm{t}}x}\|_{\max} \leq K e^{-\alpha x} \Big), \tag{5.77}$$

where we have chosen the norm in (5.77) to mean the max-norm on $\mathbb{R}^{n^2}$ (note that e^{Ax} is real if A is real, e.g. due to the series representation (4.73)). Given $C \in \mathcal{M}(n, \mathbb{R})$, define

$$B := \int_0^\infty e^{A^{\mathrm{t}}x}\, C\, e^{Ax}\, \mathrm{d}x\,. \tag{5.78}$$

To verify that $B \in \mathcal{M}(n, \mathbb{R})$ is well-defined, note that each entry of the integrand matrix of (5.78) constitutes an integrable function on $[0, \infty[$: Indeed,

$$\underset{M>0}{\exists} \quad \underset{x\geq 0}{\forall} \quad \begin{aligned} \big\|e^{A^{\mathrm{t}}x}\, C\, e^{Ax}\big\|_{\max} &\leq M\, \|e^{A^{\mathrm{t}}x}\|_{\max}\, \|C\|_{\max}\, \|e^{Ax}\|_{\max} \\ &\leq M\, \|C\|_{\max}\, K^2\, e^{-2\alpha x}, \end{aligned} \tag{5.79}$$

which is integrable on $[0, \infty[$. Next, we compute

$$\begin{aligned} -C \quad &\overset{(5.79)}{=} \quad \lim_{x\to\infty} e^{A^{\mathrm{t}}x}\, C\, e^{Ax} - C = \lim_{x\to\infty} \int_0^x \partial_s \big(e^{A^{\mathrm{t}}s}\, C\, e^{As}\big)\, \mathrm{d}s \\ &= \quad \int_0^\infty \partial_s \big(e^{A^{\mathrm{t}}s}\, C\, e^{As}\big)\, \mathrm{d}s \\ &\overset{(I.3)}{=} \quad \int_0^\infty \big(A^{\mathrm{t}}\, e^{A^{\mathrm{t}}s}\, C\, e^{As} + e^{A^{\mathrm{t}}s}\, C\, e^{As}\, A\big)\, \mathrm{d}s \\ &\overset{(I.5),(I.6)}{=} \quad A^{\mathrm{t}} \int_0^\infty e^{A^{\mathrm{t}}s}\, C\, e^{As}\, \mathrm{d}s + \left(\int_0^\infty e^{A^{\mathrm{t}}s}\, C\, e^{As}\, \mathrm{d}s \right) A \\ &\overset{(5.78)}{=} \quad = A^{\mathrm{t}}B + BA, \end{aligned} \tag{5.80}$$

showing (5.75) is satisfied. If C is positive definite and $0 \neq y \in \mathbb{R}^n$, then $y^{\mathrm{t}}Cy > 0$, implying

$$\begin{aligned} y^{\mathrm{t}}By \quad &= \quad y^{\mathrm{t}} \left(\int_0^\infty e^{A^{\mathrm{t}}x}\, C\, e^{Ax}\, \mathrm{d}x \right) y \overset{(I.5),(I.6)}{=} \int_0^\infty y^{\mathrm{t}}\, e^{A^{\mathrm{t}}x}\, C\, e^{Ax}\, y\, \mathrm{d}x \\ &\overset{\text{Prop. 4.40(c)}}{=} \int_0^\infty (e^{Ax}\, y)^{\mathrm{t}}\, C\, e^{Ax}\, y\, \mathrm{d}x \, > 0, \end{aligned} \tag{5.81}$$

showing B is positive definite as well. Finally, if C is symmetric, then

$$B^{\mathrm{t}} = \int_0^\infty \left(e^{A^{\mathrm{t}}x}\, C\, e^{Ax}\right)^{\mathrm{t}} \,\mathrm{d}x \overset{\text{Prop. 4.40(c)}}{=} \int_0^\infty e^{A^{\mathrm{t}}x}\, C\, e^{Ax} \,\mathrm{d}x = B, \tag{5.82}$$

showing B is symmetric as well. ∎

Theorem 5.44. *Let $\Omega \subseteq \mathbb{R}^n$ be open, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{R}^n$ continuously differentiable. Let $p \in \Omega$ be a fixed point (i.e. $f(p) = 0$) and $A := Df(p) \in \mathcal{M}(n, \mathbb{R})$ the derivative of f at p. If all eigenvalues of A have negative (resp. positive) real parts, then p is a positively (resp. negatively) asymptotically stable fixed point for $y' = f(y)$.*

Proof. Let all eigenvalues of A have negative real parts. We first consider the special case $p = 0$, i.e. $A = Df(0)$. By the equivalence between (ii) and (iii) of Prop. 5.43, we can choose $C := \mathrm{Id}$ in (iii) to obtain the existence of a positive definite symmetric matrix $B \in \mathcal{M}(n, \mathbb{R})$, satisfying

$$BA + A^{\mathrm{t}}B = -\,\mathrm{Id}\,. \tag{5.83}$$

The idea is now to apply the Lyapunov Th. 5.30 with V of (5.70), i.e.

$$V : \Omega \longrightarrow \mathbb{R}, \quad V(y) := y \bullet (By) = y^{\mathrm{t}}By = \sum_{k,l=1}^{n} y_k\, b_{kl}\, y_l. \tag{5.84}$$

We already know V to be continuously differentiable and positive definite. We will conclude the proof of 0 being positively asymptotically stable by showing there exists $\delta > 0$, such that

$$\dot V : B_\delta(0) \longrightarrow \mathbb{R}, \quad \dot V(y) = (\nabla V)(y) \bullet f(y), \tag{5.85}$$

is negative definite at 0, where we take $B_\delta(0)$ with respect to the 2-norm $\|\cdot\|_2$ on $\mathbb{R}^n$. The differentiability of f at 0 implies that (cf. [Phi13b, Lem. 2.21])

$$r : \Omega \longrightarrow \mathbb{R}^n, \quad r(y) := f(y) - Ay, \tag{5.86}$$

satisfies

$$\lim_{y\to 0} \frac{\|r(y)\|_2}{\|y\|_2} = 0. \tag{5.87}$$

Thus, we compute, for each $y \in \Omega$,

$$\begin{aligned} \dot V(y) &= (\nabla V)(y) \bullet f(y) \overset{(5.71b),(5.86)}{=} (\nabla V)(y) \bullet (Ay) + \big(y^{\mathrm{t}}(B + B^{\mathrm{t}})\big) \bullet r(y) \\ &\overset{(5.72),\, B=B^{\mathrm{t}}}{=} y^{\mathrm{t}}(BA + A^{\mathrm{t}}B)y + 2y^{\mathrm{t}}B\, r(y) \overset{(5.83)}{=} -\|y\|_2^2 + 2\, y \bullet B\, r(y). \end{aligned} \tag{5.88}$$

We can estimate the second summand via the Cauchy-Schwarz inequality to obtain

$$\big|y \bullet B\, r(y)\big| \le \|y\|_2\, \|B\, r(y)\|_2 \le \|y\|_2\, \|B\|\, \|r(y)\|_2, \tag{5.89}$$

and, thus, using (5.87),

$$\lim_{y\to 0} \frac{\big|y \bullet B\, r(y)\big|}{\|y\|_2^2} = 0. \tag{5.90}$$

Now choose $\delta > 0$ such that $B_\delta(0) \subseteq \Omega$ and such that

$$\underset{y \in B_\delta(0)}{\forall} \quad \frac{2\big|y \bullet B\,r(y)\big|}{\|y\|_2^2} < \frac{1}{2}. \tag{5.91}$$

Then, for each $0 \neq y \in B_\delta(0)$

$$\dot{V}(y) = -\|y\|_2^2 + 2\,y \bullet B\,r(y) \overset{(5.91)}{<} -\|y\|_2^2 + \frac{\|y\|_2^2}{2} = -\frac{\|y\|_2^2}{2} < 0, \tag{5.92}$$

showing $\dot{V}$ to be negative definite at 0, and 0 to be positively asymptotically stable. If $p \neq 0$, then consider the ODE $y' = g(y) := f(y+p)$, $g : (\Omega - p) \longrightarrow \mathbb{R}^n$. Then 0 is a fixed point for $y' = g(y)$, $Dg(0) = Df(p) = A$, i.e. 0 is positively asymptotically stable for $y' = g(y)$. But, since ψ is a solution to $y' = g(y)$ if, and only if, $\phi = \psi + p$ is a solution to $y' = f(y)$, p must be positively asymptotically stable for $y' = f(y)$. The remaining case that all eigenvalues of A have positive real parts is now treated via time reversion: If all eigenvalues of A have positive real parts, then all eigenvalues of $-A = D(-f)(p)$ have negative real parts, i.e. p is positively asymptotically stable for $y' = -f(y)$, i.e., by Rem. 5.25(b), p is negatively asymptotically stable for $y' = f(y)$. ∎

Caveat 5.45. The following example shows that the converse of Th. 5.44 does not hold: A fixed point p can be positively (resp. negatively) asymptotically stable without $A := Df(p)$ having only eigenvalues with negative (resp. positive) real parts. The same example shows that, in general, one can not infer anything regarding the stability of the fixed point p if $A := Df(p)$ is merely stable, but not asymptotically stable: Consider

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad f(y_1, y_2) := (y_2 + \mu y_1^3,\ -y_1 + \mu y_2^3), \quad \mu \in \mathbb{R}. \tag{5.93}$$

Then, independently of μ, $(0,0)$ is a fixed point and

$$Df(0,0) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{5.94}$$

with complex eigenvalues i and $-i$. Thus, the linearized system is positively and negatively stable, but not asymptotically stable, still independently of μ. However, we claim that $(0,0)$ is a positively asymptotically stable fixed point for $y' = f(y)$ if $\mu < 0$ and a positively asymptotically stable fixed point for $y' = f(y)$ if $\mu > 0$. Indeed, this can be seen by using the Lyapunov function $V : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $V(y_1, y_2) = y_1^2 + y_2^2$, which has $\nabla V(y_1, y_2) = (2y_1,\ 2y_2)$ and

$$\dot{V}(y_1, y_2) = \nabla V(y_1, y_2) \bullet f(y_1, y_2) = 2\mu\,(y_1^4 + y_2^4). \tag{5.95}$$

Thus, V is positive definite at $(0,0)$ and $\dot{V}$ is negative definite at $(0,0)$ for $\mu < 0$ and positive definite at $(0,0)$ for $\mu > 0$.

Example 5.46. Consider $(x, y, z)' = f(x, y, z)$ with

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \quad f(x, y, z) = (-x\cos y,\ -ye^z,\ x^2 - 2z). \tag{5.96}$$

The derivative is

$$Df : \mathbb{R}^3 \longrightarrow \mathcal{M}(3,\mathbb{R}), \quad Df(x,y,z) = \begin{pmatrix} -\cos y & x\sin y & 0 \\ 0 & -e^z & -ye^z \\ 2x & 0 & -2 \end{pmatrix}. \tag{5.97}$$

Clearly, $(0,0,0)$ is a fixed point and $Df(0,0,0)$ has eigenvalues -1 and -2. Thus, $(0,0,0)$ is a positively asymptotically stable fixed point for $(x,y,z)' = f(x,y,z)$ by Th. 5.44.

5.5 Limit Sets

Limit sets are important when studying the asymptotic behavior of solutions, i.e. $\phi(x)$ for $x \to \infty$ and for $x \to -\infty$. If a solution has a limit, then its corresponding limit set consists of precisely one point. In general, the limit set of a solution is defined to consist of all points that occur as limits of sequences taken along the solution's orbit (of course, the limit sets can also be empty):

Definition 5.47. Let $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{K}^n$ be such that $y' = f(y)$ admits unique maximal solutions. For each $\eta \in \Omega$, we define the *omega limit set* and the *alpha limit set* of η as follows:

$$\omega(\eta) := \omega_f(\eta) := \left\{ y \in \overline{\Omega} : \underset{(x_k)_{k\in\mathbb{N}}\subseteq\mathbb{R}}{\exists} \quad \lim_{k\to\infty} x_k = \infty \ \wedge \ \lim_{k\to\infty} Y(x_k,\eta) = y \right\}, \tag{5.98a}$$

$$\alpha(\eta) := \alpha_f(\eta) := \left\{ y \in \overline{\Omega} : \underset{(x_k)_{k\in\mathbb{N}}\subseteq\mathbb{R}}{\exists} \quad \lim_{k\to\infty} x_k = -\infty \ \wedge \ \lim_{k\to\infty} Y(x_k,\eta) = y \right\}. \tag{5.98b}$$

Remark 5.48. In the situation of Def. 5.47, consider the time-reversed version of $y' = f(y)$, i.e. $y' = -f(y)$, with its general solution $\tilde{Y}(x,\eta) = Y(-x,\eta)$, cf. (5.28). Clearly, for each $\eta \in \Omega$,

$$\omega_f(\eta) = \alpha_{-f}(\eta), \quad \alpha_f(\eta) = \omega_{-f}(\eta). \tag{5.99}$$

Proposition 5.49. *In the situation of Def. 5.47, the following hold:*

(a) *If $Y(\cdot,\eta)$ is defined on all of $\mathbb{R}_0^+$, then*

$$\omega(\eta) = \bigcap_{m=0}^{\infty} \overline{\{Y(x,\eta) : x \geq m\}}; \tag{5.100a}$$

and if $Y(\cdot,\eta)$ is defined on all of $\mathbb{R}_0^-$, then

$$\alpha(\eta) = \bigcap_{m=0}^{\infty} \overline{\{Y(x,\eta) : x \leq -m\}}. \tag{5.100b}$$

(b) *All points in the same orbit have the same omega and alpha limit sets, i.e.*

$$\underset{x\in I_{0,\eta}}{\forall} \Big(\omega(\eta) = \omega\big(Y(x,\eta)\big) \quad \wedge \quad \alpha(\eta) = \alpha\big(Y(x,\eta)\big) \Big).$$

Proof. Due to Rem. 5.48, it suffices to prove the statements involving the omega limit sets.

(a): Let $y \in \omega(\eta)$ and $m \in \mathbb{N}_0$. Then there is a sequence $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}$ such that $\lim_{k\to\infty} x_k = \infty$ and $\lim_{k\to\infty} Y(x_k,\eta) = y$. Since, for sufficiently large $k_0 \in \mathbb{N}$, the sequence $(Y(x_k,\eta))_{k\geq k_0}$ is in $\overline{\{Y(x,\eta) : x \geq m\}}$, the inclusion "$\subseteq$" of (5.100a) is proved. Conversely, assume $y \in \overline{\{Y(x,\eta) : x \geq m\}}$. Then,

$$\underset{k\in\mathbb{N}}{\forall} \quad \underset{x_k\in[k,\infty[}{\exists} \quad \big\|Y(x_k,\eta) - y\big\| < \frac{1}{k},$$

providing a sequence $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}$ such that $\lim_{k\to\infty} x_k = \infty$ and $\lim_{k\to\infty} Y(x_k,\eta) = y$, proving $y \in \omega(\eta)$ and the inclusion "$\supseteq$" of (5.100a).

(b): Let $y \in \omega(\eta)$ and $x \in I_{0,\eta}$. Choose a sequence $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}$ such that $\lim_{k\to\infty} x_k = \infty$ and $\lim_{k\to\infty} Y(x_k,\eta) = y$. Then $\lim_{k\to\infty}(x_k - x) = \infty$ and

$$\lim_{k\to\infty} Y\big(x_k - x, Y(x,\eta)\big) \overset{\text{Lem. 5.4(b)}}{=} \lim_{k\to\infty} Y(x_k,\eta) = y, \tag{5.101}$$

proving $\omega(\eta) \subseteq \omega\big(Y(x,\eta)\big)$. The reversed inclusion then also follows, since

$$Y\big(-x, Y(x,\eta)\big) \overset{\text{Lem. 5.4(b)}}{=} Y(0,\eta) = \eta, \tag{5.102}$$

concluding the proof. ∎

Example 5.50. Let $\Omega \subseteq \mathbb{K}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{K}^n$ be such that $y' = f(y)$ admits unique maximal solutions.

(a) If $\eta \in \Omega$ is a fixed point, then $\omega(\eta) = \alpha(\eta) = \{\eta\}$. More generally, if $\eta \in \Omega$ is such that $\lim_{x\to\infty} Y(x,\eta) = y \in \mathbb{K}^n$, then $\omega(\eta) = \{y\}$; if $\eta \in \Omega$ is such that $\lim_{x\to-\infty} Y(x,\eta) = y \in \mathbb{K}^n$, then $\alpha(\eta) = \{y\}$.

(b) If $A \in \mathcal{M}(n,\mathbb{C})$ is such that the conditions of Th. 5.38(b) hold (all eigenvalues have negative real parts, 0 is positively asymptotically stable), then $\omega(\eta) = \{0\}$ for each $\eta \in \mathbb{C}^n$ and $\alpha(\eta) = \emptyset$ for each $\eta \in \mathbb{C}^n \setminus \{0\}$.

(c) If $\eta \in \Omega$ is such that the orbit $\mathcal{O}(\phi)$ of $\phi := Y(\cdot,\eta)$ is periodic, then $\omega(\eta) = \alpha(\eta) = \mathcal{O}(\phi)$. For example, for (5.8),

$$\underset{\eta\in\mathbb{R}^2}{\forall} \quad \omega(\eta) = \alpha(\eta) = S_{\|\eta\|_2}(0) = \big\{y \in \mathbb{R}^2 : \|y\|_2 = \|\eta\|_2\big\}. \tag{5.103}$$

Example 5.51. As an example with nonperiodic orbits that have limit sets consisting of more than one point, consider

$$y_1' = y_2 + y_1(1 - y_1^2 - y_2^2), \tag{5.104a}$$
$$y_2' = -y_1 + y_2(1 - y_1^2 - y_2^2). \tag{5.104b}$$

We will show that, for each point except the origin (which is clearly a fixed point), the omega limit set is the unit circle, i.e.

$$\underset{\eta\in\mathbb{R}^2\setminus\{0\}}{\forall} \quad \omega(\eta) = S_1(0) = \{y \in \mathbb{R}^2 : \, y_1^2 + y_2^2 = 1\}. \tag{5.105}$$

We first verify that the general solution is

$$Y : \, D_{f,0} \longrightarrow \mathbb{R}^2, \quad Y(x, \eta_1, \eta_2) = \frac{(\eta_1 \cos x + \eta_2 \sin x, \; \eta_2 \cos x - \eta_1 \sin x)}{\sqrt{\eta_1^2 + \eta_2^2 + (1 - \eta_1^2 - \eta_2^2)e^{-2x}}}, \tag{5.106}$$

where, letting

$$\underset{\eta\in\{y\in\mathbb{R}^2:\,\|y\|_2>1\}}{\forall} \quad x_\eta := \frac{1}{2} \ln\left(\frac{\|\eta\|_2^2 - 1}{\|\eta\|_2^2}\right), \tag{5.107}$$

$$D_{f,0} = \Big(\mathbb{R} \times \{\eta \in \mathbb{R}^2 : \, \|\eta\|_2 \le 1\}\Big) \cup \Big(]x_\eta, \infty[\times \{\eta \in \mathbb{R}^2 : \, \|\eta\|_2 > 1\}\Big) : \tag{5.108}$$

For each $(\eta_1, \eta_2) \in \mathbb{R}^2$, $Y(\cdot, \eta_1, \eta_2)$ satisfies the initial condition:

$$Y(0, \eta_1, \eta_2) = \frac{(\eta_1, \, \eta_2)}{\sqrt{\eta_1^2 + \eta_2^2 + (1 - \eta_1^2 - \eta_2^2)}} = (\eta_1, \, \eta_2). \tag{5.109}$$

The following computations prepare the check that each $Y(\cdot, \eta_1, \eta_2)$ satisfies (5.104): The 2-norm squared of the numerator in (5.106) is

$$\begin{aligned} &\big\|(\eta_1 \cos x + \eta_2 \sin x, \; \eta_2 \cos x - \eta_1 \sin x)\big\|_2^2 \\ &= \eta_1^2 \cos^2 x + 2\eta_1\eta_2 \cos x \sin x + \eta_2^2 \sin^2 x + \eta_2^2 \cos^2 x - 2\eta_1\eta_2 \cos x \sin x + \eta_1^2 \sin^2 x \\ &= \eta_1^2 + \eta_2^2 = \|\eta\|_2^2. \end{aligned} \tag{5.110}$$

Thus,

$$\big\|Y(x, \eta_1, \eta_2)\big\|_2 = \frac{\|\eta\|_2}{\sqrt{\|\eta\|_2^2 + (1 - \|\eta\|_2^2)e^{-2x}}} \tag{5.111}$$

and

$$\begin{aligned} 1 - Y_1^2(x, \eta_1, \eta_2) - Y_2^2(x, \eta_1, \eta_2) = 1 - \big\|Y(x, \eta_1, \eta_2)\big\|_2^2 &= \frac{\|\eta\|_2^2 + (1 - \|\eta\|_2^2)e^{-2x} - \|\eta\|_2^2}{\|\eta\|_2^2 + (1 - \|\eta\|_2^2)e^{-2x}} \\ &= \frac{(1 - \|\eta\|_2^2)e^{-2x}}{\|\eta\|_2^2 + (1 - \|\eta\|_2^2)e^{-2x}}. \end{aligned} \tag{5.112}$$

In consequence,

$$\begin{aligned} &Y_1'(x, \eta_1, \eta_2) \\ &= \frac{(-\eta_1 \sin x + \eta_2 \cos x)\big(\|\eta\|_2^2 + (1 - \|\eta\|_2^2)e^{-2x}\big) + (\eta_1 \cos x + \eta_2 \sin x)(1 - \|\eta\|_2^2)e^{-2x}}{\big(\|\eta\|_2^2 + (1 - \|\eta\|_2^2)e^{-2x}\big)^{\frac{3}{2}}} \\ &= Y_2(x, \eta_1, \eta_2) + Y_1(x, \eta_1, \eta_2)\Big(1 - Y_1^2(x, \eta_1, \eta_2) - Y_2^2(x, \eta_1, \eta_2)\Big), \end{aligned} \tag{5.113}$$

verifying (5.104a). Similarly,

$$
\begin{aligned}
&Y_2'(x,\eta_1,\eta_2)\\
&= \frac{(-\eta_2\sin x-\eta_1\cos x)\big(\|\eta\|_2^2+(1-\|\eta\|_2^2)e^{-2x}\big)+(\eta_2\cos x-\eta_1\sin x)(1-\|\eta\|_2^2)e^{-2x}}{\big(\|\eta\|_2^2+(1-\|\eta\|_2^2)e^{-2x}\big)^{\frac{3}{2}}}\\
&= -Y_1(x,\eta_1,\eta_2)+Y_2(x,\eta_1,\eta_2)\Big(1-Y_1^2(x,\eta_1,\eta_2)-Y_2^2(x,\eta_1,\eta_2)\Big), \qquad (5.114)
\end{aligned}
$$

verifying (5.104b).

