

# 1 Wiederholung mathematischer Grundlagen

## 1.1 Der Vektorraum $\mathbb{R}^n$

Zunächst betrachten wir die Grundmenge

$$\mathbb{R}^n = \{x = (x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \in \mathbb{R}\}, \quad (1.1)$$

deren Elemente wir **Vektoren** nennen. Manchmal schreibt man die Elemente von  $\mathbb{R}^n$  auch als Spalten, wir bevorzugen jedoch hier die kompaktere Zeilenschreibweise. Das Element  $o = (0, 0, \dots, 0)$  heißt **Nullvektor**.

Um der Menge (1.1) eine Struktur zu geben, führen wir eine **Addition** von Elementen  $x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$  ein, die durch

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

komponentenweise erklärt wird.

Weiter definieren wir für  $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  und  $\alpha \in \mathbb{R}$  durch

$$\alpha \cdot x = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$$

eine **skalare Multiplikation**, die wie die Addition komponentenweise erklärt ist. Reelle Zahlen bezeichnet man daher manchmal als **Skalare**.

Die Menge (1.1) mit der so definierten Addition und skalaren Multiplikation wird als Vektorraum  $\mathbb{R}^n$  bezeichnet. In der Vorlesung werden wir meist nur mit den Vektorräumen  $\mathbb{R}^2$  und  $\mathbb{R}^3$  zu tun haben.

Mit Hilfe von Addition und skalarer Multiplikation lassen sich aus einer Menge von Vektoren  $b_1, b_2, \dots, b_k \in \mathbb{R}^n$  neue Vektoren konstruieren: Ein Vektor  $y \in \mathbb{R}^n$  der Form

$$y = \alpha_1 \cdot b_1 + \alpha_2 \cdot b_2 + \dots + \alpha_k \cdot b_k = \sum_{i=1}^k \alpha_i \cdot b_i$$

mit Skalaren  $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}$  heißt **Linearkombination (LK)** der Vektoren  $b_1, b_2, \dots, b_k$ . Man sagt auch,  $y$  kann aus  $b_1, b_2, \dots, b_k$  linear kombiniert werden.

Damit lassen sich die folgenden wichtigen Begriffe aus der linearen Algebra erklären:

**Definition 1.1** Die Vektoren  $b_1, b_2, \dots, b_k \in \mathbb{R}^n$  heißen **linear unabhängig**, falls der Nullvektor nur durch  $o = 0 \cdot b_1 + 0 \cdot b_2 + \dots + 0 \cdot b_n$  linear kombiniert werden kann, andernfalls heißen sie **linear abhängig**.

$b_1, b_2, \dots, b_k \in \mathbb{R}^n$  heißen **erzeugend**, wenn jeder Vektor  $y \in \mathbb{R}^n$  eine Linearkombination von  $b_1, b_2, \dots, b_k$  ist.

Eine Menge von linear unabhängigen und erzeugenden Vektoren eines Vektorraumes heißt **Basis** des Vektorraums.

In diesem Zusammenhang gelten die folgenden

### Aussagen

- (0) Ein Vektor  $v \in \mathbb{R}^n$  ist genau dann linear unabhängig wenn  $v \neq 0$  gilt.
- (1) Zwei Vektoren  $b_1, b_2 \in \mathbb{R}^n \setminus \{0\}$  sind genau dann linear abhängig, wenn  $b_2$  ein skalares Vielfaches von  $b_1$  ist.
- (2) Eine Menge von  $n + 1$  oder mehr Vektoren im  $\mathbb{R}^n$  ist immer linear abhängig.
- (3) Jede erzeugende Menge des  $\mathbb{R}^n$  besteht aus mindestens  $n$  Vektoren.
- (4) Ist  $B = \{b_1, b_2, \dots, b_n\} \subseteq \mathbb{R}^n$  eine Basis, so ist jeder Vektor  $x \in \mathbb{R}^n$  **eindeutig** darstellbar als Linearkombination

$$x = \alpha_1 \cdot b_1 + \alpha_2 \cdot b_2 + \dots + \alpha_n \cdot b_n$$

der Vektoren  $b_1, b_2, \dots, b_n$ . In diesem Fall heißen die Skalare  $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$  auch **Koeffizienten** von  $x$  bezüglich der Basis  $B$ .

Die speziellen Vektoren  $e_1 = (1, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1)$  bilden eine Basis von  $\mathbb{R}^n$ , die sogenannte **Standardbasis** des  $\mathbb{R}^n$ . Ein beliebiger Vektor  $x = (x_1, x_2, \dots, x_n)$  hat dann die eindeutige Darstellung

$$x = x_1 \cdot e_1 + x_2 \cdot e_2 + \dots + x_n \cdot e_n. \quad (1.2)$$

Die Komponenten  $x_1, x_2, \dots, x_n$  von  $x$  sind also genau die Koeffizienten von  $x$  bezüglich der Standardbasis.

Manchmal ist es nützlich, eine Menge von  $m$  Vektoren des  $\mathbb{R}^n$  als **Matrix**

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

zu schreiben. Genauer handelt es sich hierbei um eine  $(m \times n)$ -Matrix, das heißt eine Matrix mit  $m$  Zeilen und  $n$  Spalten. Die Menge solcher  $(m \times n)$ -Matrizen bildet den Vektorraum  $\mathbb{R}^{m \times n}$  (wieder mit komponentenweiser Addition und skalarer Multiplikation).

**Definition 1.2** Der **Rang** einer Matrix ist definiert als die maximale Anzahl linear unabhängiger Zeilen.

Der Rang einer  $(m \times n)$ -Matrix ist zugleich auch die maximale Anzahl ihrer linear unabhängigen Spalten (aufgefasst als Vektoren des  $\mathbb{R}^m$ ). Insbesondere gilt für eine Matrix  $A \in \mathbb{R}^{m \times n}$

$$\text{Rang}(A) \leq \min\{m, n\}.$$

## 1.2 Skalarprodukt und Abstände im $\mathbb{R}^n$

Um Längen und Winkel im  $\mathbb{R}^n$  zu messen, verwenden wir als Werkzeug das **Standardskalarprodukt**  $\langle \cdot, \cdot \rangle$ , das folgendermaßen definiert ist: Für  $x = (x_1, x_2, \dots, x_n)$ ,  $y = (y_1, y_2, \dots, y_n)$  setzen wir

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i. \quad (1.3)$$

Der  $\mathbb{R}^n$  mit diesem Standardskalarprodukt ist ein **euklidischer Vektorraum**. Man nennt ihn auch euklidischen Standardraum, um klarzumachen, dass das ausgezeichnete Skalarprodukt (1.3) gemeint ist.

Für  $x, y, z \in \mathbb{R}^n$  und  $\lambda, \mu \in \mathbb{R}$  gilt  $\langle x, y \rangle = \langle y, x \rangle$  und

$$\langle z, \lambda \cdot x + \mu \cdot y \rangle = \lambda \langle z, x \rangle + \mu \langle z, y \rangle.$$

Außerdem ist das Skalarprodukt **positiv definit**, das heißt  $\langle x, x \rangle \geq 0$  für beliebiges  $x \in \mathbb{R}^n$ , und

$$\langle x, x \rangle = 0 \iff x = o.$$

Deshalb lässt sich mit Hilfe des Skalarproduktes die **Norm** oder auch **Länge** eines Vektors folgendermaßen definieren:

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}.$$

Vektoren  $x \in \mathbb{R}^n$  der Länge  $\|x\| = 1$  heißen **Einheitsvektoren**.

Sind  $x, y \in \mathbb{R}^n$  und  $\lambda \in \mathbb{R}$ , so gilt für die Norm

$$\begin{aligned} \|\lambda \cdot x\| &= |\lambda| \cdot \|x\|, \\ \|x + y\| &\leq \|x\| + \|y\|. \end{aligned}$$

Für  $x, y \in \mathbb{R}^n$  definiert  $\|x - y\|$  den **Abstand** zwischen  $x$  und  $y$ .

Außerdem gilt die **Cauchy-Schwarzsche Ungleichung (CSU)**

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|.$$

Diese impliziert dass für  $x, y \in \mathbb{R}^n \setminus \{o\}$

$$\frac{|\langle x, y \rangle|}{\|x\| \cdot \|y\|} \leq 1, \quad \text{also} \quad -1 \leq \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \leq 1.$$

Somit lässt sich der **Winkel**  $\angle(x, y) \in [0, \pi]$  zwischen  $x$  und  $y$  definieren durch

$$\cos \angle(x, y) = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}.$$

Zwei Vektoren  $x, y \in \mathbb{R}^n$  heißen **orthogonal**, wenn  $\langle x, y \rangle = 0$  gilt; wir schreiben dafür auch  $x \perp y$ . Dies ist genau dann der Fall wenn  $\angle(x, y) = \pi/2$  oder  $x = o$  oder  $y = o$ . Sind  $x, y \in \mathbb{R}^n$  orthogonale Einheitsvektoren, so nennt man  $x$  und  $y$  auch **orthonormal**. Insbesondere sind die Vektoren  $e_1, e_2, \dots, e_n$  paarweise orthonormal.

Eine Basis aus paarweise orthonormalen Vektoren heißt **Orthonormalbasis**. Ein Beispiel für eine Orthonormalbasis des  $\mathbb{R}^3$  ist die Standardbasis  $e_1, e_2, e_3$ .

### 1.3 Punkte, Geraden und Ebenen im $\mathbb{R}^3$

Zu zwei beliebigen Punkten  $P$  und  $Q$  im dreidimensionalen Raum gibt es genau einen Vektor  $x = \overrightarrow{PQ} \in \mathbb{R}^3$  sodass  $Q = P + x$ . Fixiert man einen Ursprung  $O$  im dreidimensionalen Raum, so kann man jeden Punkt  $P$  eindeutig durch seinen **Ortsvektor**  $p = \overrightarrow{OP} = (p_1, p_2, p_3) \in \mathbb{R}^3$  beschreiben. Die Zahlen  $p_1, p_2, p_3 \in \mathbb{R}$  heißen **Koordinaten** von  $P$ . Ist  $\overrightarrow{OQ} = (q_1, q_2, q_3) \in \mathbb{R}^3$ , so folgt

$$\overrightarrow{PQ} = (q_1 - p_1, q_2 - p_2, q_3 - p_3).$$

Der **Abstand**  $d(P, Q)$  zwischen  $P$  und  $Q$  ist

$$d(P, Q) = \|\overrightarrow{PQ}\| = \|q - p\|.$$

Ganz allgemein gilt für drei Punkte  $P, Q$  und  $R$

$$\overrightarrow{PQ} = \overrightarrow{PR} + \overrightarrow{RQ}.$$

Eine **Gerade**  $g$  im Raum kann folgendermaßen in **Parameterform** beschrieben werden: Man wählt zwei verschiedene Punkte  $P$  und  $Q$  auf  $g$  und betrachtet die Vektoren  $a = \overrightarrow{OP}, b = \overrightarrow{PQ}$ . Dann ist die Menge aller Ortsvektoren auf  $g$  gegeben durch

$$g = \{a + s \cdot b : s \in \mathbb{R}\}.$$

Ist  $E$  eine **Ebene** im Raum, so wählen wir drei Punkte  $P, Q, R$  auf  $E$ , die nicht auf einer Geraden liegen. Setzt man nun  $a = \overrightarrow{OP}, b = \overrightarrow{PQ}$  und  $c = \overrightarrow{PR}$ , so ist die Menge aller Ortsvektoren auf  $E$  gegeben durch

$$E = \{a + s \cdot b + t \cdot c : s, t \in \mathbb{R}\}. \quad (1.4)$$

Beachte, dass  $b$  und  $c$  linear unabhängig sind, da  $P, Q$  und  $R$  nicht auf einer Geraden liegen!

Ein Vektor  $n \in \mathbb{R}^3$  mit  $\langle n, b \rangle = \langle n, c \rangle = 0$  heißt **Normalenvektor** der Ebene  $E$ ; ist außerdem  $\|n\| = 1$ , so heißt  $n$  **Normaleneinheitsvektor**. Im folgenden Abschnitt

werden wir sehen, wie man zu jeder Ebene im  $\mathbb{R}^3$  einen Normalenvektor konstruieren kann.

Wir möchten nun den Winkel zwischen einer Ebene  $E$  mit Normaleneinheitsvektor  $n$  und einer Geraden  $g = \{u + r \cdot v : r \in \mathbb{R}\}$  bestimmen: Bezeichnet  $\theta \in [0, \pi/2]$  den Winkel zwischen  $E$  und  $g$ , so ist  $\pi/2 - \theta$  der Winkel zwischen  $g$  und der Geraden durch den Schnittpunkt von  $E$  und  $g$  mit Richtungsvektor  $n$ ; dieser Winkel ist gegeben durch

$$\cos(\pi/2 - \theta) = \frac{\langle n, v \rangle}{\|v\|}.$$

Wegen  $\sin \theta = \cos(\frac{\pi}{2} - \theta)$  berechnet sich der Winkel  $\theta$  zwischen  $E$  und  $g$  durch

$$\sin \theta = \frac{\langle n, v \rangle}{\|v\|}.$$

Der Normaleneinheitsvektor  $n$  einer Ebene  $E$  ist auch nützlich zur Bestimmung der **Orthogonalprojektion**  $\Pi_E$  auf  $E$ : Seien  $n \in \mathbb{R}^3$  ein Normaleneinheitsvektor von  $E$ , und  $p = (p_1, p_2, p_3)$  der Ortsvektor eines Punktes  $P$ . Dann ist die Gerade

$$g = \{p + r \cdot n : r \in \mathbb{R}\}$$

orthogonal zu  $E$ , und ihr Schnittpunkt mit  $E$  ist genau die Orthogonalprojektion von  $P$  auf  $E$ . Wir suchen also  $r, s, t \in \mathbb{R}$  sodass die Gleichung

$$p + r \cdot n = a + s \cdot b + t \cdot c$$

erfüllt ist. Bildet man das Skalarprodukt jeder Seite mit dem Vektor  $n$ , so erhalten wir die Bedingung

$$\langle p, n \rangle + r = \langle a, n \rangle, \quad \text{also } r = \langle a - p, n \rangle.$$