For $\|\eta\|_2 \le 1$, $Y(\cdot,\eta_1,\eta_2)$ is maximal, as it is defined on $\mathbb{R}$ (the denominator in (5.106) has no zero in this case). For $\|\eta\|_2 > 1$, the denominator clearly has a zero at $x_\eta < 0$, where x_η is defined as in (5.107). For $x > x_\eta$, the expression under the square root in (5.106) is positive. Since $\lim_{x\downarrow x_\eta}\|Y(x,\eta_1,\eta_2)\|_2 = \infty$ for $\|\eta\|_2 > 1$, $Y(\cdot,\eta_1,\eta_2)$ is maximal in this case as well, completing the verification of Y, defined as in (5.106) – (5.108), being the general solution of (5.104).

It remains to prove (5.105). From (5.111), we obtain

$$\underset{\eta\in\mathbb{R}^2\setminus\{0\}}{\forall}\quad \lim_{x\to\infty}\big\|Y(x,\eta_1,\eta_2)\big\|_2 = 1, \qquad (5.115)$$

which implies

$$\underset{\eta\in\mathbb{R}^2\setminus\{0\}}{\forall}\quad \omega(\eta)\subseteq S_1(0). \qquad (5.116)$$

Conversely, consider $\eta=(\eta_1,\eta_2)\in\mathbb{R}^2\setminus\{0\}$ and $y=(y_1,y_2)\in S_1(0)$. We will show $y\in\omega(\eta)$: Since $\|y\|_2=1$,

$$\underset{\varphi_y\in[0,2\pi[}{\exists}\quad y=(\sin\varphi_y,\cos\varphi_y). \qquad (5.117)$$

Analogously,

$$\underset{\varphi_\eta\in[0,2\pi[}{\exists}\quad \eta=\|\eta\|_2\,(\sin\varphi_\eta,\cos\varphi_\eta) \qquad (5.118)$$

(the reader might note that, in (5.117) and (5.118), we have written y and η using their *polar coordinates*, cf. [Phi13b, Ex. 4.19]). Then, according to (5.106), we obtain, for each $x\ge 0$,

$$
\begin{aligned}
Y(x,\eta_1,\eta_2) &= \frac{\|\eta\|_2\,(\sin\varphi_\eta\cos x+\cos\varphi_\eta\sin x,\ \cos\varphi_\eta\cos x-\sin\varphi_\eta\sin x)}{\sqrt{\|\eta\|_2^2+(1-\|\eta\|_2^2)e^{-2x}}}\\
&= \frac{\|\eta\|_2\,\big(\sin(x+\varphi_\eta),\ \cos(x+\varphi_\eta)\big)}{\sqrt{\|\eta\|_2^2+(1-\|\eta\|_2^2)e^{-2x}}}. \qquad (5.119)
\end{aligned}
$$

Define

$$\underset{k\in\mathbb{N}}{\forall}\quad x_k := \varphi_y-\varphi_\eta+2\pi k\in\mathbb{R}^+. \qquad (5.120)$$

Then $\lim_{k\to\infty} x_k = \infty$ and

$$\begin{aligned}
&\lim_{k\to\infty} Y(x_k,\eta_1,\eta_2) \\
&\overset{(5.119)}{=} \lim_{k\to\infty} \frac{\|\eta\|_2 \big(\sin(\varphi_y - \varphi_\eta + 2\pi k + \varphi_\eta),\, \cos(\varphi_y - \varphi_\eta + 2\pi k + \varphi_\eta)\big)}{\sqrt{\|\eta\|_2^2 + (1-\|\eta\|_2^2)e^{-2x}}} \\
&= \lim_{k\to\infty} \frac{\|\eta\|_2 (\sin\varphi_y, \cos\varphi_y)}{\sqrt{\|\eta\|_2^2 + (1-\|\eta\|_2^2)e^{-2x}}} = y, \qquad (5.121)
\end{aligned}$$

showing $y \in \omega(\eta)$ and $S_1(0) \subseteq \omega(\eta)$.

Proposition 5.52. *In the situation of Def.* 5.47, *if f is locally Lipschitz, then orbits that intersect an omega or alpha limit set, must entirely remain inside that same omega or alpha limit set, i.e.*

$$\left(\underset{y\in\Omega\cap\omega(\eta)}{\forall} \quad \underset{x\in I_{0,y}}{\forall} \quad Y(x,y) \in \omega(\eta) \right) \quad \wedge \quad \left(\underset{y\in\Omega\cap\alpha(\eta)}{\forall} \quad \underset{x\in I_{0,y}}{\forall} \quad Y(x,y) \in \alpha(\eta) \right). \tag{5.122}$$

Proof. Due to Rem. 5.48, it suffices to prove the statement involving the omega limit set. Let $y \in \Omega\cap\omega(\eta)$ and $x \in I_{0,y}$. Choose a sequence $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}$ such that $\lim_{k\to\infty} x_k = \infty$ and $\lim_{k\to\infty} Y(x_k,\eta) = y$. Then $\lim_{k\to\infty}(x_k + x) = \infty$ and,

$$\lim_{k\to\infty} Y(x_k + x, \eta) \overset{\text{Lem. 5.4(b)}}{=} \lim_{k\to\infty} Y\big(x, Y(x_k,\eta)\big) \overset{(*)}{=} Y(x,y), \tag{5.123}$$

proving $Y(x,y) \in \omega(\eta)$. At "$(*)$", we have used that, due to f being locally Lipschitz by hypothesis, Y is continuous by Th. 3.35. ∎

Proposition 5.53. *In the situation of Def.* 5.47, *let $\eta \in \Omega$ be such that there exists a compact set $K \subseteq \Omega$, satisfying*

$$\{Y(x,\eta) : x \geq 0\} \subseteq K \quad \big(\textit{resp. } \{Y(x,\eta) : x \leq 0\} \subseteq K\big). \tag{5.124}$$

Then the following hold:

(a) *$\omega(\eta) \neq \emptyset$ (resp. $\alpha(\eta) \neq \emptyset$).*

(b) *$\omega(\eta)$ (resp. $\alpha(\eta)$) is compact.*

(c) *$\omega(\eta)$ (resp. $\alpha(\eta)$) is a connected set, i.e. if O_1, O_2 are disjoint open subsets of $\mathbb{K}^n$ such that $\omega(\eta) \subseteq O_1 \cup O_2$ (resp. $\alpha(\eta) \subseteq O_1 \cup O_2$), then $\omega(\eta) \cap O_1 = \emptyset$ or $\omega(\eta) \cap O_2 = \emptyset$ (resp. $\alpha(\eta) \cap O_1 = \emptyset$ or $\alpha(\eta) \cap O_2 = \emptyset$).*

Proof. Due to Rem. 5.48, it suffices to prove the statements involving the omega limit sets.

(a): Since, by hypothesis, $(Y(k,\eta))_{k\in\mathbb{N}}$ is a sequence in the compact set K, it must have a subsequence, converging to some limit $y \in K$. But then $y \in \omega(\eta)$, i.e. $\omega(\eta) \neq \emptyset$.

(b): According to (5.100a) and (5.124), $\omega(\eta)$ is a closed subset of the compact set K, implying $\omega(\eta)$ to be compact as well.

(c): Seeking a contradiction, we suppose the assertion is false, i.e. there are disjoint open subsets O_1, O_2 of $\mathbb{K}^n$ such that $\omega(\eta) \subseteq O_1 \cup O_2$, $\omega_1 := \omega(\eta) \cap O_1 \neq \emptyset$ and $\omega_2 := \omega(\eta) \cap O_2 \neq \emptyset$. Then ω_1 and ω_2 are disjoint since O_1, O_2 are disjoint. Moreover, ω_1 and ω_2 are both subsets of the compact set $\omega(\eta)$. Due to $\omega_1 = \omega(\eta) \cap (\mathbb{K}^n \setminus O_2)$ and $\omega_2 = \omega(\eta) \cap (\mathbb{K}^n \setminus O_1)$, ω_1 and ω_2 are also closed, hence, compact. Then, according to Prop. C.10, $\delta := \operatorname{dist}(\omega_1, \omega_2) > 0$. If $y_1 \in \omega_1$ and $y_2 \in \omega_2$, then there are numbers $0 < s1 < t_1 < s_2 < t_2 < \dots$ such that $\lim_{k\to\infty} s_k = \lim_{k\to\infty} t_k = \infty$ and

$$\forall_{k\in\mathbb{N}} \quad \Big(Y(s_k,\eta) \in O_1 \ \wedge Y(t_k,\eta) \in O_2\Big). \tag{5.125}$$

Define

$$\forall_{k\in\mathbb{N}} \quad \sigma_k := \sup\big\{x \geq s_k : \, Y(t,\eta) \in O_1 \text{ for each } t \in [s_k, x]\big\}. \tag{5.126}$$

Then $s_k < \sigma_k < t_k$ and the continuity of $Y(\cdot,\eta)$ yields $\eta_k := Y(\sigma_k,\eta) \in \partial O_1$. Thus $(\eta_k)_{k\in\mathbb{N}}$ is a sequence in the compact set $K \cap \partial O_1$ and, therefore, must have a convergent subsequence, converging to some $z \in K \cap \partial O_1$. But then $z \in \omega(\eta)$, but not in $O_1 \cup O_2$, in contradiction to $\omega(\eta) \subseteq O_1 \cup O_2$. ■

Theorem 5.54 (LaSalle). *Let $\Omega \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, and $f : \Omega \longrightarrow \mathbb{R}^n$ be such that $y' = f(y)$ admits unique maximal solutions. Moreover, let Ω_0 be an open subset of Ω, assume $V : \Omega_0 \longrightarrow \mathbb{R}$ is continuously differentiable, $K := \{y \in \Omega_0 : V(y) \leq r\}$ is compact for some $r \in \mathbb{R}$, and $\dot V(y) \leq 0$ (resp. $\dot V(y) \geq 0$) for each $y \in K$, where $\dot V$ is defined as in* (5.33)*. If $\eta \in \Omega_0$ is such that $V(\eta) < r$, then the following hold:*

(a) *$Y(\cdot,\eta)$ is defined on all of $\mathbb{R}_0^+$ (resp. on all of $\mathbb{R}_0^-$).*

(b) *One has $\omega(\eta) \subseteq K$ (resp. $\alpha(\eta) \subseteq K$) and V is constant on $\omega(\eta)$ (resp. on $\alpha(\eta)$).*

(c) *If f is locally Lipschitz, then, letting*

$$M := \Big\{y \in K : \, \dot V\big(Y(x,y)\big) = 0 \textit{ for each } x \geq 0 \textit{ (resp. for each } x \leq 0)\Big\},$$

one has $\omega(\eta) \subseteq M$ (resp. $\alpha(\eta) \subseteq M$). In particular, $\dot V(y) = 0$ for each $y \in \omega(\eta)$ (resp. for each $y \in \alpha(\eta)$).

Proof. As usual, it suffices to prove the assertions for $\dot V(y) \leq 0$, as the assertions for $\dot V(y) \geq 0$ then follow via time reversion.

(a): We claim

$$\forall_{x\in I_{0,\eta}\cap\mathbb{R}_0^+} \quad V\big(Y(x,\eta)\big) < r : \tag{5.127}$$

Indeed, if (5.127) does not hold, then let

$$0 < s := \sup\big\{x \geq 0 : \, V\big(Y(x,\eta)\big) < r \text{ for each } t \in [0,x]\big\} \in I_{0,\eta}, \tag{5.128}$$

and

$$r = V\big(Y(s,\eta)\big) = V(\eta) + \int_0^s \dot{V}\big(Y(t,\eta)\big)\,\mathrm{d}t \le V(\eta) < r, \tag{5.129}$$

which is impossible. Thus, (5.127) must hold. However, (5.127) implies $\mathbb{R}_0^+ \subseteq I_{0,\eta}$, since $Y(\cdot,\eta)$ is a maximal solution and $K = \{y \in \Omega_0 : V(y) \le r\}$ is compact.

(b): Let $\phi := Y(\cdot,\eta)$. During the proof of (a) above, we have shown $\phi(x) \in K$ for each $x \ge 0$. Since, then, $(V \circ \phi)'(x) = \dot{V}(\phi(x)) \le 0$ for each $x \ge 0$, $V \circ \phi$ is nonincreasing for $x \ge 0$. Since $V \circ \phi$ is also bounded on K,

$$\underset{c\in\mathbb{R}}{\exists} \quad c = \lim_{x\to\infty} V\big(\phi(x)\big). \tag{5.130}$$

If $y \in \omega(\eta)$, then there exists a sequence $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}$ such that $\lim_{k\to\infty} x_k = \infty$ and $\lim_{k\to\infty} \phi(x_k) = y$. Thus, $y \in K$ (since K is closed), and

$$V(y) = \lim_{k\to\infty} V\big(\phi(x_k)\big) \overset{(5.130)}{=} c, \tag{5.131}$$

proving (b).

(c): Let $y \in \omega(\eta)$ and $\phi := Y(\cdot,y)$. Since f is assumed to be locally Lipschitz, Prop. 5.52 applies and we obtain $\phi(x) = Y(x,y) \in \omega(\eta)$ for each $x \in \mathbb{R}_0^+$. Using (b), we know V to be constant on $\omega(\eta)$, i.e. $V \circ \phi$ must be constant on $\mathbb{R}_0^+$ as well, implying

$$\underset{x\in\mathbb{R}_0^+}{\forall} \quad \dot{V}(\phi(x)) = (V\circ\phi)'(x) = 0 \tag{5.132}$$

as claimed. ∎

Example 5.55. Let $a < 0 < b$ and let $h :]a,b[\longrightarrow \mathbb{R}$ be continuously differentiable and such that

$$h(x) \begin{cases} < 0 & \text{for } x < 0, \\ = 0 & \text{for } x = 0, \\ > 0 & \text{for } x > 0. \end{cases} \tag{5.133}$$

Consider the autonomous ODE

$$y_1' = y_2, \tag{5.134a}$$
$$y_2' = -y_1^2\, y_2 - h(y_1). \tag{5.134b}$$

The right-hand side is defined on $\Omega :=]a,b[\times\mathbb{R}$ and is clearly C^1, i.e. the ODE admits unique maximal solutions. Due to (5.133), $\mathcal{F} = \{(0,0)\}$, i.e. the origin is the only fixed point of (5.134). We will use Th. 5.54(c) to show $(0,0)$ is positively asymptotically stable: We introduce

$$H :]a,b[\longrightarrow \mathbb{R}, \quad H(x) := \int_0^x h(t)\,\mathrm{d}t\,, \tag{5.135}$$

and the Lyapunov function

$$V : \Omega \longrightarrow \mathbb{R}, \quad V(y_1,y_2) := H(y_1) + \frac{y_2^2}{2}. \tag{5.136}$$

Since H is positive definite at 0 (H is actually strictly decreasing on $]a,0]$ and strictly increasing on $[0,b[$), V is positive definite at $(0,0)$. We also obtain

$$\dot{V}: \Omega \longrightarrow \mathbb{R}, \quad \dot{V}(y_1,y_2) = \big(h(y_1),y_2\big) \bullet \big(y_2, -y_1^2\,y_2 - h(y_1)\big) = -y_1^2 y_2^2 \le 0. \tag{5.137}$$

Thus, from the Lyapunov Th. 5.30, we already know $(0,0)$ to be positively stable. However, $\dot{V}$ is not negative definite at $(0,0)$, i.e. we can not immediately conclude that $(0,0)$ is positively asymptotically stable. Instead, as promised, we apply Th. 5.54(c): To this end, using that H is continuous and positive definite at 0, we choose $r>0$ and $c,d \in \mathbb{R}$, satisfying

$$a < c < 0 < d < b \quad \text{and} \quad H(c) = H(d) = r, \tag{5.138}$$

and define

$$O := \{(y_1,y_2) \in \Omega : V(y_1,y_2) < r\}, \tag{5.139}$$

$$K := \{(y_1,y_2) \in \Omega : V(y_1,y_2) \le r\}. \tag{5.140}$$

Then O is open since V is continuous, and it suffices to show

$$\forall_{(\eta_1,\eta_2)\in O} \quad \lim_{x\to\infty} Y(x,\eta_1,\eta_2) = (0,0). \tag{5.141}$$

Moreover, the continuity of V implies K to be closed. Since $K \subseteq [c,d] \times [-\sqrt{2r}, \sqrt{2r}]$, it is also bounded, i.e. compact. Thus, Th. 5.54 applies to each $\eta \in O$. So let $\eta \in O$. We will show that $M = \{(0,0)\}$, where M is the set of Th. 5.54(c) (then $\omega(\eta) = \{(0,0)\}$ by Th. 5.54(c), which implies (5.141) as desired). To verify $M = \{(0,0)\}$, note $\dot{V}(y_1,y_2) < 0$ for $y_1, y_2 \neq 0$, showing $(y_1,y_2) \notin M$. For $y_1 = 0$, $y_2 \neq 0$, let $\phi := Y(\cdot, y_1, y_2)$. Then $\phi_2(0) = y_2 \neq 0$ and $\phi_1'(0) = y_2 \neq 0$, i.e. both ϕ_1 and ϕ_2 are nonzero on some interval $]0,\epsilon[$ with $\epsilon > 0$, showing $(y_1,y_2) \notin M$. Likewise, if $y_1 \neq 0$, $y_2 = 0$, then let ϕ be as before. This time $\phi_1(0) = y_1 \neq 0$ and $\phi_2'(0) = -h(y_1) \neq 0$, again showing both ϕ_1 and ϕ_2 are nonzero on some interval $]0,\epsilon[$ with $\epsilon > 0$, implying $(y_1,y_2) \notin M$.

A Differentiability

We provide a lemma used in the variation of constants Th. 2.3.

Lemma A.1. *Let $O \subseteq \mathbb{R}$ be open. If the function $a: O \longrightarrow \mathbb{K}$ is differentiable, then*

$$f: O \longrightarrow \mathbb{K}, \quad f(x) := e^{a(x)} \tag{A.1a}$$

is differentiable with

$$f': O \longrightarrow \mathbb{K}, \quad f'(x) := a'(x)\, e^{a(x)}. \tag{A.1b}$$

Proof. For $\mathbb{K} = \mathbb{R}$, the lemma is immediate from the one-dimensional chain rule for real-valued functions [Phi13a, (9.15)]. It remains to consider the case $\mathbb{K} = \mathbb{C}$. Note that we can not apply the chain rule for holomorphic (i.e. $\mathbb{C}$-differentiable functions), since a is only $\mathbb{R}$-differentiable and it does not need to have a holomorphic extension. However, we can argue as follows, merely using the chain rule and the product rule for real-valued functions: Write $a = b + ic$ with differentiable functions $b, c : O \longrightarrow \mathbb{R}$. Then

$$f(x) = e^{a(x)} = e^{b(x)+ic(x)} = e^{b(x)}\, e^{ic(x)} = e^{b(x)} \big(\sin c(x) + i \cos c(x) \big). \tag{A.2}$$

Thus, one computes

$$\begin{aligned} f'(x) &= b'(x)\, e^{b(x)}\, e^{ic(x)} + e^{b(x)} \big(-c'(x) \cos c(x) + ic'(x) \sin c(x) \big) \\ &= b'(x)\, e^{a(x)} + ic'(x)\, e^{b(x)} \big(i \cos c(x) + \sin c(x) \big) = b'(x)\, e^{a(x)} + ic'(x)\, e^{b(x)}\, e^{ic(x)} \\ &= \big(b'(x) + ic'(x) \big) e^{a(x)} = a'(x)\, e^{a(x)}, \end{aligned} \tag{A.3}$$

proving (A.1b). ■

B $\mathbb{K}^n$-Valued Integration

During the course of this class, we frequently need $\mathbb{K}^n$-valued integrals. In particular, for $f : I \longrightarrow \mathbb{K}^n$, I an interval in $\mathbb{R}$, we make use of the estimate $\| \int_I f \| \leq \int_I \|f\|$, for example in the proof of the Peano Th. 3.8. As mentioned in the proof of Th. 3.8, the estimate can easily be checked directly for the 1-norm on $\mathbb{K}^n$, but it does hold for *every* norm on $\mathbb{K}^n$. To verify this result is the main purpose of the present section. Throughout the class, it suffices to use Riemann integrals. However, some readers might be more familiar with Lebesgue integrals, which is a more general notion (every Riemann integrable function is also Lebesgue integrable). For convenience, the material is presented twice, first using Riemann integrals and arguments that make specific use of techniques available for Riemann integrals, then, second, using Lebesgue integrals and corresponding techniques. For Riemann integrals, the norm estimate is proved in Th. B.4, for Lebesgue integrals in Th. B.9.