Die Orthogonalprojektion von  $P$  auf  $E$  besitzt also den Ortsvektor

$$\Pi_E(p) = p + \langle a - p, n \rangle n.$$

Der **Abstand**  $d(P, E)$  des Punktes  $P$  zur Ebene  $E$  ist dann genau die Länge des Vektors  $p - \Pi_E(p) = \langle a - p, n \rangle n$ , also

$$d(P, E) = |\langle a - p, n \rangle|.$$

Beachten Sie, dass diese Formel nur gilt, wenn  $n$  ein **Einheitsvektor** ist.

Alternativ zur Parameterform (1.4) kann die Menge der Ortsvektoren  $x \in \mathbb{R}^3$  einer Ebene  $E$  auch durch die Gleichung

$$\langle x, n \rangle = d \tag{1.5}$$

beschrieben werden, wobei  $n \in \mathbb{R}^3$  der Normaleneinheitsvektor von  $E$  ist, der vom Ursprung in Richtung der Ebene zeigt, und  $d \geq 0$  der Abstand vom Ursprung zur Ebene ist. Die Darstellung einer Ebene  $E$  durch die Gleichung (1.5) heißt **Hesse-Normalform** von  $E$ .

## 1.4 Das Vektorprodukt im $\mathbb{R}^3$

Im Vektorraum  $\mathbb{R}^3$  ist außer dem Skalarprodukt eine weitere Produktbildung möglich, die zwei **Vektoren** wieder einen **Vektor** zuordnet. Diese Produktbildung ermöglicht insbesondere, auf einfache Weise zu einer Ebene einen Normalenvektor zu finden.

Für zwei Vektoren  $x = (x_1, x_2, x_3), y = (y_1, y_2, y_3) \in \mathbb{R}^3$  setzen wir

$$x \times y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

Der Vektor  $x \times y$  heißt **Vektorprodukt** (oder **Kreuzprodukt**) der Vektoren  $x$  und  $y$ .

**Achtung:** Das Vektorprodukt ist eine spezielle Konstruktion im  $\mathbb{R}^3$ ; im  $\mathbb{R}^n, n \neq 3$ , gibt es **kein Vektorprodukt**.

Das Vektorprodukt besitzt die folgenden wichtigen

**Eigenschaften:**

- (1)  $x \times y \perp x$  und  $x \times y \perp y$
- (2)  $x \times y = -(y \times x)$
- (3)  $x, y$  linear abhängig  $\iff x \times y = 0$
- (4)  $\|x \times y\| = \|x\|\|y\| \sin \angle(x, y)$ ; geometrisch ist dies der Flächeninhalt des von  $x$  und  $y$  aufgespannten Parallelogramms.

Für das mehrfache Vektorprodukt dreier Vektoren  $x, y, z \in \mathbb{R}^3$  gilt

$$x \times (y \times z) = \langle x, z \rangle y - \langle x, y \rangle z;$$

insbesondere ist zu beachten, dass im Allgemeinen

$$(x \times y) \times z \neq x \times (y \times z).$$

Weiter gilt für  $x, y, z, w \in \mathbb{R}^3$  die **Formel von Lagrange**

$$\langle x \times y, z \times w \rangle = \langle x, z \rangle \langle y, w \rangle - \langle x, w \rangle \langle y, z \rangle.$$

Ist  $E = \{a + s \cdot b + t \cdot c : s, t \in \mathbb{R}\}$  eine Ebene im  $\mathbb{R}^3$ , so ist der Vektor  $b \times c$  wegen Eigenschaft (1) ein Normalenvektor von  $E$ ; einen Normaleneinheitsvektor  $n$  von  $E$  erhält man durch Normierung

$$n = \frac{b \times c}{\|b \times c\|}.$$

Der Normaleneinheitsvektor einer Ebene ist bis auf das Vorzeichen eindeutig bestimmt.

Sind  $x, y, z$  Vektoren im  $\mathbb{R}^3$ , die nicht in einer gemeinsamen Ebene liegen, so heißt die Menge

$$\{\lambda x + \mu y + \nu z : \lambda, \mu, \nu \in [0, 1]\}$$

das von  $x, y$  und  $z$  aufgespannte **Parallelepipid** (oder auch **Spat**). Das Volumen  $V$  eines solchen Parallelepipids ist gegeben als Produkt von Grundfläche und Höhe. Da die Grundfläche genau das von  $x$  und  $y$  aufgespannte Parallelogramm ist, und die Höhe gleich der Länge der Projektion von  $z$  auf den Normaleneinheitsvektor  $n$  der von  $x$  und  $y$  aufgespannten Ebene ist, folgt

$$V = \|x \times y\| \cdot \langle z, n \rangle = \|x \times y\| \cdot \left\langle z, \frac{x \times y}{\|x \times y\|} \right\rangle = \langle z, x \times y \rangle.$$

Diese Zahl

$$\langle z, x \times y \rangle = \langle x \times y, z \rangle$$

heißt **Spatprodukt** und kann also geometrisch interpretiert werden als Volumen des von  $x, y$  und  $z$  gebildeten Parallelepipeds; dieses hat ein negatives Vorzeichen, falls  $x, y$  und  $z$  kein Rechtssystem bilden. Da bei zyklischer Vertauschung der drei Vektoren sich das Volumen nicht ändert, gelten die Gleichungen

$$\langle x \times y, z \rangle = \langle y \times z, x \rangle = \langle z \times x, y \rangle.$$

Bezeichnet  $(x \times y)_i$  die  $i$ -te Komponente des Vektorprodukts  $x \times y$ ,  $i = 1, 2, 3$ , und  $e_i$  den  $i$ -ten Standardbasisvektor des  $\mathbb{R}^3$ , so gilt speziell

$$(x \times y)_i = \langle x \times y, e_i \rangle.$$

## 1.5 Vektorwertige Funktionen

Eine Gerade  $g = \{a + t \cdot b : t \in \mathbb{R}\}$  im  $\mathbb{R}^3$  kann auch betrachtet werden als Bildmenge einer Funktion

$$x : \mathbb{R} \rightarrow \mathbb{R}^3, \quad t \mapsto x(t) = a + t \cdot b = (a_1 + tb_1, a_2 + tb_2, a_3 + tb_3) \quad (1.6)$$

mit Wertebereich im  $\mathbb{R}^3$ . Allgemeiner kann man beliebige **vektorwertige Funktionen** einer reellen Variablen betrachten: Ist  $I \subseteq \mathbb{R}$  ein Intervall, so ist die Funktion

$$x : I \rightarrow \mathbb{R}^n, \quad t \mapsto x(t) = (x_1(t), x_2(t), \dots, x_n(t)),$$

eine vektorwertige (oder genauer  **$n$ -wertige**) Funktion. Die reellwertigen Funktionen  $x_1, x_2, \dots, x_n : I \rightarrow \mathbb{R}$  heißen **Koordinatenfunktionen**.