B.1 $\mathbb{K}^n$-Valued Riemann Integral

Definition B.1. Let $a, b \in \mathbb{R}$, $I := [a, b]$. We call a function $f : I \longrightarrow \mathbb{K}^n$, $n \in \mathbb{N}$, *Riemann integrable* if, and only if, each coordinate function $f_j = \pi_j \circ f : I \longrightarrow \mathbb{K}$, $j = 1, \dots, n$, is Riemann integrable. Denote the set of all Riemann integrable functions from I into $\mathbb{K}^n$ by $\mathcal{R}(I, \mathbb{K}^n)$. If $f : I \longrightarrow \mathbb{K}^n$ is Riemann integrable, then

$$\int_I f := \left(\int_I f_1, \dots, \int_I f_n \right) \in \mathbb{K}^n \tag{B.1}$$

is the ($\mathbb{K}^n$-valued) *Riemann integral* of f over I.

Remark B.2. The linearity of the $\mathbb{K}$-valued integral implies the linearity of the $\mathbb{K}^n$-valued integral.

Theorem B.3. *Let $a, b \in \mathbb{R}$, $a \le b$, $I := [a,b]$. If $f \in \mathcal{R}(I, \mathbb{K}^n)$, $n \in \mathbb{N}$, and $\phi : f(I) \longrightarrow \mathbb{R}$ is Lipschitz continuous, then $\phi \circ f \in \mathcal{R}(I, \mathbb{R})$.*

Proof. If $\mathbb{K} = \mathbb{R}$, then $\phi \circ f = \psi \circ \iota \circ f$, where $\iota : \mathbb{R}^n \longrightarrow \mathbb{C}^n$ is the canonical imbedding, and $\psi : \mathbb{C}^n \longrightarrow \mathbb{R}$, $\psi(z_1, \dots, z_n) := \phi(\operatorname{Re} z_1, \dots, \operatorname{Re} z_n)$. Clearly, $\iota \circ f \in \mathcal{R}(I, \mathbb{C}^n)$, and, if ϕ is L-Lipschitz, $L \ge 0$, then, due to

$$\forall_{z,w \in \mathbb{C}^n} \quad \begin{aligned} |\psi(z) - \psi(w)| &= |\phi(\operatorname{Re} z) - \phi(\operatorname{Re} w)| \le L \, \|\operatorname{Re} z - \operatorname{Re} w\| \\ &\overset{(*)}{\le} CL \, \|\operatorname{Re} z - \operatorname{Re} w\|_1 = CL \sum_{j=1}^{n} |\operatorname{Re} z_j - \operatorname{Re} w_j| \\ &\overset{\text{[Phi13a, Th. 5.11(d)]}}{\le} CL \sum_{j=1}^{n} |z_j - w_j| = CL \, \|z - w\|_1 \overset{(**)}{\le} \tilde{C}CL \, \|z - w\|, \end{aligned} \tag{B.2}$$

where the estimate at $(*)$ holds with $C \in \mathbb{R}^+$, due to the equivalence of $\|\cdot\|$ and $\|\cdot\|_1$ on $\mathbb{R}^n$, and the estimate at $(**)$ holds with $\tilde{C} \in \mathbb{R}^+$, due to the equivalence of $\|\cdot\|_1$ and $\|\cdot\|$ on $\mathbb{C}^n$. Thus, by (B.2), ψ is Lipschitz as well, namely $\tilde{C}CL$-Lipschitz, and it suffices to consider the case $\mathbb{K} = \mathbb{C}$, which we proceed to do next. Once again using the equivalence of $\|\cdot\|_1$ and $\|\cdot\|$ on $\mathbb{C}^n$, there exists $c \in \mathbb{R}^+$ such that $\|z\| \le c\|z\|_1$ for each $z \in \mathbb{C}^n$. Assume ϕ to be L-Lipschitz, $L \ge 0$. If $f \in \mathcal{R}(I, \mathbb{C}^n)$, then $\operatorname{Re} f_1, \dots, \operatorname{Re} f_n \in \mathcal{R}(I, \mathbb{R})$ and $\operatorname{Im} f_1, \dots, \operatorname{Im} f_n \in \mathcal{R}(I, \mathbb{R})$, i.e., given $\epsilon > 0$, Riemann's integrability criterion of [Phi13a, Th. 10.12] provides partitions $\Delta_1, \dots, \Delta_n$ of I and $\Pi_1, \dots, \Pi_n$ of I such that

$$\forall_{j=1,\dots,n} \quad \begin{aligned} R(\Delta_j, \operatorname{Re} f_j) - r(\Delta_j, \operatorname{Re} f_j) &< \frac{\epsilon}{2ncL}, \\ R(\Pi_j, \operatorname{Im} f_j) - r(\Pi_j, \operatorname{Im} f_j) &< \frac{\epsilon}{2ncL}, \end{aligned} \tag{B.3}$$

where R and r denote upper and lower Riemann sums, respectively (cf. [Phi13a, (10.7)]). Letting Δ be a joint refinement of the $2n$ partitions $\Delta_1, \dots, \Delta_n, \Pi_1, \dots, \Pi_n$, we have (cf. [Phi13a, Def. 10.8(a),(b)] and [Phi13a, Th. 10.10(a)])

$$\forall_{j=1,\dots,n} \quad \begin{aligned} R(\Delta, \operatorname{Re} f_j) - r(\Delta, \operatorname{Re} f_j) &< \frac{\epsilon}{2ncL}, \\ R(\Delta, \operatorname{Im} f_j) - r(\Delta, \operatorname{Im} f_j) &< \frac{\epsilon}{2ncL}. \end{aligned} \tag{B.4}$$

Recalling that, for each $g : I \longrightarrow \mathbb{R}$ and $\Delta = (x_0, \dots, x_N) \in \mathbb{R}^{N+1}$, $N \in \mathbb{N}$, $a = x_0 < x_1 < \dots < x_N = b$, $I_k := [x_{k-1}, x_k]$, it is

$$r(\Delta, g) = \sum_{k=1}^{N} m_k \, |I_k| = \sum_{k=1}^{N} m_k(g)(x_k - x_{k-1}), \tag{B.5a}$$

$$R(\Delta, g) = \sum_{k=1}^{N} M_k \, |I_k| = \sum_{k=1}^{N} M_k(g)(x_k - x_{k-1}), \tag{B.5b}$$

where

$$m_k(g) := \inf\{g(x) : x \in I_k\}, \quad M_k(g) := \sup\{g(x) : x \in I_k\}, \tag{B.5c}$$

we obtain, for each $\xi_k, \eta_k \in I_k$,

$$\begin{aligned}
&\big|(\phi \circ f)(\xi_k) - (\phi \circ f)(\eta_k)\big| \leq L\, \big\|f(\xi_k) - f(\eta_k)\big\| \leq cL\, \big\|f(\xi_k) - f(\eta_k)\big\|_1 \\
&= cL \sum_{j=1}^{n} \big|f_j(\xi_k) - f_j(\eta_k)\big| \\
&\overset{\text{[Phi13a, Th. 5.11(d)]}}{\leq} cL \sum_{j=1}^{n} \big|\operatorname{Re} f_j(\xi_k) - \operatorname{Re} f_j(\eta_k)\big| + cL \sum_{j=1}^{n} \big|\operatorname{Im} f_j(\xi_k) - \operatorname{Im} f_j(\eta_k)\big| \\
&\leq cL \sum_{j=1}^{n} \big(M_k(\operatorname{Re} f_j) - m_k(\operatorname{Re} f_j)\big) + cL \sum_{j=1}^{n} \big(M_k(\operatorname{Im} f_j) - m_k(\operatorname{Im} f_j)\big).
\end{aligned} \tag{B.6}$$

Thus,

$$\begin{aligned}
&R(\Delta, \phi \circ f) - r(\Delta, \phi \circ f) = \sum_{k=1}^{N} \big(M_k(\phi \circ f) - m_k(\phi \circ f)\big)|I_k| \\
&\overset{\text{(B.6)}}{\leq} cL \sum_{k=1}^{N} \sum_{j=1}^{n} \big(M_k(\operatorname{Re} f_j) - m_k(\operatorname{Re} f_j)\big)|I_k| \\
&\quad + cL \sum_{k=1}^{N} \sum_{j=1}^{n} \big(M_k(\operatorname{Im} f_j) - m_k(\operatorname{Im} f_j)\big)|I_k| \\
&= cL \sum_{j=1}^{n} \big(R(\Delta, \operatorname{Re} f_j) - r(\Delta, \operatorname{Re} f_j)\big) + cL \sum_{j=1}^{n} \big(R(\Delta, \operatorname{Im} f_j) - r(\Delta, \operatorname{Im} f_j)\big) \\
&\overset{\text{(B.4)}}{<} 2ncL\, \frac{\epsilon}{2ncL} = \epsilon.
\end{aligned} \tag{B.7}$$

Thus, $\phi \circ f \in \mathcal{R}(I, \mathbb{R})$ by [Phi13a, Th. 10.12]. ■

Theorem B.4. *Let $a, b \in \mathbb{R}$, $a \leq b$, $I := [a, b]$. For each norm $\|\cdot\|$ on $\mathbb{K}^n$, $n \in \mathbb{N}$, and each Riemann integrable $f : I \longrightarrow \mathbb{K}^n$, it is $\|f\| \in \mathcal{R}(I, \mathbb{R})$, and the following holds:*

$$\left\| \int_I f \right\| \leq \int_I \|f\|. \tag{B.8}$$

Proof. From Th. B.3, we obtain $\|f\| \in \mathcal{R}(I, \mathbb{R})$, as the norm $\|\cdot\|$ is 1-Lipschitz by the inverse triangle inequality. Let Δ be an arbitrary partition of I. Recalling that, for each $g : I \longrightarrow \mathbb{R}$ and $\Delta = (x_0, \dots, x_N) \in \mathbb{R}^{N+1}$, $N \in \mathbb{N}$, $a = x_0 < x_1 < \dots < x_N = b$, $I_k := [x_{k-1}, x_k]$, $\xi_k \in I_k$, the intermediate Riemann sums

$$\rho(\Delta, f) = \sum_{k=1}^{N} f(t_k)\, |I_k| = \sum_{k=1}^{N} f(t_k)(x_k - x_{k-1}), \tag{B.9}$$

we obtain, for $\xi_k \in I_k$,

$$
\begin{aligned}
&\Big\|\Big(\big(\rho(\Delta,\operatorname{Re} f_1),\ \rho(\Delta,\operatorname{Im} f_1)\big),\dots,\big(\rho(\Delta,\operatorname{Re} f_n),\ \rho(\Delta,\operatorname{Im} f_n)\big)\Big)\Big\|\\
&=\left\|\left(\left(\sum_{k=1}^N \operatorname{Re} f_1(\xi_k)\,|I_k|,\ \sum_{k=1}^N \operatorname{Im} f_1(\xi_k)\,|I_k|\right),\dots,\left(\sum_{k=1}^N \operatorname{Re} f_n(\xi_k)\,|I_k|,\ \sum_{k=1}^N \operatorname{Im} f_n(\xi_k)\,|I_k|\right)\right)\right\|\\
&=\left\|\sum_{k=1}^N\Big(\Big(\operatorname{Re} f_1(\xi_k)\,|I_k|,\ \operatorname{Im} f_1(\xi_k)\,|I_k|\Big),\dots,\Big(\operatorname{Re} f_n(\xi_k)\,|I_k|,\ \operatorname{Im} f_n(\xi_k)\,|I_k|\Big)\Big)\right\|\\
&\leq\sum_{k=1}^N\Big\|\Big(\Big(\operatorname{Re} f_1(\xi_k),\ \operatorname{Im} f_1(\xi_k)\Big),\dots,\Big(\operatorname{Re} f_n(\xi_k),\ \operatorname{Im} f_n(\xi_k)\Big)\Big)\Big\|\,|I_k|\\
&=\sum_{k=1}^N\|f(\xi_k)\|\,|I_k|=\rho(\Delta,\|f\|).
\end{aligned}
\tag{B.10}
$$

Since the intermediate Riemann sums in (B.10) converge to the respective integrals by [Phi13a, (10.24b)], one obtains

$$
\begin{aligned}
\left\|\int_I f\right\| &= \lim_{|\Delta|\to 0}\Big\|\Big(\big(\rho(\Delta,\operatorname{Re} f_1),\ \rho(\Delta,\operatorname{Im} f_1)\big),\dots,\big(\rho(\Delta,\operatorname{Re} f_n),\ \rho(\Delta,\operatorname{Im} f_n)\big)\Big)\Big\|\\
&\overset{\text{(B.10)}}{\leq} \lim_{|\Delta|\to 0}\rho(\Delta,\|f\|)=\int_I\|f\|,
\end{aligned}
\tag{B.11}
$$

proving (B.8). ■

B.2 $\mathbb{K}^n$-Valued Lebesgue Integral

Definition B.5. Let $I \subseteq \mathbb{R}$ be (Lebesgue) measurable, $n \in \mathbb{N}$.

(a) A function $f : I \longrightarrow \mathbb{K}^n$ is called *(Lebesgue) measurable* (respectively, *(Lebesgue) integrable*) if, and only if, each coordinate function $f_j = \pi_j \circ f : I \longrightarrow \mathbb{K}$, $j = 1,\dots,n$, is (Lebesgue) measurable (respectively, (Lebesgue) integrable), which, for $\mathbb{K} = \mathbb{C}$, means if, and only if, each $\operatorname{Re} f_j$ and each $\operatorname{Im} f_j$, $j = 1,\dots,n$, is (Lebesgue) measurable (respectively, (Lebesgue) integrable).

(b) If $f : I \longrightarrow \mathbb{K}^n$ is integrable, then

$$\int_I f := \left(\int_I f_1,\dots,\int_I f_n\right) \in \mathbb{K}^n \tag{B.12}$$

is the ($\mathbb{K}^n$-valued) *(Lebesgue) integral* of f over I.

Remark B.6. The linearity of the $\mathbb{K}$-valued integral implies the linearity of the $\mathbb{K}^n$-valued integral.

Theorem B.7. *Let $I \subseteq \mathbb{R}$ be measurable, $n \in \mathbb{N}$. Then $f : I \longrightarrow \mathbb{K}^n$ is measurable in the sense of Def.* B.5(a) *if, and only if, $f^{-1}(O)$ is measurable for each open subset O of $\mathbb{K}^n$.*

Proof. Assume $f^{-1}(O)$ is measurable for each open subset O of $\mathbb{K}^n$. Let $j \in \{1, \dots, n\}$. If $O_j \subseteq \mathbb{K}$ is open in $\mathbb{K}$, then $O := \pi_j^{-1}(O_j) = \{z \in \mathbb{K}^n : z_j \in O_j\}$ is open in $\mathbb{K}^n$. Thus, $f_j^{-1}(O_j) = f^{-1}(O)$ is measurable, showing that each f_j is measurable, i.e. f is measurable. Now assume f is measurable, i.e. each f_j is measurable. Since every open $O \subseteq \mathbb{K}^n$ is a countable union of open sets of the form $O = O_1 \times \cdots \times O_n$ with each O_j being an open subset of $\mathbb{K}$, it suffices to show that the preimages of such open sets are measurable. So let O be as above. Then $f^{-1}(O) = \bigcap_{j=1}^n f_j^{-1}(O_j)$, showing that $f^{-1}(O)$ is measurable. ■

Corollary B.8. *Let $I \subseteq \mathbb{R}$ be measurable, $n \in \mathbb{N}$. If $f : I \longrightarrow \mathbb{K}^n$ is measurable, then $\|f\| : I \longrightarrow \mathbb{R}$ is measurable.*

Proof. If $O \subseteq \mathbb{R}$ is open, then $\|\cdot\|^{-1}(O)$ is an open subset of $\mathbb{K}^n$ by the continuity of the norm. In consequence, $\|f\|^{-1}(O) = f^{-1}\big(\|\cdot\|^{-1}(O)\big)$ is measurable. ■

Theorem B.9. *Let $I \subseteq \mathbb{R}$ be measurable, $n \in \mathbb{N}$. For each norm $\|\cdot\|$ on $\mathbb{K}^n$ and each integrable $f : I \longrightarrow \mathbb{K}^n$, the following holds:*

$$\left\| \int_I f \right\| \leq \int_I \|f\|. \tag{B.13}$$

Proof. First assume that $B \subseteq I$ is measurable, $y \in \mathbb{K}^n$, and $f = y\,\chi_B$, where χ_B is the characteristic function of B (i.e. the f_j are y_j on B and 0 on $I \setminus B$). Then

$$\left\| \int_I f \right\| = \Big\| \big(y_1 \lambda(B), \dots, y_n \lambda(B)\big) \Big\| = \lambda(B)\|y\| = \int_I \|f\|, \tag{B.14}$$

where λ denotes Lebesgue measure on $\mathbb{R}$. Next, consider the case that f is a so-called *simple function*, that means f takes only finitely many values $y_1, \dots, y_N \in \mathbb{K}^n$, $N \in \mathbb{N}$, and each preimage $B_j := f^{-1}\{y_j\} \subseteq I$ is measurable. Then

$$f = \sum_{j=1}^N y_j \chi_{B_j}, \tag{B.15}$$

where, without loss of generality, we may assume that the B_j are pairwise disjoint. We obtain

$$\begin{aligned} \left\| \int_I f \right\| &\leq \sum_{j=1}^N \left\| \int_I y_j \chi_{B_j} \right\| = \sum_{j=1}^N \int_I \|y_j \chi_{B_j}\| = \int_I \sum_{j=1}^N \|y_j \chi_{B_j}\| \\ &\overset{(*)}{=} \int_I \left\| \sum_{j=1}^N y_j \chi_{B_j} \right\| = \int_I \|f\|, \end{aligned} \tag{B.16}$$

where, at $(*)$, it was used that, as the B_j are disjoint, the integrands of the two integrals are equal at each $x \in I$.

Now, if f is integrable, then each $\operatorname{Re} f_j$ and each $\operatorname{Im} f_j$ is integrable (i.e. $\operatorname{Re} f_j, \operatorname{Im} f_j \in L^1(I)$) and there exist sequences of simple functions $\phi_{j,k} : I \longrightarrow \mathbb{R}$ and $\psi_{j,k} : I \longrightarrow \mathbb{R}$ such that $\lim_{k\to\infty} \|\phi_{j,k} - \operatorname{Re} f_j\|_{L^1(I)} = \lim_{k\to\infty} \|\psi_{j,k} - \operatorname{Im} f_j\|_{L^1(I)} = 0$. In particular,

$$0 \le \lim_{k\to\infty} \left| \int_I \phi_{j,k} - \int_I \operatorname{Re} f_j \right| \le \lim_{k\to\infty} \|\phi_{j,k} - \operatorname{Re} f_j\|_{L^1(I)} = 0, \tag{B.17a}$$

$$0 \le \lim_{k\to\infty} \left| \int_I \psi_{j,k} - \int_I \operatorname{Im} f_j \right| \le \lim_{k\to\infty} \|\psi_{j,k} - \operatorname{Im} f_j\|_{L^1(I)} = 0, \tag{B.17b}$$

and also

$$\begin{aligned} 0 &\le \lim_{k\to\infty} \|\phi_{j,k} + i\psi_{j,k} - f_j\|_{L^1(I)} \\ &\le \lim_{k\to\infty} \|\phi_{j,k} - \operatorname{Re} f_j\|_{L^1(I)} + \lim_{k\to\infty} \|\psi_{j,k} - \operatorname{Im} f_j\|_{L^1(I)} = 0. \end{aligned} \tag{B.18}$$

Thus, we obtain

$$\begin{aligned} \left\| \int_I f \right\| &= \left\| \left(\int_I f_1, \dots, \int_I f_n \right) \right\| \\ &= \left\| \left(\lim_{k\to\infty} \int_I \phi_{1,k} + i \lim_{k\to\infty} \int_I \psi_{1,k}, \dots, \lim_{k\to\infty} \int_I \phi_{n,k} + i \lim_{k\to\infty} \int_I \psi_{n,k} \right) \right\| \\ &= \lim_{k\to\infty} \left\| \int_I (\phi_k + i\,\psi_k) \right\| \le \lim_{k\to\infty} \int_I \|\phi_k + i\psi_k\| \overset{(*)}{=} \int_I \|f\|, \end{aligned} \tag{B.19}$$

where the equality at $(*)$ holds due to $\lim_{k\to\infty} \big\| \|(\phi_{1,k}, \dots, \phi_{n,k})\| - \|f\| \big\|_{L^1(I)} = 0$, which, in turn, is verified by

$$\begin{aligned} 0 &\le \int_I \big| \|\phi_k + i\psi_k\| - \|f\| \big| \le \int_I \|\phi_k + i\psi_k - f\| \le C \int_I \|\phi_k + i\psi_k - f\|_1 \\ &= C \int_I \sum_{j=1}^n \big|\phi_{j,k} + i\psi_{j,k} - f_j\big| \to 0 \quad \text{for } k \to \infty, \end{aligned} \tag{B.20}$$

with $C \in \mathbb{R}^+$ since the norms $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent on $\mathbb{K}^n$. ∎

C Metric Spaces

C.1 Distance in Metric Spaces

Lemma C.1. *The following law holds in every metric space (X, d):*

$$|d(x,y) - d(x',y')| \le d(x,x') + d(y,y') \quad \textit{for each } x, x', y, y' \in X. \tag{C.1}$$

In particular, (C.1) *states the Lipschitz continuity of $d : X^2 \longrightarrow \mathbb{R}_0^+$ (with Lipschitz constant* 1*) with respect to the metric d_1 on X^2 defined by*

$$d_1 : X^2 \times X^2 \longrightarrow \mathbb{R}_0^+, \quad d_1\big((x,y), (x',y')\big) = d(x,x') + d(y,y'). \tag{C.2}$$

Further consequences are the continuity and even uniform continuity of d, and also the continuity of d in both components.