**Beispiel 1.3 (Helix)** Es sei  $r > 0$  und  $c \in \mathbb{R}$ . Setze

$$x : \mathbb{R} \rightarrow \mathbb{R}^3, \quad x(t) = (r \cos t, r \sin t, ct). \quad (1.7)$$

Die Bildmenge von  $x$  ist eine sogenannte **Helix** (oder auch **Schraubenlinie**). Für  $c = 0$  ist die Bildmenge ein Kreis in der  $(x_1, x_2)$ -Ebene, der unendlich oft durchlaufen wird; im Allgemeinen liegt die Kurve auf einem Zylinder vom Radius  $r$  um die  $x_3$ -Achse. Die Zahl  $2\pi \cdot c$  heißt aus offensichtlichen Gründen **Ganghöhe** der Helix.

Im Folgenden bezeichne  $I \subseteq \mathbb{R}$  ein **offenes** Intervall und  $f : I \rightarrow \mathbb{R}$  eine reellwertige Funktion.

**Definition 1.4** Ist  $t_0 \in I$ , so heißt  $f_0 \in \mathbb{R}$  **Grenzwert** von  $f$  für  $t$  gegen  $t_0$ , wenn für jede Folge  $(t_k)_{k \in \mathbb{N}} \subseteq I$  mit  $\lim_{k \rightarrow \infty} t_k = t_0$

$$\lim_{k \rightarrow \infty} f(t_k) = f_0$$

gilt. In diesem Fall schreiben wir  $f_0 = \lim_{t \rightarrow t_0} f(t)$ .

$f$  heißt **stetig** an der Stelle  $t = t_0$ , falls  $\lim_{t \rightarrow t_0} f(t) = f(t_0)$ .  $f$  heißt stetig, falls  $f$  für alle  $t \in I$  stetig ist.

Eine vektorwertige Funktion ist stetig genau dann, wenn alle Koordinatenfunktionen es sind.

**Beispiel 1.5** Die vektorwertigen Funktionen (1.6) und (1.7) sind stetig.

Die vektorwertige Funktion

$$x : \mathbb{R} \rightarrow \mathbb{R}^3, \quad t \mapsto x(t) = \begin{cases} t \cdot e_1 & \text{für } t < 0 \\ e_3 + t \cdot e_2 & \text{für } t \geq 0 \end{cases}$$

ist nicht stetig an der Stelle  $t = 0$ .

**Definition 1.6**  $f$  heißt **differenzierbar** an der Stelle  $t_0 \in I$ , falls der Grenzwert

$$f'(t_0) = \frac{d}{dt} f(t_0) = \lim_{t \rightarrow t_0} \frac{f(t) - f(t_0)}{t - t_0}$$

existiert. In diesem Fall heißt  $f'(t_0)$  die **Ableitung** von  $f$  an der Stelle  $t_0$ .  $f$  heißt differenzierbar, falls  $f$  für alle  $t \in I$  differenzierbar ist.

Eine vektorwertige Funktion

$$x : I \rightarrow \mathbb{R}^n, \quad t \mapsto x(t) = (x_1(t), x_2(t), \dots, x_n(t)),$$

ist differenzierbar, wenn alle Koordinatenfunktionen es sind. In diesem Falle ist die Ableitung an der Stelle  $t_0 \in I$  gegeben durch

$$x'(t_0) = (x'_1(t_0), x'_2(t_0), \dots, x'_n(t_0)).$$

**Bemerkung 1.7** Eine differenzierbare (vektorwertige) Funktion auf einem **abgeschlossenen** Intervall  $[\alpha, \beta] \subseteq \mathbb{R}$  ist die Einschränkung einer differenzierbaren (vektorwertigen) Funktion auf einem offenen Intervall, das  $[\alpha, \beta]$  enthält.

Seien nun  $I \subseteq \mathbb{R}$  ein beliebiges Intervall,  $x, y : I \rightarrow \mathbb{R}^n$  differenzierbare vektorwertige Funktionen, und  $f : I \rightarrow \mathbb{R}$  differenzierbar. Für die Ableitung gelten die folgenden wichtigen



**Rechenregeln:**

(1) 
$$(x(t) + y(t))' = x'(t) + y'(t)$$

(2) 
$$(f(t) \cdot x(t))' = f'(t) \cdot x(t) + f(t) \cdot x'(t)$$

(3) **Produktregel für das Skalarprodukt:**

$$\langle x(t), y(t) \rangle' = \langle x'(t), y(t) \rangle + \langle x(t), y'(t) \rangle$$

(4) **Produktregel für das Vektorprodukt im Spezialfall  $n = 3$ :**

$$(x(t) \times y(t))' = x'(t) \times y(t) + x(t) \times y'(t)$$

(5) **Kettenregel:**

Ist  $t : J \rightarrow I$ ,  $s \mapsto t(s)$  differenzierbar, so ist

$$x \circ t : J \rightarrow \mathbb{R}^n, \quad s \mapsto x(t(s))$$

differenzierbar und es gilt an der Stelle  $s_0 \in J$  mit der Abkürzung  $t_0 = t(s_0) \in I$

$$\left. \frac{d}{ds} \right|_{s=s_0} x(t(s)) = \left. \frac{d}{dt} \right|_{t=t_0} x(t) \cdot \left. \frac{d}{ds} \right|_{s=s_0} t(s) = x'(t_0) \cdot \left. \frac{d}{ds} \right|_{s=s_0} t(s).$$

**Bemerkung 1.8** Die Produktregel für das Skalarprodukt hat eine für die Vorlesung wichtige Folgerung: Gilt  $\|x(t)\| = \text{const}$  für alle  $t \in I$ , so folgt durch Ableiten von  $\|x(t)\|^2$  dass  $x'(t)$  orthogonal ist zu  $x(t)$  für alle  $t \in I$ .

**Definition 1.9** Eine Funktion  $f$  heißt  $f$  **glatt** oder auch  **$C^\infty$ -Funktion**, falls  $f$  beliebig oft differenzierbar ist. Eine vektorwertige Funktion ist glatt beziehungsweise eine  $C^\infty$ -Funktion, falls alle Koordinatenfunktionen es sind.

## 1.6 Vektorwertige Funktionen mehrerer Veränderlicher

In diesem Abschnitt wiederholen wir einige bereits aus HM 2 bekannte Begriffe.

**Definition 1.10** Ist  $D \subseteq \mathbb{R}^n$  eine Teilmenge des  $\mathbb{R}^n$ , so wird durch die Vorschrift

$$F : D \rightarrow \mathbb{R}^m, \quad (x_1, x_2, \dots, x_n) \mapsto F(x_1, x_2, \dots, x_n) = \begin{pmatrix} F_1(x_1, x_2, \dots, x_n) \\ F_2(x_1, x_2, \dots, x_n) \\ \vdots \\ F_m(x_1, x_2, \dots, x_n) \end{pmatrix}$$

eine  **$m$ -wertige Funktion von  $n$  Variablen** definiert. Die reellwertigen Funktionen  $F_1, F_2, \dots, F_m$  heißen wieder **Koordinatenfunktionen**.

Im Spezialfall  $m = 1$  sprechen wir auch von einer **reellwertigen Funktion von  $n$  Variablen**.

Im weiteren Verlauf der Vorlesung werden vor allem 3-wertige Funktionen zweier Veränderlicher eine zentrale Rolle spielen. Ein einfaches Beispiel dazu ist die **lineare Abbildung**

$$E : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad (s, t) \mapsto F(s, t) = a + s \cdot b + t \cdot c$$

mit  $a, b, c \in \mathbb{R}^3$ ,  $b, c$  linear unabhängig; ihr Bild  $E(\mathbb{R}^2)$  ist genau die **Ebene** mit der Parameterform

$$\{a + s \cdot b + t \cdot c : s, t \in \mathbb{R}\}.$$

Weiter beschreibt für  $r > 0$  die 3-wertige Funktion zweier Variabler

$$F : [0, 2\pi] \times [0, \pi] \rightarrow \mathbb{R}^3, \\ (\varphi, \theta) \mapsto F(\varphi, \theta) = (r \cos \varphi \sin \theta, r \sin \varphi \sin \theta, r \cos \theta)$$

eine **Sphäre** vom Radius  $r$ .

Im Folgenden sei  $D \subseteq \mathbb{R}^n$  **offen**,  $f : D \rightarrow \mathbb{R}$  eine reellwertige Funktion sowie  $F : D \rightarrow \mathbb{R}^m$  eine  $m$ -wertige Funktion von  $n$  Veränderlichen. Analog zu Definition 1.4 werden im Spezialfall  $m = 1$  Grenzwert und Stetigkeit definiert:

**Definition 1.11** Ist  $\hat{x} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n) \in D$ , so heißt  $f_0 \in \mathbb{R}$  **Grenzwert** der **reellwertigen** Funktion  $f$  für  $x$  gegen  $\hat{x}$ , falls für jede Folge

$$(x^{(k)})_{k \in \mathbb{N}} = ((x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}))_{k \in \mathbb{N}} \subseteq D \quad \text{mit} \quad \lim_{k \rightarrow \infty} \|x^{(k)} - \hat{x}\| \rightarrow 0$$

$$\lim_{k \rightarrow \infty} f(x^{(k)}) = f_0$$

gilt. In diesem Fall schreiben wir  $f_0 = \lim_{x \rightarrow \hat{x}} f(x)$ .

$f$  heißt **stetig** an der Stelle  $\hat{x} \in D$ , falls  $\lim_{x \rightarrow \hat{x}} f(x) = f(\hat{x})$ .  $f$  heißt stetig, falls  $f$  für alle  $x \in D$  stetig ist.

Eine  **$m$ -wertige** Funktion ist stetig, wenn alle Koordinatenfunktionen es sind.

Fixiert man in  $f(x_1, x_2, \dots, x_n)$  alle bis auf eine Variable, so erhält man jeweils eine Funktion, die nur von einer Variablen abhängt. Diese kann man wie gewohnt ableiten:

**Definition 1.12**  $f$  heißt **partiell differenzierbar** nach  $x_i$ ,  $i = 1, 2, \dots, n$ , an der Stelle  $\hat{x} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n) \in D$ , falls der Grenzwert

$$f_{x_i}(\hat{x}) = \frac{\partial f}{\partial x_i}(\hat{x}) = \lim_{h \rightarrow 0} \frac{f(\hat{x}_1, \dots, \hat{x}_{i-1}, \hat{x}_i + h, \hat{x}_{i+1}, \dots, \hat{x}_n) - f(\hat{x})}{h}$$

existiert. In diesem Fall heißt  $f_{x_i}(\hat{x})$  die **partielle Ableitung** von  $f$  nach  $x_i$  an der Stelle  $\hat{x}$ . Der **Gradient** von  $f$  an der Stelle  $\hat{x} \in D$  ist definiert durch

$$\text{grad } f(\hat{x}) = \begin{pmatrix} f_{x_1}(\hat{x}) \\ f_{x_2}(\hat{x}) \\ \vdots \\ f_{x_n}(\hat{x}) \end{pmatrix} \in \mathbb{R}^n.$$

$f$  heißt **partiell differenzierbar**, falls  $f$  für alle  $\hat{x} \in D$  partiell differenzierbar nach allen  $n$  Variablen ist.  $f$  heißt **stetig partiell differenzierbar**, falls alle partiellen Ableitungen existieren und stetig sind.

$f$  heißt **zweimal (stetig) partiell differenzierbar**, wenn alle zweiten partiellen Ableitungen

$$f_{x_k x_i} = \frac{\partial^2 f}{\partial x_i \partial x_k}, \quad i, k = 1, 2, \dots, n,$$

von  $f$  existieren (und stetig sind). In diesem Fall ist die **Hessematrix**  $H_f(x)$  an der Stelle  $x \in D$  definiert durch

$$H_f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}, \quad x \in D.$$

Analog werden **(stetige) partielle Ableitungen höherer Ordnung** definiert.

$f$  heißt **glatt** oder  **$C^\infty$ -Funktion**, falls alle partiellen Ableitungen beliebig hoher Ordnung existieren.

**Bemerkung 1.13** Wie im Fall einer reellen Variablen heißt eine auf einer **nicht offenen** Teilmenge  $D \subseteq \mathbb{R}^n$  definierte Funktion  $f : D \rightarrow \mathbb{R}$  (stetig) partiell differenzierbar, falls  $f$  die Einschränkung einer (stetig) partiell differenzierbaren Funktion auf einer offenen Teilmenge des  $\mathbb{R}^n$  ist, die  $D$  enthält.