Proof. First, note $d(x,y) \leq d(x,x') + d(x',y') + d(y',y)$, i.e.

$$d(x,y) - d(x',y') \leq d(x,x') + d(y',y). \tag{C.3a}$$

Second, $d(x',y') \leq d(x',x) + d(x,y) + d(y,y')$, i.e.

$$d(x',y') - d(x,y) \leq d(x',x) + d(y,y'). \tag{C.3b}$$

Taken together, (C.3a) and (C.3b) complete the proof of (C.1). ■

Definition C.2. Let (X,d) be a nonempty metric space. For each $A, B \subseteq X$ define the *distance* between A and B by

$$\operatorname{dist}(A,B) := \inf\{d(a,b) : \ a \in A, \ b \in B\} \in [0,\infty] \tag{C.4}$$

and

$$\underset{x \in X}{\forall} \quad \operatorname{dist}(x,B) := \operatorname{dist}(\{x\},B) \quad \text{and} \quad \operatorname{dist}(A,x) := \operatorname{dist}(A,\{x\}). \tag{C.5}$$

Remark C.3. Clearly, for $\operatorname{dist}(A,B)$ as defined in (C.4), we have

$$\operatorname{dist}(A,B) < \infty \quad \Leftrightarrow \quad A \neq \emptyset \text{ and } B \neq \emptyset. \tag{C.6}$$

Theorem C.4. *Let (X,d) be a nonempty metric space. If $A \subseteq X$ and $A \neq \emptyset$, then the functions*

$$\delta, \tilde{\delta} : \ X \longrightarrow \mathbb{R}_0^+, \quad \delta(x) := \operatorname{dist}(x,A), \quad \tilde{\delta}(x) := \operatorname{dist}(A,x), \tag{C.7}$$

are both Lipschitz continuous with Lipschitz constant 1 *(in particular, they are both continuous and even uniformly continuous).*

Proof. Since $\operatorname{dist}(x,A) = \operatorname{dist}(A,x)$, it suffices to verify the Lipschitz continuity of δ. We need to show

$$\underset{x,y \in X}{\forall} \quad |\operatorname{dist}(x,A) - \operatorname{dist}(y,A)| \leq d(x,y). \tag{C.8}$$

To this end, let $x, y \in X$ and $a \in A$ be arbitrary. Then

$$\operatorname{dist}(x,A) \leq d(x,a) \leq d(x,y) + d(y,a) \tag{C.9}$$

and

$$\operatorname{dist}(x,A) - d(x,y) \leq d(y,a), \tag{C.10}$$

implying

$$\operatorname{dist}(x,A) - d(x,y) \leq \operatorname{dist}(y,A) \tag{C.11}$$

and

$$\operatorname{dist}(x,A) - \operatorname{dist}(y,A) \leq d(x,y). \tag{C.12}$$

Since $x, y \in X$ were arbitrary, (C.12) also yields

$$\operatorname{dist}(y,A) - \operatorname{dist}(x,A) \leq d(x,y), \tag{C.13}$$

where (C.12) and (C.13) together are precisely (C.8). ■

Definition C.5. Let (X,d) be a metric space, $A \subseteq X$, and $\epsilon \in \mathbb{R}^+$. Define

$$A_\epsilon := \{x \in X : d(x,A) < \epsilon\}, \tag{C.14a}$$

$$\overline{A}_\epsilon := \{x \in X : d(x,A) \leq \epsilon\}. \tag{C.14b}$$

We call A_ϵ the *open ϵ-fattening* of A, and $\overline{A}_\epsilon$ the *closed ϵ-fattening* of A.

Lemma C.6. *Let (X,d) be a metric space, $A \subseteq X$, and $\epsilon \in \mathbb{R}^+$. Then A_ϵ, the open ϵ-fattening of A, is, indeed, open, and $\overline{A}_\epsilon$, the closed ϵ-fattening of A, is, indeed, closed.*

Proof. Since the distance function $\delta : X \longrightarrow \mathbb{R}_0^+$, $\delta(x) := \operatorname{dist}(x,A)$, is continuous by Th. C.4, $A_\epsilon = \delta^{-1}[0,\epsilon[$ is open as the continuous preimage of an open set (note that $[0,\epsilon[$ is, indeed, (relatively) open in $\mathbb{R}_0^+$); $\overline{A}_\epsilon = \delta^{-1}[0,\epsilon]$ is closed as the continuous preimage of a closed set. ■

Lemma C.7. *Let (X,d) be a metric space, $A \subseteq X$, and $\epsilon \in \mathbb{R}^+$. If A is bounded, then so are the fattenings A_ϵ and $\overline{A}_\epsilon$.*

Proof. If A is bounded, then there exist $x \in X$ and $r > 0$ such that $A \subseteq B_r(x)$. Let $s := r + \epsilon + 1$. If $y \in \overline{A}_\epsilon$, then there exists $a \in A$ such that $d(a,y) < \epsilon + 1$. Thus,

$$d(x,y) \leq d(x,a) + d(a,y) < r + \epsilon + 1 = s, \tag{C.15}$$

showing $A_\epsilon \subseteq \overline{A}_\epsilon \subseteq B_s(x)$, i.e. A_ϵ and $\overline{A}_\epsilon$ are bounded. ■

Proposition C.8. *Let (X,d) be a metric space, $A \subseteq X$, and $0 < \epsilon_1 < \epsilon_2$.*

(a) *Then $A \subseteq A_{\epsilon_1} \subseteq \overline{A}_{\epsilon_1} \subseteq A_{\epsilon_2} \subseteq \overline{A}_{\epsilon_2}$ always holds.*

(b) *If $(X, \|\cdot\|)$ is a normed space with d being the induced metric, $\emptyset \neq A \subseteq X$, and there exists $x \notin A$, satisfying $\delta := d(x,A) \geq \epsilon_2$, then all the inclusions in* (a) *are strict: $A \subsetneq A_{\epsilon_1} \subsetneq \overline{A}_{\epsilon_1} \subsetneq A_{\epsilon_2} \subsetneq \overline{A}_{\epsilon_2}$. Caveat: For general metric spaces X and A satisfying all the hypotheses, the inclusions do not need to be strict (consider discrete metric spaces for simple examples).*

Proof. (a) is immediate from (C.14).

To prove (b), consider the maps

$$\phi : [0,1] \longrightarrow X, \quad \phi(t) := tx + (1-t)a, \tag{C.16a}$$

$$f : [0,1] \longrightarrow \mathbb{R}, \quad f(t) := d\big(\phi(t), A\big). \tag{C.16b}$$

If $(s_n)_{n\in\mathbb{N}}$ is a sequence in $[0,1]$ such that $\lim_{n\to\infty} s_n = s \in [0,1]$, then $\lim_{n\to\infty} \phi(s_n) = sx+(1-s)a = \phi(s)$, i.e. ϕ is continuous. Then, using Th. C.4, f is also continuous. Thus, since $f(0) = d(a,A) = 0$ and $f(1) = d(x,A) = \delta \geq \epsilon_2$, one can use the intermediate value theorem [Phi13a, Th. 7.57] to obtain, for each $\epsilon \in [0,\epsilon_2]$, some $\tau \in [0,1]$, satisfying $f(\tau) = \epsilon$. If $\epsilon > 0$, then $d(\phi(\tau), A) = f(\tau) = \epsilon > 0$, i.e $\phi(\tau) \in \overline{A}_\epsilon \setminus A$ and $\phi(\tau) \in \overline{A}_\epsilon \setminus A_\epsilon$, showing $A \subsetneq A_{\epsilon_1}$, $A_{\epsilon_1} \subsetneq \overline{A}_{\epsilon_1}$, and $A_{\epsilon_2} \subsetneq \overline{A}_{\epsilon_2}$. If $\epsilon := (\epsilon_1 + \epsilon_2)/2$, then $\epsilon_1 < \epsilon = f(\tau) = d(\phi(\tau), A) < \epsilon_2$, i.e. $A_{\epsilon_2} \setminus \overline{A}_{\epsilon_1}$, showing $\overline{A}_{\epsilon_1} \subsetneq A_{\epsilon_2}$. ■

C.2 Compactness in Metric Spaces

Definition C.9. A subset C of a metric space X is called *compact* if, and only if, every sequence in C has a subsequence that converges to some limit $c \in C$.

Proposition C.10. *Let (X, d) be a metric space, $C, A \subseteq X$. If C is compact, A is closed, and $A \cap C = \emptyset$, then* $\operatorname{dist}(C, A) > 0$.

Proof. Proceeding by contraposition, we show that $\operatorname{dist}(C, A) = 0$ implies $A \cap C \neq \emptyset$. If $\operatorname{dist}(C, A) = 0$, then there exists a sequence $((c_k, a_k))_{k\in\mathbb{N}}$ in $C \times A$ such that

$$\lim_{k\to\infty} d(c_k, a_k) = 0. \tag{C.17}$$

As C is compact, we may assume

$$\lim_{k\to\infty} c_k = c \in C, \tag{C.18}$$

also implying

$$\lim_{k\to\infty} a_k = c, \quad \text{since} \quad \forall_{k\in\mathbb{N}} \quad d(a_k, c) \leq d(a_k, c_k) + d(c_k, c). \tag{C.19}$$

Since A is closed, (C.19) yields $c \in A$, i.e. $c \in A \cap C$. ■

Proposition C.11. *Let (X, d) be a metric space and $C \subseteq X$.*

(a) *If C is compact, then C is closed and bounded.*

(b) *If C is compact and $A \subseteq C$ is closed, then A is compact.*

Proof. (a): Suppose C is compact. Let $(x^k)_{k\in\mathbb{N}}$ be a sequence in C that converges in X, i.e. $\lim_{k\to\infty} x^k = x \in X$. Since C is compact, $(x^k)_{k\in\mathbb{N}}$ must have a subsequence that converges to some $c \in C$, implying $x = c \in C$ and showing C is closed. If C is not bounded, then, for each $x \in X$, there is a sequence $(x^k)_{k\in\mathbb{N}}$ in C such that $\lim_{k\to\infty} d(x, x^k) = \infty$. If $y \in X$, then $d(x, x^k) \leq d(x, y) + d(y, x^k)$, i.e. $d(y, x^k) \geq d(x, x^k) - d(x, y)$, showing that $\lim_{k\to\infty} d(y, x^k) = \infty$ as well. Thus, y can not be a limit of any subsequence of $(x^k)_{k\in\mathbb{N}}$. As y was arbitrary, C can not be compact.

(b): If $(x^k)_{k\in\mathbb{N}}$ is a sequence in A, then $(x^k)_{k\in\mathbb{N}}$ is a sequence in C. Since C is compact, it must have a subsequence that converges to some $c \in C$. However, as A is closed, c must be in A, showing that $(x^k)_{k\in\mathbb{N}}$ has a subsequence that converges to some $c \in A$, i.e. A is compact. ■

Corollary C.12. *A subset C of $\mathbb{K}^n$, $n \in \mathbb{N}$, is compact if, and only if, C is closed and bounded.*

Proof. Every compact set is closed and bounded by Prop. C.11(a). If C is closed and bounded, and $(x^k)_{k\in\mathbb{N}}$ is a sequence in C, then the boundedness and the Bolzano-Weierstrass theorem yield a subsequence that converges to some $x \in \mathbb{K}^n$. However, since C is closed, $x \in C$, showing that C is compact. ■

The following examples show that, in general, sets can be closed and bounded without being compact.

Example C.13. **(a)** If (X, d) is a noncomplete metric space, than it contains a Cauchy sequence that does not converge. It is not hard to see that such a sequence can not have a convergent subsequence, either. This shows that no noncomplete metric space can be compact. Moreover, the closure of every bounded subset of X that contains such a nonconvergent Cauchy sequence is an example of a closed and bounded set that is noncompact. Concrete examples are given by $\mathbb{Q} \cap [a, b]$ for each $a, b \in \mathbb{R}$ with $a < b$ (these sets are $\mathbb{Q}$-closed, but not $\mathbb{R}$-closed!) and $]a, b[$ for each $a, b \in \mathbb{R}$ with $a < b$, in each case endowed with the usual metric $d(x, y) := |x - y|$.

(b) There can also be closed and bounded sets in complete spaces that are not compact. Consider the space X of all bounded sequences $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$, endowed with the sup-norm $\|(x_n)_{n\in\mathbb{N}}\|_{\sup} := \sup\{|x_n| : n \in \mathbb{N}\}$. It is not too difficult to see that X with the sup-norm is a Banach space: Let $(x^k)_{k\in\mathbb{N}}$ with $x^k = (x^k_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in X. Then, for each $n \in \mathbb{N}$, $(x^k_n)_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{K}$, and, thus, it has a limit $y_n \in \mathbb{K}$. Let $y := (y_n)_{n\in\mathbb{N}}$. Then

$$\|x^k - y\|_{\sup} = \sup\{|x^k_n - y_n| : n \in \mathbb{N}\}.$$

Let $\epsilon > 0$. As $(x^k)_{k\in\mathbb{N}}$ is a Cauchy sequence with respect to the sup-norm, there is $N \in \mathbb{N}$ such that $\|x^k - x^l\|_{\sup} < \epsilon$ for all $k, l > N$. Fix some $l > N$ and some $n \in \mathbb{N}$. Then $\epsilon \geq \lim_{k\to\infty} |x^k_n - x^l_n| = \lim_{k\to\infty} |y_n - x^l_n|$. Since this is valid for each $n \in \mathbb{N}$, we get $\|x^l - y\|_{\sup} \leq \epsilon$ for each $l > N$, showing $\lim_{l\to\infty} x^l = y$, i.e. X is complete and a Banach space.

Now consider the sequence $(e^k)_{k\in\mathbb{N}}$ with

$$e^k_n := \begin{cases} 1 & \text{for } k = n, \\ 0 & \text{otherwise.} \end{cases}$$

Then $(e^k)_{k\in\mathbb{N}}$ constitutes a sequence in X with $\|e^k\|_{\sup} = 1$ for each $k \in \mathbb{N}$. In particular, $(e^k)_{k\in\mathbb{N}}$ is a sequence inside the closed unit ball $\overline{B}_1(0)$, and, hence, bounded. However, if $k, l \in \mathbb{N}$ with $k \neq l$, then $\|e^k - e^l\|_{\sup} = 1$. Thus, neither $(e^k)_{k\in\mathbb{N}}$ nor any subsequence can be a Cauchy sequence. In particular, no subsequence can converge, showing that the closed and bounded unit ball $\overline{B}_1(0)$ is not compact.

Note: There is an important result that shows that a normed vector space is finite-dimensional if, and only if, the closed unit ball $\overline{B}_1(0)$ is compact (see, e.g., [Str08, Th. 28.14]).

Theorem C.14. *If (X, d_X) and (Y, d_Y) are metric spaces, $C \subseteq X$ is compact, and $f : C \longrightarrow Y$ is continuous, then $f(C)$ is compact.*

Proof. If $(y^k)_{k\in\mathbb{N}}$ is a sequence in $f(C)$, then, for each $k \in \mathbb{N}$, there is some $x^k \in C$ such that $f(x^k) = y^k$. As C is compact, there is a subsequence $(a^k)_{k\in\mathbb{N}}$ of $(x^k)_{k\in\mathbb{N}}$ with $\lim_{k\to\infty} a^k = a$ for some $a \in C$. Then $(f(a^k))_{k\in\mathbb{N}}$ is a subsequence of $(y^k)_{k\in\mathbb{N}}$ and the continuity of f yields $\lim_{k\to\infty} f(a^k) = f(a) \in f(C)$, showing that $(y^k)_{k\in\mathbb{N}}$ has a convergent subsequence with limit in $f(C)$. We have therefore established that $f(C)$ is compact. ■

Theorem C.15. *If (X, d) is a metric space, $C \subseteq X$ is compact, and $f : C \longrightarrow \mathbb{R}$ is continuous, then f assumes its max and its min, i.e. there are $x_m \in C$ and $x_M \in C$ such that f has a global min at x_m and a global max at x_M.*

Proof. Since C is compact and f is continuous, $f(C) \subseteq \mathbb{R}$ is compact according to Th. C.14. Then, by [Phi13a, Lem. 7.53], $f(C)$ contains a smallest element m and a largest element M. This, in turn, implies that there are $x_m, x_M \in C$ such that $f(x_m) = m$ and $f(x_M) = M$. ■

Theorem C.16. *If (X, d_X) and (Y, d_Y) are metric spaces, $C \subseteq X$ is compact, and $f : C \longrightarrow Y$ is continuous, then f is uniformly continuous.*

Proof. If f is not uniformly continuous, then there must be some $\epsilon > 0$ such that, for each $k \in \mathbb{N}$, there exist $x^k, y^k \in C$ satisfying $d_X(x^k, y^k) < 1/k$ and $d_Y(f(x^k), f(y^k)) \geq \epsilon$. Since C is compact, there is $a \in C$ and a subsequence $(a^k)_{k\in\mathbb{N}}$ of $(x^k)_{k\in\mathbb{N}}$ such that $a = \lim_{k\to\infty} a^k$. Then there is a corresponding subsequence $(b^k)_{k\in\mathbb{N}}$ of $(y^k)_{k\in\mathbb{N}}$ such that $d_X(a^k, b^k) < 1/k$ and $d_Y(f(a^k), f(b^k)) \geq \epsilon$ for all $k \in \mathbb{N}$. Using the compactness of C again, there $b \in C$ and a subsequence $(v^k)_{k\in\mathbb{N}}$ of $(b^k)_{k\in\mathbb{N}}$ such that $b = \lim_{k\to\infty} v^k$. Now there is a corresponding subsequence $(u^k)_{k\in\mathbb{N}}$ of $(a^k)_{k\in\mathbb{N}}$ such that $d_X(u^k, v^k) < 1/k$ and $d_Y(f(u^k), f(v^k)) \geq \epsilon$ for all $k \in \mathbb{N}$. Note that we still have $a = \lim_{k\to\infty} v^k$. Given $\alpha > 0$, there is $N \in \mathbb{N}$ such that, for each $k > N$, one has $d_X(a, u^k) < \alpha/3$, $d_X(b, v^k) < \alpha/3$, and $d_X(u^k, v^k) < 1/k < \alpha/3$. Thus, $d_X(a, b) < d_X(a, u^k) + d_X(u^k, v^k) + d_X(b, v^k) < \alpha$, implying $d(a, b) = 0$ and $a = b$. Finally, the continuity of f implies $f(a) = \lim_{k\to\infty} f(u^k) = \lim_{k\to\infty} f(v^k)$ in contradiction to $d_Y(f(u^k), f(v^k)) \geq \epsilon$. ■

Theorem C.17. *If (X, d_X) and (Y, d_Y) are metric spaces, $C \subseteq X$ is compact, and $f : C \longrightarrow Y$ is continuous and one-to-one, then $f^{-1} : f(C) \longrightarrow C$ is continuous.*

Proof. Let $(y^k)_{k\in\mathbb{N}}$ be a sequence $f(C)$ such that $\lim_{k\to\infty} y^k = y \in f(C)$. Then there is a sequence $(x^k)_{k\in\mathbb{N}}$ in C such that $f(x^k) = y^k$ for each $k \in \mathbb{N}$. Let $x := f^{-1}(y)$. It remains to prove that $\lim_{k\to\infty} x^k = x$. As C is compact, there is $a \in C$ and a subsequence $(a^k)_{k\in\mathbb{N}}$ of $(x^k)_{k\in\mathbb{N}}$ such that $a = \lim_{k\to\infty} a^k$. The continuity of f yields $f(a) = \lim_{k\to\infty} f(a^k) = \lim_{k\to\infty} y^k = y = f(x)$ since $(f(a^k))_{k\in\mathbb{N}}$ is a subsequence of $(y^k)_{k\in\mathbb{N}}$. It now follows that $a = x$ since f is one-to-one. The same argument shows that *every* convergent subsequence of $(x^k)_{k\in\mathbb{N}}$ has to converge to x. If $(x^k)_{k\in\mathbb{N}}$ did not converge to x, then there had to be some $\epsilon > 0$ such that infinitely man x^k are not in $B_\epsilon(x)$. However, the compactness of C would provide a convergent subsequence whose limit could not be x, in contradiction to x having to be the limit of all convergent subsequences of $(x^k)_{k\in\mathbb{N}}$. ■

Definition C.18. A subset A of a metric space (X, d) is called *precompact* or *totally bounded* if, and only if, for each $\epsilon > 0$, A can be covered by finitely many ϵ-balls, i.e. if, and only if, there exist finitely many points $a_1, \dots, a_N \in A$, $N \in \mathbb{N}$, such that

$$A \subseteq \bigcup_{j=1}^{N} B_\epsilon(a_j). \tag{C.20}$$

Theorem C.19. *For a subset C of a metric space (X, d), the following statements are equivalent:*

(i) *C is compact as defined in Def.* C.9.

(ii) *C has the* Heine-Borel property*, i.e. every open cover of C has a finite subcover, i.e. if $(O_j)_{j\in I}$ is a family of open sets $O_j \subseteq C$, satisfying*

$$A \subseteq \bigcup_{j\in I}^{N} O_j, \tag{C.21}$$

then there exist $j_1, dots, j_N \in I$, $N \in \mathbb{N}$, such that $A \subseteq \bigcup_{j=1}^{N} O_j$.

(iii) *C is precompact (i.e. totally bounded) as defined in Def.* C.18 *and* complete*, i.e. every Cauchy sequence in C converges to a limit in C.*

Proof. We show (i) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i).

"(i) $\Rightarrow$ (iii)": Let $(c_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in C. As C is compact, $(c_n)_{n\in\mathbb{N}}$ has a subsequence $(c_{n_j})_{j\in\mathbb{N}}$ such that $\lim_{j\to\infty} c_{n_j} = c \in C$. Given $\epsilon > 0$ choose $K \in \mathbb{N}$ such that, for each $m, n \geq K$, $d(c_m, c_n) < \frac{\epsilon}{2}$, and such that, for each $n_j \geq K$, $d(c_{n_j}, c) < \frac{\epsilon}{2}$. Then, fixing some $n_j \geq K$,

$$\forall_{n\geq K} \quad d(c_n, c) \leq d(c_n, c_{n_j}) + d(c_{n_j}, c) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \tag{C.22}$$

showing $\lim_{n\to\infty} c_n = c$ and the completeness of C. We now show C to be also totally bounded. We proceed by contraposition and assume C not to be totally bounded, i.e. there exists $\epsilon > 0$ such that C is not contained in any finite union of ϵ-balls. Inductively, we construct a sequence $(c_n)_{n\in\mathbb{N}}$ in C such that

$$\forall_{\substack{m,n\in\mathbb{N},\\ m\neq n}} \quad d(c_m, c_n) \geq \epsilon : \tag{C.23}$$

To start with, we note $C \neq \emptyset$ and choose some arbitrary $c_1 \in C$. Assuming $c_1, \dots, c_k \in C$, $k \in \mathbb{N}$, have already been constructed such that $d(c_m, c_n) \geq \epsilon$ holds for each $m, n \in \{1, \dots, k\}$, there must be

$$c \in C \setminus \bigcup_{j=1}^{k} B_\epsilon(c_j). \tag{C.24}$$

Choosing $c_{k+1} := c$, (C.24) guarantees (C.23) now holds for each $m, n \in \{1, \dots, k+1\}$. Due to (C.23), no subsequence of $(c_n)_{n\in\mathbb{N}}$ can be a Cauchy sequence, i.e. $(c_n)_{n\in\mathbb{N}}$ does not have a convergent subsequence, proving C is not compact.

"(iii) ⇒ (ii)": Assume C to be precompact and complete. For each $k \in \mathbb{N}$, the precompactness yields points $c_1^k, \dots, c_{N_k}^k \in C$, $N_k \in \mathbb{N}$, such that

$$C \subseteq \bigcup_{j=1}^{N_k} B_{\frac{1}{k}}(c_j^k). \tag{C.25}$$

Seeking a contradiction, assume C does not have the Heine-Borel property, i.e. there exists an open cover $(O_j)_{j\in I}$ of C which does not have a finite subcover. Inductively, we construct a decreasing sequence of subsets C_k of C, $C \supseteq C_1 \supseteq C_2 \supseteq \dots$, such that no C_k can be covered by a finite subcover of $(O_j)_{j\in I}$ and such that

$$\underset{k\in\mathbb{N}}{\forall} \quad \underset{j\in\{1,\dots,N_k\}}{\exists} \quad C_k \subseteq B_{\frac{1}{k}}(c_j^k): \tag{C.26}$$

To start out, we note that (C.25) implies at least one of the finitely many sets $C \cap B_1(c_1^1), \dots, C\cap B_1(c_{N_1}^1)$ can not be covered by a finite subcover of $(O_j)_{j\in I}$, say, $C\cap B_1(c_{j_1}^1)$. Define $C_1 := C\cap B_1(c_{j_1}^1)$. Then, given $C_1, \dots, C_k$ have already been constructed for some $k \in \mathbb{N}$, since C_k can not be covered by a finite subcover of $(O_j)_{j\in I}$ and

$$C_k \subseteq C \subseteq \bigcup_{j=1}^{N_{k+1}} B_{\frac{1}{k+1}}(c_j^{k+1}), \tag{C.27}$$

there exists $j_{k+1} \in \{1, \dots, N_{k+1}\}$ such that $C_k \cap B_{\frac{1}{k+1}}(c_{j_{k+1}}^{k+1})$ can not be covered by a finite subcover of $(O_j)_{j\in I}$, either. Define $C_{k+1} := C_k \cap B_{\frac{1}{k+1}}(c_{j_{k+1}}^{k+1})$. For each $k \in \mathbb{N}$, choose some $s_k \in C_k$ (note $C_k \neq \emptyset$, as it can not be covered by finitely many O_j). Given $\epsilon > 0$, there is $K \in \mathbb{N}$ such that $\frac{2}{K} < \epsilon$. If $k, l \geq K$, then $s_k, s_l \in C_K \subseteq B_{\frac{1}{K}}(c_j^K)$ for some suitable $j \in \{1, \dots, N_K\}$. In particular, $d(s_k, s_l) < \frac{2}{K} < \epsilon$, showing $(s_k)_{k\in\mathbb{N}}$ is a Cauchy sequence. As $(s_k)_{k\in\mathbb{N}}$ is a Cauchy sequence in C and C is complete, there exists $c \in C$ such that $\lim_{k\to\infty} s_k = c$. However, then there must exist some $j \in I$ such that $c \in O_j$ and, since O_j is open, there is $\epsilon > 0$ with $B_\epsilon(c) \subseteq O_j$, and $B_\epsilon(c)$ must contain almost all of the s_k. Choose k sufficiently large such that $\frac{1}{k} < \frac{\epsilon}{4}$ and $d(s_k, c) < \frac{\epsilon}{2}$. Then, since

$$s_k \in C_k \subseteq B_{\frac{1}{k}}(c_j^k), \tag{C.28}$$

one has

$$\underset{x\in B_{\frac{1}{k}}(c_j^k)}{\forall} \quad d(x,c) \leq d(x, s_k) + d(s_k, c) < \frac{2}{k} + \frac{\epsilon}{2} < \frac{2\epsilon}{4} + \frac{\epsilon}{2} = \epsilon, \tag{C.29}$$

showing $C_k \subseteq B_{\frac{1}{k}}(c_j^k) \subseteq B_\epsilon(c) \subseteq O_j$, in contradiction to C_k not being coverable by finitely many O_j.

"(ii) ⇒ (i)": Assume C has the Heine-Borel property. Seeking a contradiction, assume C is not compact, that means there exists a sequence $(c_n)_{n\in\mathbb{N}}$ in C such that no subsequence of $(c_n)_{n\in\mathbb{N}}$ converges to a limit in C. According to [Phi13b, Prop. 1.38(d)], no $c \in C$ can be a cluster point of $(c_n)_{n\in\mathbb{N}}$, i.e., for each $c \in C$, there exists $\epsilon_c > 0$ such that $B_{\epsilon_c}(c)$ contains only finitely many of the c_n. Since $C \subseteq \bigcup_{c\in C} B_{\epsilon_c}(c)$, the family $\big(B_{\epsilon_c}(c)\big)_{c\in C}$ constitutes an open cover of C. As C has the Heine-Borel property, there exist finitely many points $c_1, \dots, c_N \in C$, $N \in \mathbb{N}$, such that $C \subseteq \bigcup_{j=1}^N B_{\epsilon_{c_j}}(c_j)$, i.e. C contains only finitely many of the c_n, in contradiction to $(c_n)_{n\in\mathbb{N}}$ being a sequence in C. ■

Caveat C.20. In general topological spaces, one *defines* compactness via the Heine-Borel property (a topological space C is defined to be *compact* if, and only if, C has the Heine-Borel property). Moreover, a topological space C is defined to be *sequentially compact* if, and only if, every sequence in C has a convergent subsequence. Using this terminology, one can rephrase the equivalence between (i) and (ii) in Th. C.19 by stating that a metric space is sequentially compact if, and only if, it is compact. However, in general topological spaces, *neither* implication remains true ((iii) of Th. C.19 does not even make sense in general topological spaces, as the concepts of boundedness, total boundedness, and Cauchy sequences are, in general, not available): For an example of a topological space that is compact, but not sequentially compact, see, e.g. [Pre75, 7.2.10(a)]; for an example of a topological space that is sequentially compact, but not compact, see, e.g. [Pre75, 7.2.10(c)].

Theorem C.21 (Lebesgue Number)**.** *Let (X, d) be a metric space and $C \subseteq X$. If C is compact and $(O_j)_{j\in I}$ is an open cover of C, then there exists a* Lebesgue number *δ for the open cover, i.e. some $\delta > 0$ such that, for each $A \subseteq C$ with* $\operatorname{diam} A < \delta$*, there exists $j_0 \in I$, where $A \subseteq O_{j_0}$. Recall that*

$$\operatorname{diam} A = \begin{cases} 0 & \text{for } A = \emptyset, \\ \sup\big\{d(x,y) : x, y \in A\big\} & \text{for } \emptyset \neq A. \end{cases} \tag{C.30}$$

Proof. Seeking a contradiction, assume there is no Lebesgue number for the open cover $(O_j)_{j\in I}$. Then there is a sequence of pairs $(x_k, y_k) \in C^2$ such that

$$d(x_k, y_k) < \frac{1}{k} \quad \text{but} \quad \underset{k\in\mathbb{N}}{\forall} \quad \underset{j\in I}{\forall} \quad \{x_k, y_k\} \subsetneq O_j. \tag{C.31}$$

As C is compact, we may assume that $\lim_{k\to\infty} x_k = c \in C$, implying $\lim_{k\to\infty} y_k = c$ as well. But then there must be O_j such that $c \in O_j$ and, due to the openness of O_j and the convergences of both sequences to c, O_j must contain almost all of the x_k as well as almost all of the y_k in contradiction to (C.31). ■

D Local Lipschitz Continuity

In Prop. 3.13, it was shown that a continuous function is locally Lipschitz with respect to y if, and only if, it is globally Lipschitz with respect to y on every compact set. The following Prop. D.1 shows that this equivalence holds even if f is not continuous, provided that each projection G_x as in (D.1) below is convex. On the other hand, Ex. D.2 shows that, in general, there exist discontinuous functions that are locally Lipschitz with respect to y without being globally Lipschitz with respect to y on every compact set.

Proposition D.1. *Let $m, n \in \mathbb{N}$, $G \subseteq \mathbb{R} \times \mathbb{K}^m$, and $f : G \longrightarrow \mathbb{K}^n$. If G is such that each projection*

$$G_x := \{y \in \mathbb{K}^m : (x, y) \in G\}, \quad x \in \mathbb{R}, \tag{D.1}$$

is convex (in particular, if G itself is convex), then f is locally Lipschitz with respect to y if, and only if, f is (globally) Lipschitz with respect to y on every compact subset K of G.

Proof. The proof of Prop. 3.13 shows, whithout making use of the continuity of f, that (global) Lipschitz continuity with respect to y on every compact subset K of G implies local Lipschitz continuity on G. Thus, assume f to be locally Lipschitz with respect to y and assume each G_x to be convex. The proof of Prop. 3.13 shows, whithout making use of the continuity of f, that, for each $K \subseteq G$ compact

$$\underset{\substack{\delta>0,\\ L\geq 0}}{\exists} \quad \underset{(x,y),(x,\bar{y})\in K}{\forall} \quad \Big(\|y-\bar{y}\| < \delta \quad \Rightarrow \quad \|f(x,y)-f(x,\bar{y})\| \leq L\|y-\bar{y}\| \Big). \tag{D.2}$$

If $(x,y),(x,\bar{y}) \in K$ are arbitrary with $y \neq \bar{y}$, then the convexity of G_x implies

$$\{(x,(1-t)y)+(x,t\bar{y}) : t \in [0,1]\} \subseteq G. \tag{D.3}$$

Choose $N \in \mathbb{N}$ such that $N > 2\|y-\bar{y}\|/\delta$ and set $h := \|y-\bar{y}\|/N$. Then

$$h < \delta/2. \tag{D.4}$$

Define

$$\underset{k=0,\dots,N}{\forall} \quad y_k := \frac{kh}{\|y-\bar{y}\|}\, y + \left(1 - \frac{kh}{\|y-\bar{y}\|}\right) \bar{y}. \tag{D.5}$$

Then

$$\underset{k=0,\dots,N-1}{\forall} \quad \|y_{k+1}-y_k\| = \left\| \frac{h}{\|y-\bar{y}\|}\, y + \frac{h}{\|y-\bar{y}\|}\, \bar{y} \right\| = h < \delta \tag{D.6}$$

and

$$\begin{aligned} \|f(x,y)-f(x,\bar{y})\| &\leq \sum_{k=0}^{N-1} \|f(x,y_k)-f(x,y_{k+1})\| \overset{\text{(D.2)}}{\leq} L \sum_{k=0}^{N-1} \|y_k - y_{k+1}\| \\ &= L\,N\,h = L\,\|y-\bar{y}\|, \end{aligned} \tag{D.7}$$

showing f to be (globally) L-Lipschitz with respect to y on K. ■

Example D.2. We provide two examples that show that, in general, a discontinuous function can be locally Lipschitz with respect to y without being globally Lipschitz with respect to y on every compact set.

(a) Consider

$$G :=]-2,2[\times \big( ]-4,-1[\cup]1,4[\big) \tag{D.8}$$

and $f : G \longrightarrow \mathbb{R}$,

$$f(x,y) := \begin{cases} 1/x & \text{for } x \neq 0,\ y \in]-4,-1[, \\ 0 & \text{for } x = 0,\ y \in]-4,-1[, \\ 0 & \text{for } y \in]1,4[. \end{cases} \tag{D.9}$$

For the following open balls with respect to the max norm $\|(x,y)\| := \max\{|x|,|y|\}$, one has $B_1(x,y) \cap G \subseteq]-2,2[\times]-4,-1[$ for $y \in]-4,-1[$, and$B_1(x,y) \cap G \subseteq]-2,2[\times]1,4[$ for $y \in]1,4[$. Thus, $f(x,\cdot)$ is constant on each set $B_1(x,y)\cap G$ (either constantly equal to $1/x$ or constantly equal to 0), i.e. 0-Lipschitz with respect to y. In particular, f is locally Lipschitz with respect to y. However, f is not Lipschitz continuous with respect to y on the compact set

$$K := [-1,1] \times \big([-3,-2] \cup [2,3]\big) : \tag{D.10}$$

For the sequence $((x_k, y_k, \overline{y}_k))_{k\in\mathbb{N}}$, where

$$\forall_{k\in\mathbb{N}} \quad x_k := 1/k, \quad y_k := -2, \quad \overline{y}_k := 2, \tag{D.11}$$

one has

$$\lim_{k\to\infty} \frac{|f(x_k,y_k) - f(x_k,\overline{y}_k)|}{|y_k - \overline{y}_k|} = \lim_{k\to\infty} \frac{k-0}{2-(-2)} = \infty, \tag{D.12}$$

showing f is not Lipschitz continuous with respect to y on K.

(b) If one increases the dimension by 1, then one can modify the example in (a) such the set G is even connected (this variant was pointed out by Anton Sporrer): Let

$$A := \big(]-4,-1[\times]-2,2[\big) \cup \big(]-4,4[\times]-2,0[\big) \cup \big(]1,4[\times]-2,2[\big) \subseteq \mathbb{R}^2. \tag{D.13}$$

Then A is open and connected (but not convex) and the same holds for

$$G :=]-2,2[\times A \subseteq \mathbb{R}^3. \tag{D.14}$$

Define

$$f : G \longrightarrow \mathbb{R}, \quad f(x,y_1,y_2) := \begin{cases} 1/x & \text{for } x \neq 0,\ y_1 \in]-4,-1[,\ y_2 > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{D.15}$$

Then everything works essentially as in (a) (it might be helpful to graphically visualize the set A and the behavior of the function f): For the following open balls with respect to the max norm $\|(x,y)\| := \max\{|x|,|y_1|,|y_2|\}$, one has

$$\forall_{\substack{(x,y_1,y_2)\in G,\\ y_1\in]-4,-1[}} \quad \Big((\xi,\eta_1,\eta_2) \in B_1(x,y_1,y_2) \cap G \quad \Rightarrow \quad \eta_1 < -1+1 = 0 < 1 \Big). \tag{D.16}$$

Analogous to (a), $f(x,\cdot)$ is constant on each set $B_1(x,y_1,y_2)\cap G$ (either constantly equal to $1/x$ or constantly equal to 0), i.e. 0-Lipschitz with respect to y. In particular, f is locally Lipschitz with respect to y. However, f is not Lipschitz continuous with respect to y on the compact set

$$K := [-1,1] \times \big([-3,-2] \cup [2,3]\big) \times [-1,1] : \tag{D.17}$$

For the sequence $((x_k, y_{1,k}, \overline{y}_{1,k}), y_{2,k})_{k\in\mathbb{N}}$ with

$$\forall_{k\in\mathbb{N}} \quad x_k := 1/k, \quad y_{1,k} := -2, \quad \overline{y}_{1,k} := 2, \quad y_{2,k} := 0, \tag{D.18}$$

one has

$$\lim_{k\to\infty} \frac{|f(x_k,y_{1,k},y_{2,k}) - f(x_k,\overline{y}_{1,k},y_{2,k})|}{\|(y_{1,k},y_{2,k}) - (\overline{y}_{1,k},y_{2,k})\|_{\max}} = \lim_{k\to\infty} \frac{k-0}{\max\{4,0\}} = \infty, \tag{D.19}$$

showing f is not Lipschitz continuous with respect to y on K.

E Maximal Solutions on Nonopen Intervals

In Def. 3.20, we required a maximal solution to an ODE to be defined on an *open* interval. The following Ex. E.1 shows it can occur that such a maximal solution has an extension to a larger *nonopen* interval. In such cases, one might want to call the solution on the nonopen interval maximal rather than the solution on the smaller open interval. However, this would make the treatment of maximal solutions more cumbersome in some places, without adding any real substance, which is why we stick to our requirement for maximal solutions to always be defined on an open interval.

Example E.1. (a) Let

$$G := [0,1] \times \mathbb{R}, \quad f : G \longrightarrow \mathbb{R}, \quad f(x,y) := 0. \tag{E.1}$$

Then, for each $(x_0, y_0) \in G$, the function

$$\phi : [0,1] \longrightarrow \mathbb{R}, \quad \phi \equiv y_0, \tag{E.2}$$

is a solution to the initial value problem

$$y' = f(x,y), \quad y(x_0) = y_0. \tag{E.3}$$

However, the maximal solution of (E.3) according to Def. 3.20 is $\phi{\restriction}_{]0,1[}$.

(b) The following modification of (a) allows f to be defined on all of $\mathbb{R}^2$: Let

$$G := \mathbb{R}^2, \quad f : G \longrightarrow \mathbb{R}, \quad f(x,y) := \begin{cases} 0 & \text{for } x \in [0,1], \\ 1 & \text{for } x \notin [0,1]. \end{cases} \tag{E.4}$$

Then, for each $(x_0, y_0) \in [0,1] \times \mathbb{R}$, the function ϕ of (E.2) is a solution to the initial value problem (E.3), but, again, the maximal solution of (E.3) according to Def. 3.20 is $\phi{\restriction}_{]0,1[}$.

F Paths in $\mathbb{R}^n$

Definition F.1. A *path* or *curve* in $\mathbb{R}^n$, $n \in \mathbb{N}$, is a continuous map $\psi : I \longrightarrow \mathbb{R}^n$, where $I \subseteq \mathbb{R}$ is an interval. One calls the path differentiable, continuously differentiable, etc. if, and only if, the function ψ has the respective property.

Definition F.2. If $a, b \in \mathbb{R}$, $a \leq b$, and $I := [a,b]$, then we call

$$|I| := b - a = |a - b|, \tag{F.1}$$

the *length* of I.

Definition F.3. Given a real interval $I := [a,b] \subseteq \mathbb{R}$, $a,b \in \mathbb{R}$, $a < b$, the $(N+1)$-tuple $\Delta := (x_0,\dots,x_N) \in \mathbb{R}^{N+1}$, $N \in \mathbb{N}$, is called a *partition* of I if, and only if, $a = x_0 < x_1 < \dots < x_N = b$. The set of all partitions of I is denoted by $\Pi(I)$ or by $\Pi[a,b]$. Given a partition Δ of I as above and letting $I_j := [x_{j-1},x_j]$, the number

$$|\Delta| := \max\big\{|I_j| : j \in \{1,\dots,N\}\big\}, \tag{F.2}$$

is called the *mesh size* of Δ.

Notation F.4. Given $a,b \in \mathbb{R}$, $a < b$, a path $\psi : [a,b] \longrightarrow \mathbb{R}^n$, $n \in \mathbb{N}$, and a partition $\Delta = (x_0,\dots,x_N)$, $N \in \mathbb{N}$, of $[a,b]$, we consider the approximation of ψ by the polygon, connecting the points $\psi(x_0),\dots,\psi(x_N)$, where we denote the polygon's length by

$$p_\psi(\Delta) := p_\psi(x_0,\dots,x_N) := \sum_{k=0}^{N-1} \|\psi(x_{k+1}) - \psi(x_k)\|_2, \tag{F.3}$$

using $\|\cdot\|_2$ to denote the 2-Norm on $\mathbb{R}^n$, i.e. the Euclidean norm.