Sei von nun an also  $D \subseteq \mathbb{R}^n$  eine **beliebige** Menge. Wir formulieren zunächst den wichtigen

**Satz 1.14 (Satz von Schwarz)** Ist  $f : D \rightarrow \mathbb{R}$  zweimal **stetig** partiell differenzierbar, so gilt

$$f_{x_i x_k} = f_{x_k x_i} \quad \text{für alle } i, k = 1, 2, \dots, n.$$

Das heißt die Hessematrix von  $f$  ist in jedem Punkt **symmetrisch**.

Für eine  $m$ -wertige Funktion  $F : D \rightarrow \mathbb{R}^m$  hat man die naheliegende

**Definition 1.15** Die  $m$ -wertige Funktion  $F : D \rightarrow \mathbb{R}^m$  heißt **(stetig) partiell differenzierbar**, wenn alle Koordinatenfunktionen es sind. In diesem Fall ist die partielle Ableitung nach  $x_i$  (für  $i = 1, 2, \dots, n$ ) an der Stelle  $x \in D$  gegeben durch

$$F_{x_i}(x) = \frac{\partial F}{\partial x_i}(x) = \begin{pmatrix} \frac{\partial F_1}{\partial x_i}(x) \\ \frac{\partial F_2}{\partial x_i}(x) \\ \vdots \\ \frac{\partial F_m}{\partial x_i}(x) \end{pmatrix}.$$

Die **Jacobimatrix** von  $F$  an der Stelle  $x \in D$  ist definiert durch

$$J_F(x) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x) & \dots & \frac{\partial F_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(x) & \dots & \frac{\partial F_m}{\partial x_n}(x) \end{pmatrix}.$$

$F$  heißt **glatt** oder  **$C^\infty$ -Funktion**, falls alle Koordinatenfunktionen glatt sind.

**Bemerkung 1.16** Für eine partiell differenzierbare Funktion  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $n$  Variablen ist die Jacobimatrix  $J_f$  gegeben durch

$$J_f(x) = \text{grad } f(x)^\top, \quad x \in D;$$

ist  $f$  sogar zweimal partiell differenzierbar, so ist die Hessematrix genau die Jacobimatrix des Gradienten von  $f$ , das heißt

$$H_f(x) = J_{\text{grad} f}(x), \quad x \in D.$$

Für eine differenzierbare  $m$ -wertige Funktion **einer** reellen Variablen

$$F : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n, \quad t \mapsto F(t) = \begin{pmatrix} F_1(t) \\ F_2(t) \\ \vdots \\ F_m(t) \end{pmatrix},$$

ist die Jacobimatrix  $J_F$  an der Stelle  $t \in I$  gegeben durch

$$J_F(t) = F'(t) = \begin{pmatrix} F'_1(t) \\ F'_2(t) \\ \vdots \\ F'_m(t) \end{pmatrix}.$$

Analog zum bereits in Abschnitt 1.5 behandelten Fall einer reellen Variablen haben wir die **verallgemeinerte Kettenregel** für die Differentiation zusammengesetzter Funktionen:

**Satz 1.17 (Verallgemeinerte Kettenregel)** Sind  $F : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $G : U \subseteq \mathbb{R}^p \rightarrow D$  stetig partiell differenzierbar, so auch die Verkettung

$$H = F \circ G : U \rightarrow \mathbb{R}^m, \quad x \mapsto F(G(x));$$

für die Jacobimatrix gilt

$$J_H(x) = J_F(G(x)) \cdot J_G(x), \quad x \in U.$$

Beachten Sie, dass die Jacobimatrix der Verkettung  $H = F \circ G$  als Matrixprodukt einer  $(m \times n)$ -Matrix mit einer  $(n \times p)$ -Matrix eine  $(m \times p)$ -Matrix ist.

Ein für die Vorlesung äußerst wichtiger Spezialfall ist die

**Folgerung 1.18** Sind  $\varphi_1, \varphi_2 : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$  differenzierbare Funktionen,  $D \subseteq \mathbb{R}^2$ ,

$$F : D \rightarrow \mathbb{R}^3, \quad (x_1, x_2) \mapsto F(x_1, x_2),$$

eine vektorwertige Funktion von zwei Variablen mit stetigen partiellen Ableitungen, und gilt  $(\varphi_1(I), \varphi_2(I)) \subseteq D$ , so ist die vektorwertige Funktion

$$F \circ (\varphi_1, \varphi_2) : I \rightarrow \mathbb{R}^3, \quad t \mapsto F(\varphi_1(t), \varphi_2(t)),$$

differenzierbar. Die Ableitung an der Stelle  $t_0 \in I$  ist gegeben durch

$$\left. \frac{d}{dt} \right|_{t=t_0} F(\varphi_1(t), \varphi_2(t)) = F_{x_1}(\hat{x}) \cdot \varphi_1'(t_0) + F_{x_2}(\hat{x}) \cdot \varphi_2'(t_0),$$

wobei  $\hat{x} = (\varphi_1(t_0), \varphi_2(t_0))$ .

**BEWEIS.** Für die 2-wertige Funktion einer reellen Variablen

$$G : I \rightarrow \mathbb{R}^2, \quad t \mapsto \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \end{pmatrix},$$

berechnen wir

$$J_G(t) = G'(t) = \begin{pmatrix} \varphi_1'(t) \\ \varphi_2'(t) \end{pmatrix}, \quad t \in I.$$

Weiter ist die Jacobimatrix von  $F$  an der Stelle  $x = (x_1, x_2) \in D$  gegeben durch

$$J_F(x) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x) & \frac{\partial F_1}{\partial x_2}(x) \\ \frac{\partial F_2}{\partial x_1}(x) & \frac{\partial F_2}{\partial x_2}(x) \\ \frac{\partial F_3}{\partial x_1}(x) & \frac{\partial F_3}{\partial x_2}(x) \end{pmatrix};$$

somit gilt für die Jacobimatrix der Verkettung  $H = F \circ G : I \rightarrow \mathbb{R}^3$  an der Stelle  $t_0 \in I$

$$\begin{aligned} J_H(t_0) &= J_F(G(t_0)) \cdot J_G(t_0) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(G(t_0)) & \frac{\partial F_1}{\partial x_2}(G(t_0)) \\ \frac{\partial F_2}{\partial x_1}(G(t_0)) & \frac{\partial F_2}{\partial x_2}(G(t_0)) \\ \frac{\partial F_3}{\partial x_1}(G(t_0)) & \frac{\partial F_3}{\partial x_2}(G(t_0)) \end{pmatrix} \cdot \begin{pmatrix} \varphi_1'(t_0) \\ \varphi_2'(t_0) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(G(t_0)) \cdot \varphi_1'(t_0) + \frac{\partial F_1}{\partial x_2}(G(t_0)) \cdot \varphi_2'(t_0) \\ \frac{\partial F_2}{\partial x_1}(G(t_0)) \cdot \varphi_1'(t_0) + \frac{\partial F_2}{\partial x_2}(G(t_0)) \cdot \varphi_2'(t_0) \\ \frac{\partial F_3}{\partial x_1}(G(t_0)) \cdot \varphi_1'(t_0) + \frac{\partial F_3}{\partial x_2}(G(t_0)) \cdot \varphi_2'(t_0) \end{pmatrix}. \end{aligned}$$