Definition F.5. Given $a,b \in \mathbb{R}$, $a \le b$, for each path $\psi : [a,b] \longrightarrow \mathbb{R}^n$, $n \in \mathbb{N}$, define

$$l(\psi) := \begin{cases} 0 & \text{for } a = b,\\ \sup\Big\{p_\psi(\Delta) : \Delta \in \Pi[a,b]\Big\} \in [0,\infty] & \text{for } a < b. \end{cases} \tag{F.4}$$

The path ψ is called *rectifyable* with *arc length* $l(\psi)$ if, and only if, $l(\psi) < \infty$.

Proposition F.6. *Let $a,b \in \mathbb{R}$, $a < b$, and let $\psi : [a,b] \longrightarrow \mathbb{R}^n$ be a path, $n \in \mathbb{N}$.*

(a) *If ψ is affine, i.e. there exist $y_0, y_1 \in \mathbb{R}^n$ such that*

$$\underset{x\in[a,b]}{\forall} \quad \psi(x) = y_0 + x\,y_1, \tag{F.5}$$

then ψ is rectifyable with arc length

$$l(\psi) = \|y_1\|_2\,(b-a) = \|\psi(b) - \psi(a)\|_2. \tag{F.6}$$

(b) *If the path ψ is L-Lipschitz with $L \ge 0$, then ψ is rectifyable and*

$$l(\psi) \le L\,(b-a). \tag{F.7}$$

(c) *If the paths $\phi, \psi : [a,b] \longrightarrow \mathbb{R}^n$ are both rectifyable, then*

$$\big|l(\phi) - l(\psi)\big| \le l(\phi - \psi). \tag{F.8}$$

(d) *For each $\xi \in [a,b]$, it holds that*

$$l(\psi) = l(\psi{\restriction}_{[a,\xi]}) + l(\psi{\restriction}_{[\xi,b]}). \tag{F.9}$$

Proof. (a): For each partition $(x_0,\dots,x_N)$, $N\in\mathbb{N}$, of $[a,b]$, we have

$$\begin{aligned}p_\psi(x_0,\dots,x_N) &= \sum_{k=0}^{N-1}\|\psi(x_{k+1})-\psi(x_k)\|_2 = \sum_{k=0}^{N-1}\|x_{k+1}\,y_1 - x_k\,y_1\|_2\\ &= \|y_1\|_2\sum_{k=0}^{N-1}(x_{k+1}-x_k) = \|y_1\|_2\,(b-a),\end{aligned} \tag{F.10}$$

proving (F.6).

(b): For each partition $(x_0,\dots,x_N)$, $N\in\mathbb{N}$, of $[a,b]$, we have

$$\begin{aligned}p_\psi(x_0,\dots,x_N) &= \sum_{k=0}^{N-1}\|\psi(x_{k+1})-\psi(x_k)\|_2 \le \sum_{k=0}^{N-1}L\,\|x_{k+1}-x_k\|_2\\ &= L\sum_{k=0}^{N-1}(x_{k+1}-x_k) = L\,(b-a),\end{aligned} \tag{F.11}$$

proving (F.7).

(c): For each partition $\Delta=(x_0,\dots,x_N)$, $N\in\mathbb{N}$, of $[a,b]$, we have

$$\begin{aligned}\big|p_\phi(\Delta)-p_\psi(\Delta)\big| &= \left|\sum_{k=0}^{N-1}\|\phi(x_{k+1})-\phi(x_k)\|_2 - \sum_{k=0}^{N-1}\|\psi(x_{k+1})-\psi(x_k)\|_2\right|\\ &\le \sum_{k=0}^{N-1}\Big|\|\phi(x_{k+1})-\phi(x_k)\|_2 - \|\psi(x_{k+1})-\psi(x_k)\|_2\Big|\\ &\le \sum_{k=0}^{N-1}\Big\|\phi(x_{k+1})-\psi(x_{k+1})-\big(\phi(x_k)-\psi(x_k)\Big\|_2\\ &= p_{\phi-\psi}(\Delta),\end{aligned} \tag{F.12}$$

proving (F.8) (the last estimate in (F.12) holds true due to the inverse triangle inequality).

(d): If $\xi=a$ or $\xi=b$, then there is nothing to prove. Thus, assume $a<\xi<b$. If $\Delta_1:=(x_0,\dots,x_N)$ is a partition of $[a,\xi]$ and $\Delta_2:=(x_N,\dots,x_M)$ is a partition of $[\xi,b]$, $N,M\in\mathbb{N}$, $M>N$, then $\Delta:=(x_0,\dots,x_M)$ is a partition of $[a,b]$. Moreover,

$$p_\psi(\Delta)=p_\psi(\Delta_1)+p_\psi(\Delta_2) \tag{F.13}$$

is immediate from (F.3), implying

$$l(\psi)\ge l(\psi{\restriction}_{[a,\xi]})+l(\psi{\restriction}_{[\xi,b]}). \tag{F.14}$$

On the other hand, if $\Delta=(x_0,\dots,x_M)$ $M\in\mathbb{N}$, is a partition of $[a,b]$, then, either there is $0<N<M$ such that $\xi=N$, in which case (F.13) holds once again, where Δ_1 and Δ_2 are defined as before. Otherwise, there is $N\in\{0,\dots,M-1\}$ such that $x_N<\xi<x_{N+1}$

and, in this case, $\Delta_1 := (x_0, \dots, x_N, \xi)$ is a partition of $[a, \xi]$ and $\Delta_2 := (\xi, x_{N+1}, \dots, x_M)$ is a partition of $[\xi, b]$. Moreover,

$$\begin{aligned} p_\psi(\Delta) &= \sum_{k=0}^{M-1} \|\psi(x_{k+1}) - \psi(x_k)\|_2 \\ &= \sum_{k=0}^{N-1} \|\psi(x_{k+1}) - \psi(x_k)\|_2 + \|x_{N+1} - x_N\| + \sum_{k=N+1}^{M-1} \|\psi(x_{k+1}) - \psi(x_k)\|_2 \\ &\le p_\psi(\Delta_1) + p_\psi(\Delta_2), \end{aligned} \tag{F.15}$$

showing

$$l(\psi) \le l(\psi{\restriction}_{[a,\xi]}) + l(\psi{\restriction}_{[\xi,b]}) \tag{F.16}$$

and concluding the proof. ∎

Theorem F.7. *Given $a, b \in \mathbb{R}$, $a < b$, each continuously differentiable path $\psi : [a, b] \longrightarrow \mathbb{R}^n$, $n \in \mathbb{N}$, is rectifyable with arc length*

$$l(\psi) = \int_a^b \big\|\psi'(x)\big\|_2 \, \mathrm{d}x \,. \tag{F.17}$$

Proof. Since ψ is continuously differentiable, it follows from [Phi13b, Th. C.3] that ψ is Lipschitz continuous on $[a, b]$, i.e. ψ is rectifyable by Prop. F.6(b) above. To prove (F.17), according to the fundamental theorem of calculus [Phi13a, Th. 10.19(b)], it suffices to show the function

$$\lambda : [a, b] \longrightarrow \mathbb{R}_0^+, \quad \lambda(x) := l(\psi{\restriction}_{[a,x]}), \tag{F.18}$$

is differentiable with derivative $\lambda'(x) = \|\psi'(x)\|_2$. To this end, first note the continuous function ψ' is even uniformly continuous by Th. C.16. Thus,

$$\underset{\epsilon>0}{\forall} \ \underset{\delta>0}{\exists} \ \underset{x_0,x\in[a,b]}{\forall} \Big(|x_0 - x| < \delta \quad \Rightarrow \quad \|\psi(x_0) - \psi(x)\|_2 < \epsilon. \Big). \tag{F.19}$$

Fix $x_0 \in [a, b[$ and consider $x_1 \in]a, b[$ such that $x_0 < x_1 < x_0 + \delta$. Define the affine path

$$\alpha : [x_0, x_1] \longrightarrow \mathbb{R}^n, \quad \alpha(x) := \psi(x_0) + (x - x_0)\psi'(x_0). \tag{F.20}$$

According to Prop. F.6(a), we have

$$l(\alpha) = \|\psi'(x_0)\|_2 \, (x_1 - x_0). \tag{F.21}$$

Moreover, for the path $\psi - \alpha$, we have

$$\underset{x\in[x_0,x_1]}{\forall} \quad \|\psi'(x) - \alpha'(x)\|_2 = \|\psi'(x) - \psi'(x_0)\|_2 \overset{\text{(F.19)}}{<} \epsilon. \tag{F.22}$$

Thus, it follows from [Phi13b, Th. C.3] that $\psi - \alpha$ is ϵ-Lipschitz on $[a, b]$ and, then, Prop. F.6(b) yields

$$l(\psi{\restriction}_{[x_0,x_1]} - \alpha) \le \epsilon(x_1 - x_0), \tag{F.23}$$

Prop. F.6(c), in turn, yields

$$\left| l(\psi\restriction_{[x_0,x_1]}) - l(\alpha) \right| \leq l(\psi\restriction_{[x_0,x_1]} - \alpha) \leq \epsilon(x_1 - x_0). \tag{F.24}$$

Putting everything together, we obtain

$$\left| \frac{l(\psi\restriction_{[a,x_1]}) - l(\psi\restriction_{[a,x_0]})}{x_1 - x_0} - \|\psi'(x_0)\|_2 \right| \overset{\text{Prop. F.6(d), (F.21)}}{=} \left| \frac{l(\psi\restriction_{[x_0,x_1]})}{x_1 - x_0} - \frac{l(\alpha)}{x_1 - x_0} \right|$$
$$\overset{\text{(F.24)}}{\leq} \frac{\epsilon(x_1 - x_0)}{x_1 - x_0} = \epsilon, \tag{F.25}$$

showing the function λ from (F.18) has a right-hand derivative at x_0 and the value of that right-hand derivative at x_0 is the desired $\|\psi'(x_0)\|_2$. Repeating the above argument with $x_0, x_1 \in]a, b]$ such that $x_0 - \delta < x_1 < x_0$ shows λ to have a left-hand derivative at each $x_0 \in]a, b]$ with value $\|\psi'(x_0)\|_2$, which completes the proof. ■

Remark F.8. An example of a differentiable nonrectifyable path is given by (cf. [Wal02, Ex. 5.14.6])

$$\psi : [0,1] \longrightarrow \mathbb{R}^2, \quad \psi(x) := \begin{cases} x^2 \cos \frac{\pi}{x^2} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases} \tag{F.26}$$

G Operator Norms and Matrix Norms

For the present ODE class, we are mostly interested in linear maps from $\mathbb{K}^n$ into itself. However, introducing the relevant notions for linear maps between general normed vector spaces does not provide much additional difficulty, and, hopefully, even some extra clarity.

Definition G.1. Let $A : X \longrightarrow Y$ be a linear map between two normed vector spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ over $\mathbb{K}$. Then A is called bounded if, and only if, A maps bounded sets to bounded sets, i.e. if, and only if, $A(B)$ is a bounded subset of Y for each bounded $B \subseteq X$. The vector space of all bounded linear maps between X and Y is denoted by $\mathcal{L}(X, Y)$.

Definition G.2. Let $A : X \longrightarrow Y$ be a linear map between two normed vector spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ over $\mathbb{K}$. The number

$$\|A\| := \sup\left\{ \frac{\|Ax\|_Y}{\|x\|_X} : x \in X, \ x \neq 0 \right\}$$
$$= \sup\left\{ \|Ax\|_Y : x \in X, \ \|x\|_X = 1 \right\} \in [0, \infty] \tag{G.1}$$

is called the *operator norm* of A induced by $\|\cdot\|_X$ and $\|\cdot\|_Y$ (strictly speaking, the term operator norm is only justified if the value is finite, but it is often convenient to use the term in the generalized way defined here).

In the special case, where $X = \mathbb{K}^n$, $Y = \mathbb{K}^m$, and A is given via a real $m \times n$ matrix, the operator norm is also called *matrix norm*.

From now on, the space index of a norm will usually be suppressed, i.e. we write just $\|\cdot\|$ instead of both $\|\cdot\|_X$ and $\|\cdot\|_Y$.

Theorem G.3. *For a linear map $A : X \longrightarrow Y$ between two normed vector spaces $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ over $\mathbb{K}$, the following statements are equivalent:*

(a) *A is bounded.*

(b) *$\|A\| < \infty$.*

(c) *A is Lipschitz continuous.*

(d) *A is continuous.*

(e) *There is $x_0 \in X$ such that A is continuous at x_0.*

Proof. Since every Lipschitz continuous map is continuous and since every continuous map is continuous at every point, "(c) ⇒ (d) ⇒ (e)" is clear.

"(e) ⇒ (c)": Let $x_0 \in X$ be such that A is continuous at x_0. Thus, for each $\epsilon > 0$, there is $\delta > 0$ such that $\|x - x_0\| < \delta$ implies $\|Ax - Ax_0\| < \epsilon$. As A is linear, for each $x \in X$ with $\|x\| < \delta$, one has $\|Ax\| = \|A(x + x_0) - Ax_0\| < \epsilon$, due to $\|x + x_0 - x_0\| = \|x\| < \delta$. Moreover, one has $\|(\delta x)/2\| \leq \delta/2 < \delta$ for each $x \in X$ with $\|x\| \leq 1$. Letting $L := 2\epsilon/\delta$, this means that $\|Ax\| = \|A((\delta x)/2)\|/(\delta/2) < 2\epsilon/\delta = L$ for each $x \in X$ with $\|x\| \leq 1$. Thus, for each $x, y \in X$ with $x \neq y$, one has

$$\|Ax - Ay\| = \|A(x - y)\| = \|x - y\| \left\| A\left(\frac{x - y}{\|x - y\|}\right) \right\| < L\,\|x - y\|. \tag{G.2}$$

Together with the fact that $\|Ax - Ay\| \leq \|x - y\|$ is trivially true for $x = y$, this shows that A is Lipschitz continuous.

"(c) ⇒ (b)": As A is Lipschitz continuous, there exists $L \in \mathbb{R}_0^+$ such that $\|Ax - Ay\| \leq L\,\|x - y\|$ for each $x, y \in X$. Considering the special case $y = 0$ and $\|x\| = 1$ yields $\|Ax\| \leq L\,\|x\| = L$, implying $\|A\| \leq L < \infty$.

"(b) ⇒ (c)": Let $\|A\| < \infty$. We will show

$$\|Ax - Ay\| \leq \|A\|\,\|x - y\| \text{ for each } x, y \in X. \tag{G.3}$$

For $x = y$, there is nothing to prove. Thus, let $x \neq y$. One computes

$$\frac{\|Ax - Ay\|}{\|x - y\|} = \left\| A\left(\frac{x - y}{\|x - y\|}\right) \right\| \leq \|A\| \tag{G.4}$$

as $\left\| \frac{x-y}{\|x-y\|} \right\| = 1$, thereby establishing (G.3).

"(b) ⇒ (a)": Let $\|A\| < \infty$ and let $M \subseteq X$ be bounded. Then there is $r > 0$ such that $M \subseteq B_r(0)$. Moreover, for each $0 \neq x \in M$:

$$\frac{\|Ax\|}{\|x\|} = \left\| A\left(\frac{x}{\|x\|}\right) \right\| \leq \|A\| \tag{G.5}$$

as $\left\|\frac{x}{\|x\|}\right\| = 1$. Thus $\|Ax\| \leq \|A\|\|x\| \leq r\|A\|$, showing that $A(M) \subseteq B_{r\|A\|}(0)$. Thus, $A(M)$ is bounded, thereby establishing the case.

"(a) $\Rightarrow$ (b)": Since A is bounded, it maps the bounded set $B_1(0) \subseteq X$ into some bounded subset of Y. Thus, there is $r > 0$ such that $A(B_1(0)) \subseteq B_r(0) \subseteq Y$. In particular, $\|Ax\| < r$ for each $x \in X$ satisfying $\|x\| = 1$, showing $\|A\| \leq r < \infty$. ■

Remark G.4. For linear maps between finite-dimensional spaces, the equivalent properties of Th. G.3 always hold: Each linear map $A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$, $(n, m) \in \mathbb{N}^2$, is continuous (this follows, for example, from the fact that each such map is (trivially) differentiable, and every differentiable map is continuous). In particular, each linear map $A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$, has all the equivalent properties of Th. G.3.

Theorem G.5. *Let X and Y be normed vector spaces over $\mathbb{K}$.*

(a) *The operator norm does, indeed, constitute a norm on the set of bounded linear maps $\mathcal{L}(X, Y)$.*

(b) *If $A \in \mathcal{L}(X, Y)$, then $\|A\|$ is the smallest Lipschitz constant for A, i.e. $\|A\|$ is a Lipschitz constant for A and $\|Ax - Ay\| \leq L\,\|x - y\|$ for each $x, y \in X$ implies $\|A\| \leq L$.*

Proof. (a): If $A = 0$, then, in particular, $Ax = 0$ for each $x \in X$ with $\|x\| = 1$, implying $\|A\| = 0$. Conversely, $\|A\| = 0$ implies $Ax = 0$ for each $x \in X$ with $\|x\| = 1$. But then $Ax = \|x\|\,A(x/\|x\|) = 0$ for every $0 \neq x \in X$, i.e. $A = 0$. Thus, the operator norm is positive definite. If $A \in \mathcal{L}(X, Y)$, $\lambda \in \mathbb{K}$, and $x \in X$, then

$$\big\|(\lambda A)x\big\| = \big\|A(\lambda x)\big\| = \big\|\lambda(Ax)\big\| = |\lambda|\,\big\|Ax\big\|, \tag{G.6}$$

yielding

$$\begin{aligned}\|\lambda A\| &= \sup\big\{\|(\lambda A)x\| : x \in X,\ \|x\| = 1\big\} = \sup\big\{|\lambda|\,\|Ax\| : x \in X,\ \|x\| = 1\big\}\\ &= |\lambda|\,\sup\big\{\|Ax\| : x \in X,\ \|x\| = 1\big\} = |\lambda|\,\|A\|,\end{aligned} \tag{G.7}$$

showing that the operator norm is homogeneous of degree 1. Finally, if $A, B \in \mathcal{L}(X, Y)$ and $x \in X$, then

$$\|(A + B)x\| = \|Ax + Bx\| \leq \|Ax\| + \|Bx\|, \tag{G.8}$$

yielding

$$\begin{aligned}\|A + B\| &= \sup\big\{\|(A + B)x\| : x \in X,\ \|x\| = 1\big\}\\ &\leq \sup\big\{\|Ax\| + \|Bx\| : x \in X,\ \|x\| = 1\big\}\\ &\leq \sup\big\{\|Ax\| : x \in X,\ \|x\| = 1\big\} + \sup\big\{\|Bx\| : x \in X,\ \|x\| = 1\big\}\\ &= \|A\| + \|B\|,\end{aligned} \tag{G.9}$$

showing that the operator norm also satisfies the triangle inequality, thereby completing the verification that it is, indeed, a norm.

(b): That $\|A\|$ is a Lipschitz constant for A was already shown in the proof of "(b) $\Rightarrow$ (c)" of Th. G.3. Now let $L \in \mathbb{R}_0^+$ be such that $\|Ax - Ay\| \leq L\,\|x - y\|$ for each $x, y \in X$. Specializing to $y = 0$ and $\|x\| = 1$ implies $\|Ax\| \leq L\,\|x\| = L$, showing $\|A\| \leq L$. ■

Remark G.6. Even though it is beyond the scope of the present class, let us mention as an outlook that one can show that $\mathcal{L}(X,Y)$ with the operator norm is a Banach space (i.e. a complete normed vector space) provided that Y is a Banach space (even if X is not a Banach space).

Lemma G.7. *If* $\mathrm{Id} : X \longrightarrow X$, $\mathrm{Id}(x) := x$, *is the identity map on a normed vector space* X *over* $\mathbb{K}$, *then* $\|\mathrm{Id}\| = 1$ *(in particular, the operator norm of a unit matrix is always* 1*). Caveat: In principle, one can consider two different norms on* X *simultaneously, and then the operator norm of the identity can differ from* 1.

Proof. If $\|x\| = 1$, then $\|\mathrm{Id}(x)\| = \|x\| = 1$. ■

Lemma G.8. *Let* X, Y, Z *be normed vector spaces and consider linear maps* $A \in \mathcal{L}(X,Y)$, $B \in \mathcal{L}(Y,Z)$. *Then*

$$\|BA\| \leq \|B\|\,\|A\|. \tag{G.10}$$

Proof. Let $x \in X$ with $\|x\| = 1$. If $Ax = 0$, then $\|B(A(x))\| = 0 \leq \|B\|\,\|A\|$. If $Ax \neq 0$, then one estimates

$$\big\|B(Ax)\big\| = \|Ax\| \left\|B\left(\frac{Ax}{\|Ax\|}\right)\right\| \leq \|A\|\,\|B\|, \tag{G.11}$$

thereby establishing the case. ■

Example G.9. Let $m, n \in \mathbb{N}$ and let $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be the linear map given by the $m \times n$ matrix $(a_{kl})_{(k,l)\in\{1,\dots,m\}\times\{1,\dots,n\}}$. Then

$$\|A\|_\infty := \max\left\{\sum_{l=1}^{n} |a_{kl}| : k \in \{1,\dots,m\}\right\} \tag{G.12a}$$

is called the *row sum norm* of A, and

$$\|A\|_1 := \max\left\{\sum_{k=1}^{m} |a_{kl}| : l \in \{1,\dots,n\}\right\} \tag{G.12b}$$

is called the *column sum norm* of A. It is an exercise to show that $\|A\|_\infty$ is the operator norm induced if $\mathbb{R}^n$ and $\mathbb{R}^m$ are endowed with the ∞-norm, and $\|A\|_1$ is the operator norm induced if $\mathbb{R}^n$ and $\mathbb{R}^m$ are endowed with the 1-norm.