Mit den Abkürzungen  $\hat{x} = G(t_0) = (\varphi_1(t_0), \varphi_2(t_0)) \in D$  sowie

$$F_{x_i}(\hat{x}) = \frac{\partial F}{\partial x_i}(\hat{x}) = \begin{pmatrix} \frac{\partial F_1}{\partial x_i}(\hat{x}) \\ \frac{\partial F_2}{\partial x_i}(\hat{x}) \\ \frac{\partial F_3}{\partial x_i}(\hat{x}) \end{pmatrix}, \quad i = 1, 2, 3,$$

folgt daraus die Behauptung. ■

## 1.7 Differentialgleichungen

Unter einer Differentialgleichung versteht man eine Bestimmungsgleichung zum Auffinden einer Funktion, in der neben der Funktion und ihrer Veränderlichen auch die Ableitung auftritt, gegebenenfalls auch Ableitungen höherer Ordnung. Genauer haben wir die folgende

**Definition 1.19** Eine **Differentialgleichung  $n$ -ter Ordnung** für eine  $n$ -mal differenzierbare Funktion  $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$  ist eine Gleichung der Form

$$\begin{aligned} G(t, f, f', \dots, f^{(n-1)}, f^{(n)}) &= 0 \quad (\text{implizite Form}) \\ \text{bzw. } f^{(n)} &= H(t, f, f', \dots, f^{(n-1)}) \quad (\text{explizite Form}), \end{aligned} \tag{1.8}$$

wobei  $G$  und  $H$  Funktionen sind, die auf einer Teilmenge des  $\mathbb{R}^{n+2}$  bzw. des  $\mathbb{R}^{n+1}$  definiert sind.

Unter dem **Anfangswertproblem**

$$f^{(n)} = H(t, f, f', \dots, f^{(n-1)}) \quad \text{mit} \quad (t_0, f_0, f_1, \dots, f_{n-1}) \quad (1.9)$$

versteht man die Aufgabe, eine Lösung  $\tilde{f}$  der Differentialgleichung (1.8) zu finden, die die Anfangsbedingungen

$$\tilde{f}(t_0) = f_0, \quad \tilde{f}'(t_0) = f_1, \dots, \quad \tilde{f}^{(n-1)}(t_0) = f_{n-1},$$

erfüllt.

Eine Funktion  $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$  heißt **allgemeine Lösung** der Differentialgleichung (1.8)  $n$ -ter Ordnung, wenn ihre Zuordnungsvorschrift neben der Variablen  $t$  noch  $n$  frei wählbare Konstanten  $c_0, c_1, \dots, c_{n-1} \in \mathbb{R}$  enthält, und jede Lösung der Differentialgleichung durch Einsetzen geeigneter Zahlenwerte für diese frei wählbaren Konstanten aus  $f$  entsteht.

Wir wollen zunächst eine wichtige Lösungsmethode für gewisse Differentialgleichungen erster Ordnung vorstellen, die sogenannte **Separation der Variablen**: Hat man eine Differentialgleichung der speziellen Form

$$f'(t) = g(f) \cdot h(t)$$

mit stetigen Funktionen  $g, h$ ,  $g \neq 0$ , so kann man die Veränderlichen  $f$  und  $t$  trennen, d.h. auf verschiedene Seiten des Gleichheitszeichens bekommen.

Zur Lösung schreibt man daher zunächst formal  $f' = \frac{df}{dt}$ , dividiert durch  $g$  und multipliziert mit  $dt$ , also

$$\frac{df}{g(f)} = h(t)dt.$$

Dabei muss man natürlich die Nullstellen von  $g$  ausschließen.

Die Lösung der Differentialgleichung erhält man nun, indem man für jede Seite eine Stammfunktion bestimmt, d.h. durch

$$\int \frac{df}{g(f)} = \int h(t)dt.$$

Wir betrachten dazu das folgende

**Beispiel 1.20** Gesucht ist die allgemeine Lösung der Differentialgleichung

$$f' = 3t^2 f.$$

Zunächst bemerken wir, dass die konstante Funktion  $f = 0$  eine Lösung ist. Für  $f \neq 0$  schreiben wir

$$\frac{df}{f} = 3t^2 dt,$$

erhalten also durch Integration

$$\ln|f| = t^3 + C \quad \text{mit } C \in \mathbb{R}.$$

Die allgemeine Lösung ist somit gegeben durch

$$f(t) = c_0 \cdot e^{t^3} \quad \text{mit } c_0 \in \mathbb{R}.$$

Wir betrachten nun allgemeiner Systeme von Differentialgleichungen, bei denen mehrere Funktionen derselben unabhängigen Variablen gesucht sind; man kann solche Systeme auch als Menge von Differentialgleichungen für eine vektorwertige Funktion

$y : I \subset \mathbb{R} \rightarrow \mathbb{R}^k$  auffassen.

**Definition 1.21** Ein **Differentialgleichungssystem** (oder genauer ein System von  $m$  Differentialgleichungen) für eine vektorwertige Funktion

$$y : I \subset \mathbb{R} \rightarrow \mathbb{R}^k, \quad t \mapsto \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_k(t) \end{pmatrix},$$

ist eine Gleichung der Form

$$G(t, y_1, \dots, y_k, y_1', \dots, y_k', \dots, y_1^{(n)}, \dots, y_k^{(n)}) = 0 \quad (1.10)$$

mit  $G : U \subset \mathbb{R}^{(n+1) \cdot k+1} \rightarrow \mathbb{R}^m$ . Die höchste auftretende Ableitung heißt **Ordnung** des Differentialgleichungssystems.

Bei einem Differentialgleichungssystem handelt es sich also um  $m$  Differentialgleichungen  $n$ -ter Ordnung für die  $k$  Funktionen  $y_1, y_2, \dots, y_k : I \subset \mathbb{R} \rightarrow \mathbb{R}$ . Man sagt auch, die Funktionen  $y_1, y_2, \dots, y_k$  sind über das Differentialgleichungssystem (1.10) **gekoppelt**.

Wir wollen nun ein Beispiel für ein solches Differentialgleichungssystem betrachten:

### Beispiel 1.22

$$\begin{aligned} 0 &= u'' + 2 \tanh v \cdot u' \cdot v' \\ 0 &= \cosh v \cdot v' - 1 \end{aligned}$$

In diesem Beispiel handelt es sich um ein Differentialgleichungssystem zweiter Ordnung bestehend aus zwei Gleichungen (also  $m = 2$ ) für zwei unbekannte Funktionen (d.h.  $k = 2$ ).



Wir beschränken uns im Folgenden auf Differentialgleichungssysteme **erster Ordnung**. Der einfachste Spezialfall eines solchen Systems – nämlich  $m = k = 1$  – entspricht genau einer Differentialgleichung erster Ordnung für eine Funktion einer Variablen in impliziter Form. Bei folgendem Beispiel handelt es sich um eine Differentialgleichung erster Ordnung für zwei Funktionen derselben Variablen:

**Beispiel 1.23**

$$(u')^2 + (1 - u) \cdot (v')^2 = 0$$

Im Allgemeinen besitzen Differentialgleichungen und Differentialgleichungssysteme nicht immer eine Lösung. Andererseits können – wie man am vorigen Beispiel sieht – trotz vorgegebener Anfangsbedingungen verschiedene Lösungen existieren. Wir wenden uns daher der Frage zu, wann ein gegebenes Anfangswertproblem eine eindeutige Lösung besitzt. Dazu betrachten wir den Spezialfall  $k = m$ , also Systeme von  $m$  Differentialgleichungen erster Ordnung für eine vektorwertige Funktion  $y : I \subset \mathbb{R} \rightarrow \mathbb{R}^m$  mit  $m$  Komponentenfunktionen  $y_1, y_2, \dots, y_m$  in expliziter Form:

$$y'(t) = H(t, y(t)), \quad \text{wobei } H : U \subset \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m,$$

oder auch komponentenweise

$$\begin{aligned} y_1'(t) &= H_1(t, y_1(t), \dots, y_m(t)), \\ y_2'(t) &= H_2(t, y_1(t), \dots, y_m(t)), \\ &\vdots \\ y_m'(t) &= H_m(t, y_1(t), \dots, y_m(t)). \end{aligned} \tag{1.11}$$

Der folgende Satz besagt, dass das Anfangswertproblem

$$y'(t) = H(t, y(t)) \quad \text{mit } (t_0, y_0)$$

unter gewissen Voraussetzungen zumindest lokal eine eindeutige Lösung besitzt.

**Satz 1.24 (Satz von Picard-Lindelöf)** *Gegeben sei  $t_0 \in \mathbb{R}$ ,  $\varepsilon > 0$ ,  $y_0 \in \mathbb{R}^m$  und  $D \subset \mathbb{R}^m$  eine Kugel um  $y_0$ . Sei weiter  $H : [t_0 - \varepsilon, t_0 + \varepsilon] \times D \rightarrow \mathbb{R}^m$ , und es existiere eine Konstante  $L > 0$  mit der Eigenschaft*

$$\|H(t, x) - H(t, z)\| \leq L \cdot \|x - z\|$$

für alle  $x, z \in D$  und  $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ . Dann besitzt das Anfangswertproblem

$$y'(t) = H(t, y(t)), \quad y(t_0) = y_0,$$

genau eine stetig differenzierbare Lösung  $y : J \rightarrow \mathbb{R}^m$  auf einem geeigneten Teilintervall  $J \subseteq [t_0 - \varepsilon, t_0 + \varepsilon]$ .

Die Aussage dieses Satzes ist insbesondere für die folgende wichtige Klasse von Differentialgleichungssystemen richtig:

**Definition 1.25** Ein mit  $m(m+1)$  stetigen Funktionen

$$a_{11}, a_{12}, \dots, a_{1m}, a_{21}, \dots, a_{2m}, \dots, a_{mm}, b_1, \dots, b_m : I \subset \mathbb{R} \rightarrow \mathbb{R}$$

gebildetes Differentialgleichungssystem der Form

$$y'(t) = \underbrace{\begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1m}(t) \\ a_{21}(t) & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{m1}(t) & \cdots & \cdots & a_{mm}(t) \end{pmatrix}}_{=:A(t)} \cdot y(t) + \underbrace{\begin{pmatrix} b_1(t) \\ b_2(t) \\ \vdots \\ b_m(t) \end{pmatrix}}_{=:b(t)} \quad (1.12)$$

heißt **lineares Differentialgleichungssystem** erster Ordnung. Die darin auftretende Matrix  $A$  mit den Einträgen  $a_{ik}$ ,  $1 \leq i, k \leq m$ , heißt **Koeffizientenmatrix**, die vektorwertige Funktion  $b$  mit den Komponenten  $b_i$ ,  $1 \leq i \leq m$ , heißt **Störvektor**.

Ist  $b_1(t) = b_2(t) = \dots = b_m(t) = 0$  für alle  $t \in I$ , so heißt das Differentialgleichungssystem **homogen**, andernfalls **inhomogen**. Sind alle Einträge der Koeffizientenmatrix konstante Funktionen, so nennt man (1.12) ein lineares Differentialgleichungssystem mit konstanten Koeffizienten.

Da stetige Funktionen auf einem kompakten Intervall beschränkt sind, besitzt nach dem Satz von Picard-Lindelöf das Anfangswertproblem

$$y'(t) = A(t) \cdot y(t) + b(t), \quad y(t_0) = y_0,$$

mit  $t_0 \in I$  und  $y_0 \in \mathbb{R}^m$  stets genau eine stetig differenzierbare Lösung auf einem geeigneten Teilintervall von  $I$ .

**Beispiel 1.26** Bei dem System

$$y'(t) = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix} \cdot y(t)$$

handelt es sich um ein homogenes lineares Differentialgleichungssystem mit konstanten Koeffizienten; in diesem Fall ist  $m = 3$ .

Für lineare Differentialgleichungssysteme ist die Struktur der Lösungsmenge genau bekannt; wir wollen zunächst **homogene** Systeme betrachten:

**Satz 1.27** Sind  $u_1, u_2, \dots, u_m : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$  linear unabhängige Lösungen des homogenen linearen Differentialgleichungssystems  $y'(t) = A(t) \cdot y(t)$  wie in (1.12), so ist die allgemeine Lösung gegeben durch

$$y_h(t) = c_1 \cdot u_1(t) + c_2 \cdot u_2(t) + \dots + c_m \cdot u_m(t),$$

wobei  $c_1, c_2, \dots, c_m \in \mathbb{R}$  frei wählbare Konstanten sind.

Für **inhomogene** Systeme gilt der folgende

**Satz 1.28** Die allgemeine Lösung eines inhomogenen linearen Differentialgleichungssystems (1.12) ergibt sich als Summe der allgemeinen Lösung  $y_h$  des zugehörigen homogenen linearen Differentialgleichungssystems  $y'(t) = A(t) \cdot y(t)$  plus einer beliebigen Lösung  $y_p$  des inhomogenen Systems  $y'(t) = A(t) \cdot y(t) + b(t)$ .

Wie man lineare Gleichungssysteme lösen kann, haben Sie in Abschnitt 2.3 von HM 2 gelernt. Da bei uns im Wesentlichen nichtlineare Systeme auftreten werden, gehen wir auf diese Lösungsstrategien nicht nochmals ein.