H The Vandermonde Determinant

Theorem H.1. *Let* $n \in \mathbb{N}$ *and* $\lambda_0, \lambda_1, \dots, \lambda_n \in \mathbb{C}$. *Moreover, let*

$$V := \begin{pmatrix} 1 & \lambda_0 & \dots & \lambda_0^n \\ 1 & \lambda_1 & \dots & \lambda_1^n \\ \vdots & & & \vdots \\ 1 & \lambda_n & \dots & \lambda_n^n \end{pmatrix} \tag{H.1}$$

be the corresponding Vandermonde matrix*. Then its determinant, the so-called* Vandermonde determinant *is given by*

$$\det(V) = \prod_{\substack{k,l=0\\k>l}}^{n} (\lambda_k - \lambda_l). \tag{H.2}$$

Proof. The proof can be conducted by induction with respect to n: For $n = 1$, we have

$$\det(V) = \begin{vmatrix} 1 & \lambda_0 \\ 1 & \lambda_1 \end{vmatrix} = \lambda_1 - \lambda_0 = \prod_{\substack{k,l=0\\k>l}}^{1} (\lambda_k - \lambda_l), \tag{H.3}$$

showing (H.2) holds for $n = 1$. Now let $n > 1$. We know from Linear Algebra that the value of a determinant does not change if we add a multiple of a column to a different column. Adding the $(-\lambda_0)$-fold of the nth column to the $(n+1)$st column, we obtain in the $(n+1)$st column

$$\begin{pmatrix} 0 \\ \lambda_1^n - \lambda_1^{n-1}\lambda_0 \\ \vdots \\ \lambda_n^n - \lambda_n^{n-1}\lambda_0 \end{pmatrix}. \tag{H.4}$$

Next, one adds the $(-\lambda_0)$-fold of the $(n-1)$st column to the nth column, and, successively, the $(-\lambda_0)$-fold of the mth column to the $(m+1)$st column. One finishes, in the nth step, by adding the $(-\lambda_0)$-fold of the first column to the second column, obtaining

$$\det(V) = \begin{vmatrix} 1 & \lambda_0 & \dots & \lambda_0^n \\ 1 & \lambda_1 & \dots & \lambda_1^n \\ \vdots & & & \vdots \\ 1 & \lambda_n & \dots & \lambda_n^n \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & \lambda_1 - \lambda_0 & \lambda_1^2 - \lambda_1\lambda_0 & \dots & \lambda_1^n - \lambda_1^{n-1}\lambda_0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_n - \lambda_0 & \lambda_n^2 - \lambda_n\lambda_0 & \dots & \lambda_n^n - \lambda_n^{n-1}\lambda_0 \end{vmatrix}. \tag{H.5}$$

Applying the rule for determinants of block matrices to (H.5) yields

$$\det(V) = 1 \cdot \begin{vmatrix} \lambda_1 - \lambda_0 & \lambda_1^2 - \lambda_1\lambda_0 & \dots & \lambda_1^n - \lambda_1^{n-1}\lambda_0 \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_n - \lambda_0 & \lambda_n^2 - \lambda_n\lambda_0 & \dots & \lambda_n^n - \lambda_n^{n-1}\lambda_0 \end{vmatrix}. \tag{H.6}$$

As we also know from Linear Algebra that determinants are linear in each row, for each k, we can factor out $(\lambda_k - \lambda_0)$ from the kth row of (H.6), arriving at

$$\det(V) = \prod_{k=1}^{n} (\lambda_k - \lambda_0) \begin{vmatrix} 1 & \lambda_1 & \dots & \lambda_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_n & \dots & \lambda_n^{n-1} \end{vmatrix}. \tag{H.7}$$

However, the determinant in (H.7) is precisely the Vandermonde determinant of the $n-1$ numbers $\lambda_1, \dots, \lambda_n$, which is given according to the induction hypothesis, implying

$$\det(V) = \prod_{k=1}^{n} (\lambda_k - \lambda_0) \prod_{\substack{k,l=1\\k>l}}^{n} (\lambda_k - \lambda_l) = \prod_{\substack{k,l=0\\k>l}}^{n} (\lambda_k - \lambda_l), \tag{H.8}$$

completing the induction proof of (H.2). ■

I Matrix-Valued Functions

Notation I.1. Given $m, n \in \mathbb{N}$, let $\mathcal{M}(m, n, \mathbb{K})$ denote the set of $m \times n$ matrices over $\mathbb{K}$.

I.1 Product Rule

Proposition I.2. *Let $I \subseteq \mathbb{R}$ be a nontrivial interval, let $m, n, l \in \mathbb{N}$, and suppose*

$$A : I \longrightarrow \mathcal{M}(m, n, \mathbb{K}), \qquad A(x) = \big(a_{\alpha\beta}(x)\big), \tag{I.1a}$$

$$B : I \longrightarrow \mathcal{M}(n, l, \mathbb{K}), \qquad B(x) = \big(b_{\alpha\beta}(x)\big), \tag{I.1b}$$

are differentiable. Then

$$C : I \longrightarrow \mathcal{M}(m, l, \mathbb{K}), \quad C(x) := A(x)B(x), \tag{I.2}$$

is differentiable, and one has the product rule

$$\underset{x\in I}{\forall} \quad C'(x) = A'(x)B(x) + A(x)B'(x). \tag{I.3}$$

Proof. Writing $C(x) = \big(c_{\alpha\beta}(x)\big)$ and using the one-dimensional product rule together with the definition of matrix multiplication, one computes, for each $(\alpha, \beta) \in \{1, \dots, m\} \times \{1, \dots, l\}$,

$$\begin{aligned} c'_{\alpha\beta}(x) &= \left(\sum_{\gamma=1}^{n} a_{\alpha\gamma}(x)\, b_{\gamma\beta}(x)\right)' \\ &= \sum_{\gamma=1}^{n} a'_{\alpha\gamma}(x)\, b_{\gamma\beta}(x) + \sum_{\gamma=1}^{n} a_{\alpha\gamma}(x)\, b'_{\gamma\beta}(x) \\ &= \big(A'(x)B(x)\big)_{\alpha\beta} + \big(A(x)B'(x)\big)_{\alpha\beta}, \end{aligned} \tag{I.4}$$

proving the proposition. ■

I.2 Integration and Matrix Multiplication Commute

Proposition I.3. *Let $m, n, p \in \mathbb{N}$, let $I \subseteq \mathbb{R}$ be measurable (e.g. an interval), let $A : I \longrightarrow \mathcal{M}(m, n, \mathbb{K})$, $x \mapsto A(x) = (a_{kl}(x))$, be integrable (i.e. all $\operatorname{Re} a_{kl}, \operatorname{Im} a_{kl} : I \longrightarrow \mathbb{R}$ are integrable).*

(a) *If* $B = (b_{jk}) \in \mathcal{M}(p, m, \mathbb{K})$, *then*

$$B \int_I A(x)\,\mathrm{d}x = \int_I B\,A(x)\,\mathrm{d}x\,. \tag{I.5}$$

(b) *If* $B = (B_{lj}) \in \mathcal{M}(n, p, \mathbb{K})$, *then*

$$\left(\int_I A(x)\,\mathrm{d}x\right) B = \int_I A(x)\,B\,\mathrm{d}x\,. \tag{I.6}$$

Proof. (a): One computes, for each $(j, l) \in \{1, \dots, p\} \times \{1, \dots, n\}$,

$$\begin{aligned}\left(B \int_I A(x)\,\mathrm{d}x\right)_{jl} &= \sum_{k=1}^{m} b_{jk} \int_I a_{kl}(x)\,\mathrm{d}x = \int_I \left(\sum_{k=1}^{m} b_{jk} a_{kl}(x)\right) \mathrm{d}x \\ &= \int_I (BA(x))_{jl}\,\mathrm{d}x = \left(\int_I B\,A(x)\,\mathrm{d}x\right)_{jl}, \end{aligned} \tag{I.7}$$

proving (I.5).

(b): One computes, for each $(k, j) \in \{1, \dots, m\} \times \{1, \dots, p\}$,

$$\begin{aligned}\left(\left(\int_I A(x)\,\mathrm{d}x\right) B\right)_{kj} &= \sum_{l=1}^{n} \left(\int_I a_{kl}(x)\,\mathrm{d}x\right) b_{lj} = \int_I \left(\sum_{l=1}^{n} a_{kl}(x) b_{lj}\right) \mathrm{d}x \\ &= \int_I (A(x)B)_{kj}\,\mathrm{d}x = \left(\int_I A(x)\,B\,\mathrm{d}x\right)_{kj}, \end{aligned} \tag{I.8}$$

proving (I.6). ■

J Autonomous ODE

J.1 Equivalence of Autonomous and Nonautonomous ODE

Theorem J.1. *Let* $G \subseteq \mathbb{R} \times \mathbb{K}^n$, $n \in \mathbb{N}$, *and* $f : G \longrightarrow \mathbb{K}^n$. *Then the nonautonomous ODE*

$$y' = f(x, y) \tag{J.1}$$

is equivalent to the autonomous ODE

$$y' = g(y), \tag{J.2}$$

where

$$g : \mathbb{R} \times G \longrightarrow \mathbb{K}^{n+1}, \quad g(y_1, \dots, y_{n+1}) := \big(1, f(y_1, y_2, \dots, y_{n+1})\big), \tag{J.3}$$

in the following sense:

(a) *If* $\phi : I \longrightarrow \mathbb{K}^n$ *is a solution to* (J.1)*, then* $\psi : I \longrightarrow \mathbb{K}^{n+1}$*,* $\psi(x) := (x, \phi(x))$*, is a solution to* (J.2)*.*

(b) *If* $\psi : I \longrightarrow \mathbb{K}^{n+1}$ *is a solution to* (J.2) *with the property*

$$\underset{x_0 \in I}{\exists} \quad \psi_1(x_0) = x_0, \tag{J.4}$$

then $\phi : I \longrightarrow \mathbb{K}^n$*,* $\phi(x) := (\psi_2(x), \dots, \psi_{n+1}(x))$*, is a solution to* (J.1)*.*

Proof. (a): If $\phi : I \longrightarrow \mathbb{K}^n$ is a solution to (J.1) and $\psi : I \longrightarrow \mathbb{K}^{n+1}$, $\psi(x) := (x, \phi(x))$, then

$$\underset{x \in I}{\forall} \quad \psi'(x) = (1, \phi'(x)) = \big(1, f(x, \phi(x))\big) = g(x, \phi(x)) = g(\psi(x)), \tag{J.5}$$

showing ψ is a solution to (J.2).

(b): If $\psi : I \longrightarrow \mathbb{K}^{n+1}$ is a solution to (J.2) with the property (J.4) and $\phi : I \longrightarrow \mathbb{K}^n$, $\phi(x) := (\psi_2(x), \dots, \psi_{n+1}(x))$, then (J.4) implies $\psi_1(x) = x$ for each $x \in I$ and, thus,

$$\underset{x \in I}{\forall} \quad \phi'(x) = (\psi_2'(x), \dots, \psi_{n+1}'(x)) = f(x, \psi_2(x), \dots, \psi_{n+1}(x)) = f(x, \phi(x)), \tag{J.6}$$

showing ϕ is a solution to (J.1). ■

While Th. J.1 is somewhat striking and of theoretical interest, it has few useful applications in practise, due to the unbounded first component of solutions to (J.2) (cf. Rem. 5.2).

J.2 Integral for ODE with Discontinuous Right-Hand Side

The following Example J.2, provided by Anton Sporrer, shows Lem. 5.19 becomes false if the hypothesis that every initial value problem for the considered ODE $y' = f(y)$ has at least one solution is omitted:

Example J.2. Consider

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f(y) := \begin{cases} 0 & \text{for } y \in \mathbb{Q}, \\ 1 & \text{for } y \in \mathbb{R} \setminus \mathbb{Q}, \end{cases} \tag{J.7}$$

and the autonomous ODE $y' = f(y)$. If $(x_0, y_0) \in \mathbb{R} \times \mathbb{Q}$, then the initial value problem $y(x_0) = y_0$ has the unique solution $\phi : \mathbb{R} \longrightarrow \mathbb{R}$, $\phi \equiv y_0 \in \mathbb{Q}$. However, if $(x_0, y_0) \in \mathbb{R} \times (\mathbb{R} \setminus \mathbb{Q})$, then the initial value problem $y(x_0) = y_0$ has no solution. Since $y' = f(y)$ has only constant solutions, *every* function $E : \mathbb{R} \longrightarrow \mathbb{R}$ is an integral for this ODE according to Def. 5.18. However, not every differentiable function $E : \mathbb{R} \longrightarrow \mathbb{R}$ satisfies (5.15): For example, if $E(y) := y$, then $E' \equiv 1$, i.e.

$$\underset{y \in \mathbb{R} \setminus \mathbb{Q}}{\forall} \quad E'(y) f(y) = 1 \neq 0, \tag{J.8}$$

showing that Lem. 5.19 does not hold for $y' = f(y)$ with f according to (J.7).

K Polar Coordinates

Recall the following functions, used in *polar coordinates* of the plane:

$$r : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{R}^+, \quad r(y_1, y_2) := \sqrt{y_1^2 + y_2^2}, \tag{K.1a}$$

$$\varphi : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow [0, 2\pi[, \quad \varphi(y_1, y_2) := \begin{cases} 0 & \text{for } y_2 = 0,\ y_1 > 0, \\ \operatorname{arccot}(y_1/y_2) & \text{for } y_2 > 0, \\ \pi & \text{for } y_2 = 0,\ y_1 < 0, \\ \pi + \operatorname{arccot}(y_1/y_2) & \text{for } y_2 < 0. \end{cases} \tag{K.1b}$$

Theorem K.1. *Consider $f : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{R}^2$ and the corresponding $\mathbb{R}^2$-valued autonomous ODE*

$$y_1' = f_1(y_1, y_2), \tag{K.2a}$$
$$y_2' = f_2(y_1, y_2), \tag{K.2b}$$

together with its polar coordinate version

$$r' = g_1(r, \varphi), \tag{K.3a}$$
$$\varphi' = g_2(r, \varphi), \tag{K.3b}$$

where $g : \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}^2$,

$$g_1 : \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R},$$
$$g_1(r, \varphi) := f_1(r\cos\varphi, r\sin\varphi)\ \cos\varphi + f_2(r\cos\varphi, r\sin\varphi)\ \sin\varphi, \tag{K.4a}$$
$$g_2 : \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R},$$
$$g_2(r, \varphi) := \frac{1}{r}\Big(f_2(r\cos\varphi, r\sin\varphi)\ \cos\varphi - f_1(r\cos\varphi, r\sin\varphi)\ \sin\varphi\Big). \tag{K.4b}$$

Let $\mu : I \longrightarrow \mathbb{R}^2$ be a solution to (K.3)*.*

(a) *Then*

$$\phi : I \longrightarrow \mathbb{R}^2, \quad \phi(x) := \big(\mu_1(x)\ \cos\mu_2(x),\ \mu_1(x)\ \sin\mu_2(x)\big), \tag{K.5}$$

is a solution to (K.2)*.*

(b) *If μ satisfies the initial condition*

$$\mu_1(0) = \rho, \quad \rho \in \mathbb{R}^+, \tag{K.6a}$$
$$\mu_2(0) = \tau, \quad \tau \in \mathbb{R}, \tag{K.6b}$$

and if

$$\eta_1 = \rho\cos\tau, \tag{K.7a}$$
$$\eta_2 = \rho\sin\tau, \tag{K.7b}$$

then ϕ satisfies the initial condition

$$\phi_1(0) = \eta_1, \tag{K.8a}$$
$$\phi_2(0) = \eta_2. \tag{K.8b}$$

Note that $\rho > 0$ implies $(\eta_1, \eta_2) \neq (0,0)$, and that, for $(\rho, \tau) \in \mathbb{R}^+ \times [0, 2\pi[$, (K.7) *is equivalent to*

$$r(\eta_1, \eta_2) = \rho, \tag{K.9a}$$
$$\varphi(\eta_1, \eta_2) = \tau \tag{K.9b}$$

(cf. the computations of $\phi^{-1} \circ \phi$ and of $\phi \circ \phi^{-1}$ in [Phi13b, Ex. 4.19]*).*

Proof. Exercise. ■

Example K.2. Consider the autonomous ODE (K.2) with

$$f_1 : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{R},$$
$$f_1(y_1, y_2) := y_1 \big(1 - r(y_1, y_2)\big) - \frac{y_2 \,(r(y_1, y_2) - y_1)}{2\, r(y_1, y_2)}, \tag{K.10a}$$
$$f_2 : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{R},$$
$$f_2(y_1, y_2) := y_2 \big(1 - r(y_1, y_2)\big) + \frac{y_1 \,(r(y_1, y_2) - y_1)}{2\, r(y_1, y_2)}, \tag{K.10b}$$

where r is the radial polar coordinate function as defined in (K.1a). Using $g : \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}^2$ as defined in (K.4), one obtains, for each $(\rho, \varphi) \in \mathbb{R}^+ \times \mathbb{R}$,

$$\begin{aligned} g_1(\rho, \varphi) &= \rho \cos\varphi \,(1 - \rho)\, \cos\varphi - \frac{\rho \sin\varphi \,(\rho - \rho\cos\varphi)\, \cos\varphi}{2\rho} \\ &\quad + \rho \sin\varphi \,(1-\rho)\, \sin\varphi + \frac{\rho\cos\varphi\,(\rho - \rho\cos\varphi)\, \sin\varphi}{2\rho} \\ &= \rho\,(1-\rho), \end{aligned} \tag{K.11a}$$

$$\begin{aligned} g_2(\rho, \varphi) &= \sin\varphi \,(1-\rho)\, \cos\varphi + \frac{\cos\varphi\,(\rho - \rho\cos\varphi)\, \cos\varphi}{2\rho} \\ &\quad - \cos\varphi\,(1-\rho)\, \sin\varphi + \frac{\sin\varphi\,(\rho - \rho\cos\varphi)\, \sin\varphi}{2\rho} \\ &= \frac{1 - \cos\varphi}{2}, \end{aligned} \tag{K.11b}$$

such that the autonomous ODE

$$r' = r\,(1 - r), \tag{K.12a}$$
$$\varphi' = \frac{1 - \cos\varphi}{2} \overset{\text{[Phi13a, (D.1c)]}}{=} \sin^2 \frac{\varphi}{2}, \tag{K.12b}$$

is the polar coordinate version of (K.2) as defined in Th. K.1.

Claim 1. The general solution to

$$r' = p(r), \quad p : \mathbb{R}^+ \longrightarrow \mathbb{R}, \quad p(r) := r\,(1-r), \tag{K.13}$$

is

$$Y_p : D_{p,0} \longrightarrow \mathbb{R}^+, \quad Y_p(x,\rho) := \frac{\rho}{\rho + (1-\rho)\,e^{-x}}, \tag{K.14}$$

where

$$D_{p,0} = \big(\mathbb{R}\times]0,1]\big) \cup \left\{(x,\rho) : \rho > 1,\, x \in \left] -\ln\frac{\rho}{\rho-1}, \infty\right[\right\}. \tag{K.15}$$

Proof. The initial condition is satisfied, since

$$\forall_{\rho\in\mathbb{R}^+} \quad Y_p(0,\rho) = \frac{\rho}{\rho + (1-\rho)} = \rho. \tag{K.16}$$

The ODE is satisfied, since

$$\forall_{(x,\rho)\in D_{p,0}} \quad Y_p'(x,\rho) = \frac{\rho(1-\rho)e^{-x}}{\big(\rho + (1-\rho)\,e^{-x}\big)^2} = Y_p(x,\rho)\,\big(1 - Y_p(x,\rho)\big). \tag{K.17}$$

To verify the form of $D_{p,0}$, we note that the denominator in (K.14) is positive for each $x \in \mathbb{R}$ if $0 < \rho \leq 1$. If $\rho > 1$, then the function $a : \mathbb{R} \longrightarrow \mathbb{R}$, $a(x) := \rho + (1-\rho)\,e^{-x}$, is strictly increasing (note $a'(x) = (\rho - 1)\,e^{-x} > 0$) and has a unique zero at $x = -\ln\frac{\rho}{\rho-1}$. Thus $\lim_{x\downarrow -\ln\frac{\rho}{\rho-1}} Y(x,\rho) = \infty$, proving the maximality of $Y(\cdot,\rho)$. ▲

Claim 2. Letting, for each $k \in \mathbb{Z}$,

$$\begin{aligned}
x_0(\tau) &:= \frac{2\cos\tau + 2}{\sin\tau} \quad \text{for} \quad \tau \in \mathbb{R}\setminus\{l\pi : l \in \mathbb{Z}\}, && \text{(K.18a)}\\
R_k &:=]0,\pi[+ 2k\pi, && \text{(K.18b)}\\
L_k &:=]-\pi,0[+ 2k\pi, && \text{(K.18c)}\\
A_0 &:= \mathbb{R}\times\{2k\pi : k \in \mathbb{Z}\}, && \text{(K.18d)}\\
A_{0,k} &:= \mathbb{R}^-\times\{\pi + 2k\pi\}, && \text{(K.18e)}\\
B_{0,k} &:= \mathbb{R}^+\times\{\pi + 2k\pi\}, && \text{(K.18f)}\\
A_{1,k} &:= \big\{(x,\tau)\in\mathbb{R}^2 : x \in]-\infty, x_0(\tau)[,\, \tau \in R_k\big\}, && \text{(K.18g)}\\
A_{2,k} &:= \big\{(x,\tau)\in\mathbb{R}^2 : x \in]x_0(\tau), \infty[,\, \tau \in L_k\big\}, && \text{(K.18h)}\\
B_{1,k} &:= \big\{(x,\tau)\in\mathbb{R}^2 : x \in]x_0(\tau), \infty[,\, \tau \in R_k\big\}, && \text{(K.18i)}\\
B_{2,k} &:= \big\{(x,\tau)\in\mathbb{R}^2 : x \in]-\infty, x_0(\tau)[,\, \tau \in L_k\big\}, && \text{(K.18j)}\\
C_{1,k} &:= \big\{(x,\tau)\in\mathbb{R}^2 : x = x_0(\tau),\, \tau \in R_k\big\}, && \text{(K.18k)}\\
C_{2,k} &:= \big\{(x,\tau)\in\mathbb{R}^2 : x = x_0(\tau),\, \tau \in L_k\big\}, && \text{(K.18l)}
\end{aligned}$$

the general solution to

$$\varphi' = q(\varphi), \quad q : \mathbb{R} \longrightarrow \mathbb{R}, \quad q(\varphi) := \frac{1-\cos\varphi}{2} = \sin^2\frac{\varphi}{2}, \tag{K.19}$$

is

$$Y_q : \mathbb{R}^2 \longrightarrow \mathbb{R}^+,$$

$$Y_q(x,\tau) := \begin{cases} \tau & \text{for } (x,\tau) \in A_0, \\ 2\left(k\pi + \arctan\left(-\frac{2}{x}\right)\right) & \text{for } (x,\tau) \in A_{0,k},\ k \in \mathbb{Z}, \\ 2\left((k+1)\pi + \arctan\left(-\frac{2}{x}\right)\right) & \text{for } (x,\tau) \in B_{0,k},\ k \in \mathbb{Z}, \\ \pi + 2k\pi & \text{for } (x,\tau) = (0, \pi + 2k\pi),\ k \in \mathbb{Z}, \\ 2\left(k\pi + \arctan\left(\frac{2\,\sin\tau}{2\,\cos\tau - x\,\sin\tau + 2}\right)\right) & \text{for } (x,\tau) \in A_{1,k} \cup A_{2,k},\ k \in \mathbb{Z}, \\ \pi + 2k\pi & \text{for } (x,\tau) \in C_{1,k},\ k \in \mathbb{Z}, \\ 2\left((k+1)\pi + \arctan\left(\frac{2\,\sin\tau}{2\,\cos\tau - x\,\sin\tau + 2}\right)\right) & \text{for } (x,\tau) \in B_{1,k},\ k \in \mathbb{Z}, \\ -\pi + 2k\pi & \text{for } (x,\tau) \in C_{2,k},\ k \in \mathbb{Z}, \\ 2\left((k-1)\pi + \arctan\left(\frac{2\,\sin\tau}{2\,\cos\tau - x\,\sin\tau + 2}\right)\right) & \text{for } (x,\tau) \in B_{2,k},\ k \in \mathbb{Z}. \end{cases} \tag{K.20}$$

Proof. One observes that Y_q is well-defined, since

$$\mathbb{R}^2 = A_0 \,\dot\cup \bigcup_{k\in\mathbb{Z}} \Big(\{(0, \pi + 2k\pi)\} \,\dot\cup\, A_{0,k} \,\dot\cup\, B_{0,k} \,\dot\cup\, A_{1,k} \,\dot\cup\, A_{2,k} \,\dot\cup\, B_{1,k} \,\dot\cup\, B_{2,k} \,\dot\cup\, C_{1,k} \,\dot\cup\, C_{2,k} \Big) \tag{K.21}$$

and, introducing the auxiliary function

$$\Delta : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad \Delta(x,\tau) := 2\,\cos\tau - x\,\sin\tau + 2, \tag{K.22}$$

one has

$$\Delta(x,\tau) \neq 0 \quad \text{for each } (x,\tau) \in A_{1,k} \cup A_{2,k} \cup B_{1,k} \cup B_{2,k},\ k \in \mathbb{Z}. \tag{K.23}$$

It remains to show that, for each $\tau \in \mathbb{R}$, the function $x \mapsto Y_q(x,\tau)$ is differentiable on $\mathbb{R}$, satisfying

$$\underset{x\in\mathbb{R}}{\forall} \quad Y_q'(x,\tau) = \frac{1 - \cos Y_q(x,\tau)}{2}, \tag{K.24}$$

and the initial condition

$$Y_q(0,\tau) = \tau. \tag{K.25}$$

The initial condition (K.25) is satisfied, since

$$\underset{\tau\in\{k\pi:\,k\in\mathbb{Z}\}}{\forall} \quad Y_q(0,\tau) = \tau, \tag{K.26a}$$

$$\begin{aligned} \underset{\tau\in R_k\cup L_k,\,k\in\mathbb{Z}}{\forall} \quad Y_q(0,\tau) &= 2k\pi + 2\arctan\frac{2\,\sin\tau}{2\,\cos\tau + 2} \overset{\text{[Phi13a, (D.1d)]}}{=} 2k\pi + 2\arctan\left(\tan\frac{\tau}{2}\right) \\ &= 2k\pi + 2\left(\frac{\tau}{2} - k\pi\right) = \tau. \end{aligned} \tag{K.26b}$$

Next, we show that, for each $\tau \in \mathbb{R}$, the function $x \mapsto Y_q(x,\tau)$ is differentiable on $\mathbb{R}$ and satisfies (K.24): For $\tau \in \{2k\pi : k \in \mathbb{Z}\}$, $x \mapsto Y_q(x,\tau)$ is constant, i.e. differentiability is clear, and

$$\underset{x\in\mathbb{R}}{\forall} \quad \frac{1 - \cos Y_q(x,\tau)}{2} = \frac{1 - \cos(2k\pi)}{2} = 0 = Y_q'(x,\tau) \tag{K.27}$$

proves (K.24).

For each $\tau \in \{2(k+1)\pi : k \in \mathbb{Z}\}$, differentiability is clear in each $x \in \mathbb{R} \setminus \{0\}$. Moreover,

$$\underset{x\in\mathbb{R}\setminus\{0\}}{\forall} \quad \left(2\arctan\left(-\frac{2}{x}\right)\right)' = \frac{4}{x^2}\,\frac{1}{1+\frac{4}{x^2}} = \frac{4}{4+x^2}, \tag{K.28}$$

and, thus, for each $x \in \mathbb{R} \setminus \{0\}$,

$$\begin{aligned} \frac{1-\cos Y_q(x,\tau)}{2} &= \frac{1}{2} - \frac{1}{2}\cos\left(2\arctan\left(-\frac{2}{x}\right)\right) \overset{\text{[Phi13a, (D.1e)]}}{=} \frac{1}{2} - \frac{1}{2}\,\frac{1-\frac{4}{x^2}}{1+\frac{4}{x^2}} \\ &= \frac{1}{2} - \frac{1}{2}\,\frac{x^2-4}{x^2+4} = \frac{8}{2(4+x^2)} = \frac{4}{4+x^2} \overset{\text{(K.28)}}{=} Y_q'(x,\tau), \end{aligned} \tag{K.29}$$

proving (K.24) for each $x \in \mathbb{R} \setminus \{0\}$. It remains to consider $x = 0$. One has, by L'Hôpital's rule [Phi13a, Th. 9.23(a)],

$$\begin{aligned} \lim_{x\uparrow 0} \frac{Y_q(0,\tau) - Y_q(x,\tau)}{0-x} &= \lim_{x\uparrow 0} \frac{\pi + 2k\pi - 2\left(k\pi + \arctan\left(-\frac{2}{x}\right)\right)}{-x} \\ &\overset{\text{[Phi13a, (9.28)],(K.28)}}{=} \lim_{x\uparrow 0} \frac{-\frac{4}{4+x^2}}{-1} = 1 \end{aligned} \tag{K.30}$$

and

$$\begin{aligned} \lim_{x\downarrow 0} \frac{Y_q(0,\tau) - Y_q(x,\tau)}{0-x} &= \lim_{x\downarrow 0} \frac{\pi + 2k\pi - 2\left((k+1)\pi + \arctan\left(-\frac{2}{x}\right)\right)}{-x} \\ &\overset{\text{[Phi13a, (9.28)],(K.28)}}{=} \lim_{x\downarrow 0} \frac{-\frac{4}{4+x^2}}{-1} = 1, \end{aligned} \tag{K.31}$$

showing $x \mapsto Y_q(x,\tau)$ to be differentiable in $x = 0$ with $Y_q'(0,\tau) = 1$. Due to

$$\frac{1-\cos(\pi + 2k\pi)}{2} = 1 = Y_q'(0, \pi + 2k\pi), \tag{K.32}$$

(K.24) also holds.

For each $\tau \in R_k \cup L_k$, the differentiability is clear in each $x \in \mathbb{R} \setminus \{x_0(\tau)\}$. Moreover, recalling $\Delta(x,\tau)$ from (K.22), one has, for each $x \in \mathbb{R} \setminus \{x_0(\tau)\}$,

$$\left(2\arctan\left(\frac{2\sin\tau}{\Delta(x,\tau)}\right)\right)' = \frac{4(\sin\tau)^2}{(\Delta(x,\tau))^2}\,\frac{1}{1+\frac{4(\sin\tau)^2}{(\Delta(x,\tau))^2}} = \frac{4(\sin\tau)^2}{4(\sin\tau)^2 + (\Delta(x,\tau))^2}, \tag{K.33}$$

and, thus, for each $x \in \mathbb{R} \setminus \{x_0(\tau)\}$,

$$\begin{aligned} \frac{1-\cos Y_q(x,\tau)}{2} &= \frac{1}{2} - \frac{1}{2}\cos\left(2\arctan\left(\frac{2\sin\tau}{\Delta(x,\tau)}\right)\right) \overset{\text{[Phi13a, (D.1e)]}}{=} \frac{1}{2} - \frac{1}{2}\,\frac{1-\frac{4(\sin\tau)^2}{(\Delta(x,\tau))^2}}{1+\frac{4(\sin\tau)^2}{(\Delta(x,\tau))^2}} \\ &= \frac{1}{2} - \frac{1}{2}\,\frac{(\Delta(x,\tau))^2 - 4(\sin\tau)^2}{(\Delta(x,\tau))^2 + 4(\sin\tau)^2} = \frac{8(\sin\tau)^2}{2(4(\sin\tau)^2 + (\Delta(x,\tau))^2)} \\ &= \frac{4(\sin\tau)^2}{4(\sin\tau)^2 + (\Delta(x,\tau))^2} \overset{\text{(K.33)}}{=} Y_q'(x,\tau), \end{aligned} \tag{K.34}$$

proving (K.24) for each $x \in \mathbb{R} \setminus \{x_0(\tau)\}$. It remains to consider $x = x_0(\tau)$. For $\tau \in R_k$, we have $\sin\tau > 0$ and $x_0(\tau) > 0$, and, thus, by L'Hôpital's rule [Phi13a, Th. 9.23(a)],

$$\lim_{x \uparrow x_0(\tau)} \frac{Y_q\big(x_0(\tau),\tau\big) - Y_q(x,\tau)}{x_0(\tau) - x} = \lim_{x \uparrow x_0(\tau)} \frac{\pi + 2k\pi - 2\left(k\pi + \arctan\left(\frac{2\sin\tau}{2\cos\tau - x\sin\tau + 2}\right)\right)}{x_0(\tau) - x}$$
$$\overset{\text{[Phi13a, (9.28)],(K.33)}}{=} \lim_{x \uparrow x_0(\tau)} \frac{-\frac{4(\sin\tau)^2}{4(\sin\tau)^2 + (\Delta(x,\tau))^2}}{-1} = 1 \tag{K.35}$$

and

$$\lim_{x \downarrow x_0(\tau)} \frac{Y_q\big(x_0(\tau),\tau\big) - Y_q(x,\tau)}{x_0(\tau) - x} = \lim_{x \downarrow x_0(\tau)} \frac{\pi + 2k\pi - 2\left((k+1)\pi + \arctan\left(\frac{2\sin\tau}{2\cos\tau - x\sin\tau + 2}\right)\right)}{x_0(\tau) - x}$$
$$\overset{\text{[Phi13a, (9.28)],(K.33)}}{=} \lim_{x \downarrow x_0(\tau)} \frac{-\frac{4(\sin\tau)^2}{4(\sin\tau)^2 + (\Delta(x,\tau))^2}}{-1} = 1 \tag{K.36}$$

showing $x \mapsto Y_q(x,\tau)$ to be differentiable in $x = x_0(\tau)$ with $Y_q'(x_0(\tau),\tau) = 1$. Due to

$$\frac{1 - \cos(\pi + 2k\pi)}{2} = 1 = Y_q'(x_0(\tau),\tau), \tag{K.37}$$

(K.24) also holds. For $\tau \in L_k$, we have $\sin\tau < 0$ and $x_0(\tau) < 0$, and, thus, by L'Hôpital's rule [Phi13a, Th. 9.23(a)],

$$\lim_{x \uparrow x_0(\tau)} \frac{Y_q\big(x_0(\tau),\tau\big) - Y_q(x,\tau)}{x_0(\tau) - x}$$
$$= \lim_{x \uparrow x_0(\tau)} \frac{-\pi + 2k\pi - 2\left((k-1)\pi + \arctan\left(\frac{2\sin\tau}{2\cos\tau - x\sin\tau + 2}\right)\right)}{x_0(\tau) - x}$$
$$\overset{\text{[Phi13a, (9.28)],(K.33)}}{=} \lim_{x \uparrow x_0(\tau)} \frac{-\frac{4(\sin\tau)^2}{4(\sin\tau)^2 + (\Delta(x,\tau))^2}}{-1} = 1 \tag{K.38}$$

and

$$\lim_{x \downarrow x_0(\tau)} \frac{Y_q\big(x_0(\tau),\tau\big) - Y_q(x,\tau)}{x_0(\tau) - x} = \lim_{x \downarrow x_0(\tau)} \frac{-\pi + 2k\pi - 2\left(k\pi + \arctan\left(\frac{2\sin\tau}{2\cos\tau - x\sin\tau + 2}\right)\right)}{x_0(\tau) - x}$$
$$\overset{\text{[Phi13a, (9.28)],(K.33)}}{=} \lim_{x \downarrow x_0(\tau)} \frac{-\frac{4(\sin\tau)^2}{4(\sin\tau)^2 + (\Delta(x,\tau))^2}}{-1} = 1 \tag{K.39}$$

showing $x \mapsto Y_q(x,\tau)$ to be differentiable in $x = x_0(\tau)$ with $Y_q'(x_0(\tau),\tau) = 1$. Due to

$$\frac{1 - \cos(-\pi + 2k\pi)}{2} = 1 = Y_q'(x_0(\tau),\tau), \tag{K.40}$$

(K.24) also holds. ▲

Claim 3. The general solution to (K.2) with f_1, f_2 according to (K.10) is

$$\begin{aligned} Y : D_{f,0} &\longrightarrow \mathbb{R}^2, \\ Y(x,\eta_1,\eta_2) &:= \Big(Y_p\big(x, r(\eta_1,\eta_2)\big) \cos Y_q\big(x,\varphi(\eta_1,\eta_2)\big), \\ &\qquad Y_p\big(x, r(\eta_1,\eta_2)\big) \sin Y_q\big(x,\varphi(\eta_1,\eta_2)\big)\Big), \end{aligned} \tag{K.41}$$

where r and φ are given by (K.1), and

$$\begin{aligned} D_{f,0} = &\Big(\mathbb{R} \times \{\eta \in \mathbb{R}^2 : 0 < \|\eta\|_2 \le 1\}\Big) \\ &\cup \left\{ (x,\eta) \in \mathbb{R} \times \mathbb{R}^2 : \|\eta\|_2 > 1,\ x \in \left] -\ln \frac{\|\eta\|_2}{\|\eta\|_2 - 1}, \infty \right[\right\}. \end{aligned} \tag{K.42}$$

Proof. Since (K.12) is the polar coordinate version of (K.2) with f_1, f_2 according to (K.10), everything follows from combining Th. K.1 with Claims 1 and 2. ▲

Claim 4. The autonomous ODE (K.2) with f_1, f_2 according to (K.10) has $(1,0)$ as its only fixed point, and $(1,0)$ satisfies Def. 5.24(iii) for $x \to \infty$ (even for each $\eta \in \mathbb{R}^2 \setminus \{0\}$) without satisfying Def. 5.24(ii) (i.e. without being positively stable).

Proof. For each $\eta \in \mathbb{R}^2 \setminus \{0\}$, it is $r(\eta) > 0$, and, thus

$$\underset{\eta \in \mathbb{R}^2 \setminus \{0\}}{\forall} \quad \lim_{x\to\infty} Y_p\big(x, r(\eta)\big) = \lim_{x\to\infty} \frac{r(\eta)}{r(\eta) + \big(1 - r(\eta)\big)\, e^{-x}} = 1. \tag{K.43}$$

Fix $\eta \in \mathbb{R}^2 \setminus \{0\}$. If $\varphi(\eta) = 0$, then

$$\lim_{x\to\infty} Y_q\big(x, 0\big) = \lim_{x\to\infty} 0 = 0. \tag{K.44a}$$

If $\varphi(\eta) = \pi$, then

$$\lim_{x\to\infty} Y_q\big(x, \pi\big) = \lim_{x\to\infty} 2\left(\pi + \arctan\left(-\frac{2}{x}\right)\right) = 2(\pi + 0) = 2\pi. \tag{K.44b}$$

If $0 < \varphi(\eta) < \pi$ or $\pi < \varphi(\eta) < 2\pi$, then $\sin\varphi(\eta) \neq 0$ and, thus,

$$\begin{aligned} \lim_{x\to\infty} Y_q\big(x, \varphi(\eta)\big) &= \lim_{x\to\infty} 2\left(\pi + \arctan\left(\frac{2\,\sin\varphi(\eta)}{2\,\cos\varphi(\eta) - x\,\sin\varphi(\eta) + 2}\right)\right) \\ &= 2(\pi + 0) = 2\pi. \end{aligned} \tag{K.44c}$$

Using (K.43) and (K.44) in (K.41) yields

$$\underset{\eta \in \mathbb{R}^2 \setminus \{0\}}{\forall} \quad \lim_{x\to\infty} Y(x,\eta) = (1,0). \tag{K.45}$$

While $(1,0)$ is clearly a fixed point for (K.2) with f_1, f_2 according to (K.10), (K.45) shows that no other $\eta \in \mathbb{R}^2 \setminus \{0\}$ can be a fixed point.

For each $\tau \in]0, \pi[$ and $\eta := (\cos\tau, \sin\tau)$, it is $\varphi(\eta) = \tau$ and $Y_q(0, \varphi(\eta)) = \tau$. Thus, due to (K.44c) and the intermediate value theorem, the continuous function $Y_q(\cdot, \varphi(\eta))$ must attain every value between τ and 2π, in particular, there is $x_\pi > 0$ that $Y_q(x_\tau, \varphi(\eta)) = \pi$ and $Y(x_\tau, \eta) = (\cos\pi, \sin\pi) = (-1, 0)$. Since every neighborhood of $(1, 0)$ contains points $\eta = (\cos\tau, \sin\tau)$ with $\tau \in]0, \pi[$, this shows that $(1, 0)$ does not satisfy Def. 5.24(ii) for $x \geq 0$. ▲

References

[Aul04] Bernd Aulbach. *Gewöhnliche Differenzialgleichungen*, 2nd ed. Spektrum Akademischer Verlag, Heidelberg, Germany, 2004 (German).

[KÖ04] Konrad Königsberger. *Analysis 2*, 5th ed. Springer-Verlag, Berlin, 2004 (German).

[Koe03] Max Koecher. *Lineare Algebra und analytische Geometrie*, 4th ed. Springer-Verlag, Berlin, 2003 (German), 1st corrected reprint.

[Mar04] Nelson G. Markley. *Principles of Differential Equations*. Pure and Applied Mathematics, Wiley-Interscience, Hoboken, NJ, USA, 2004.

[Oss09] E. Ossa. *Topologie*. Vieweg+Teubner, Wiesbaden, Germany, 2009 (German).

[Phi13a] P. Philip. *Calculus I for Computer Science and Statistics Students*. Lecture Notes, Ludwig-Maximilians-Universität, Germany, 2012/2013, available in PDF format at `http://www.math.lmu.de/~philip/publications/lectureNotes/calc1_forInfAndStatStudents.pdf`.

[Phi13b] P. Philip. *Calculus II for Statistics Students*. Lecture Notes, Ludwig-Maximilians-Universität, Germany, 2013, available in PDF format at `http://www.math.lmu.de/~philip/publications/lectureNotes/calc2_forStatStudents.pdf`.

[Pre75] G. Preuss. *Allgemeine Topologie*, 2nd ed. Springer-Verlag, Berlin, 1975 (German).

[Put66] E.J. Putzer. *Avoiding the Jordan Canonical Form in the Discussion of Linear Systems with Constant Coefficients*. The American Mathematical Monthly **73** (1966), No. 1, 2–7.

[Str08] Gernot Stroth. *Lineare Algebra*, 2nd ed. Berliner Studienreihe zur Mathematik, Vol. 7, Heldermann Verlag, Lemgo, Germany, 2008 (German).

[Wal02] Wolfgang Walter. *Analysis 2*, 5th ed. Springer-Verlag, Berlin, 2002 (German).