

Mathematik für Chemiker 1

Prof. Dr. habil. Thomas Runst,
Mathematisches Institut, Friedrich–Schiller–Universität Jena

Ziel der Vorlesung:

- teilweise Wiederholung und Vertiefung der Kenntnisse aus dem Mathematikunterricht (aufbauend auf Vorkurs)
- Anfangsgründe der höheren Mathematik
- Vermittlung von mathematischen Lehrinhalten, die für die Chemieausbildung erforderlich sind
- Schwerpunkte: Verständnis von mathematischen Denkweisen und Modellen, Herausbildung von rechnerischen Fertigkeiten und deren Anwendungen in der Chemie

1 Komplexe Zahlen

Die folgenden Überlegungen führen zu einer notwendigen Erweiterung der Menge der reellen Zahlen. Wir betrachten die Lösbarkeit von *reellen* quadratischen Gleichungen der Form

$$ax^2 + bx + c = 0,$$

wobei $a, b, c \in \mathbb{R}$ und $a \neq 0$ gilt (ansonsten hätten wir eine lineare Gleichung vorliegen). Durch äquivalente Umformung erhält man

$$x^2 + \frac{b}{a}x + \frac{c}{a} = 0.$$

Setzen wir $b/a = p$, $c/a = q$, so erhalten wir die *Normalform* einer quadratischen Gleichung, d.h.

$$x^2 + px + q = 0.$$

Aus der Schule ist bekannt, dass das Lösungsverhalten von der Diskriminante

$$D = \frac{p^2}{4} - q$$

abhängt. Es gilt

- $D > 0$: zwei reelle Lösungen $x_{1,2} = -\frac{p}{2} \pm \sqrt{D}$,
- $D = 0$: eine reelle (Doppel-)Lösung $x = -\frac{p}{2}$,
- $D < 0$: keine reelle Lösung, d.h., in \mathbb{R} nicht lösbar.

Unser Ziel besteht jetzt darin, die reellen Zahlen \mathbb{R} so zu erweitern, dass jede beliebige quadratische Gleichung mit reellen Koeffizienten lösbar ist. Dieser Ansatz, der uns zunächst als innermathematisch erscheint, führt uns später zur Menge \mathbb{C} der *komplexen* Zahlen. Jedoch existieren vielfältige Anwendungen in der Mathematik und in den Naturwissenschaften, z.B. bei der Beschreibung von Schwingungsvorgängen in der Mechanik und der Elektrodynamik.

Andererseits hat sich in der Mathematik die Funktionentheorie, die sich mit der Untersuchung von Funktionen einer *komplexen* Variablen beschäftigt, als eigenständige Theorie herausgebildet. Diese wird sowohl in der nichteuklidischen Geometrie, in der Strömungstheorie (Hydrodynamik), Integrationstheorie und der Theorie der Differentialgleichungen angewendet.

Die einfachste quadratische Gleichung, die in \mathbb{R} nicht lösbar ist, lautet $x^2 + 1 = 0$.

Definition 1 Die Zahl i heißt *imaginäre Einheit* und ist Lösung der quadratischen Gleichung $x^2 + 1 = 0$, d.h. es gilt $i^2 = -1$. Wir setzen $i = \sqrt{-1}$.

Das formale Vorgehen ist also

$$i^2 = -1 \rightarrow i = \sqrt{-1}.$$

Bemerkung 1 Es ist leicht nachzuweisen, dass $-i = -\sqrt{-1}$ die zweite Lösung ist.

$$(-i)(-i) = (-1)^2 i^2 = -1.$$

Allgemein erhalten wir unter Verwendung von Definition 1 als Lösungen für die quadratische Gleichung $x^2 + px + q = 0$ mit reellen Koeffizienten p und q im Falle

$$D = \frac{p^2}{4} - q < 0 \iff -D > 0$$
$$x_{1/2} = -\frac{p}{2} \pm \sqrt{D} = -\frac{p}{2} \pm \sqrt{(-1)(-D)} = -\frac{p}{2} \pm i\sqrt{-D},$$

wobei $\sqrt{-D} \in \mathbb{R}$ gilt. Durch Einsetzen ergibt sich leicht, dass x_1 und x_2 wirklich Lösungen der quadratischen Gleichung $x^2 + px + q = 0$ sind. Somit finden wir im Falle einer negativen Diskriminante D ebenfalls zwei (nichtreelle) Lösungen

$$x_1 = a + bi, \quad x_2 = a - bi,$$

wobei

$$a = -\frac{p}{2} \in \mathbb{R}, \quad b = \sqrt{-D} = \sqrt{q - \frac{p^2}{4}} \in \mathbb{R}$$

gilt. Diese Überlegungen führen uns jetzt zur Menge der komplexen Zahlen.

Definition 2 Eine komplexe Zahl z ist ein geordnetes Paar reeller Zahlen, d.h.

$$z = (a, b), \quad a, b \in \mathbb{R}.$$

Durch $\mathbb{C} = \{(a, b) : a, b \in \mathbb{R}\}$ definieren wir die Menge der komplexen Zahlen. Falls $z = (a, b) \in \mathbb{C}$ ist, so bezeichnet $a = \operatorname{Re} z$ den Realteil von z und $b = \operatorname{Im} z$ den Imaginärteil von z . $0 = (0, 0)$ ist das Nullelement.

Für $z = (a, b) \in \mathbb{C}$ wird mit $z = a + bi = a + ib$ die Normaldarstellung von z eingeführt.

Bemerkung 2 (a) Es gilt $i = (0, 1)$ sowie $0 = (0, 0) = 0 + 0i$ in diesem Sinne.

(b) Eine reelle Zahl a wird in \mathbb{C} mit dem Paar $(a, 0) = a + 0i$ identifiziert, d.h. es gilt $\mathbb{R} \subset \mathbb{C}$.

(c) Für geordnete Paare gilt:

$$z_1 = (a_1, b_1) = (a_2, b_2) = z_2,$$

genau dann, wenn

$$\operatorname{Re} z_1 = a_1 = a_2 = \operatorname{Re} z_2, \quad \operatorname{Im} z_1 = b_1 = b_2 = \operatorname{Im} z_2$$

gilt, d.h. wir haben insbesondere $(a, b) \neq (b, a)$, falls $a \neq b$.

Für komplexe Zahlen wird folgendermaßen eine Addition, Subtraktion und Multiplikation eingeführt.

Falls $z_1 = a_1 + b_1i = (a_1, b_1)$ und $z_2 = a_2 + b_2i = (a_2, b_2)$, so gelte

$$z_1 \pm z_2 = (a_1 \pm a_2, b_1 \pm b_2) = a_1 \pm a_2 + i(b_1 \pm b_2)$$

sowie

$$z_1 \cdot z_2 = (a_1a_2 - b_1b_2, a_1b_2 + a_2b_1) = (a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1).$$

Bemerkung 3 (a) Aus diesen Definitionen folgen insbesondere die bekannten Rechenregeln für die reellen Zahlen, da

$$(a_1, 0) \pm (a_2, 0) = (a_1 \pm a_2, 0) = a_1 \pm a_2 + 0i$$

und

$$(a_1, 0) \cdot (a_2, 0) = (a_1a_2, 0) = a_1a_2 + 0i.$$

(b) Bei der Definition der Multiplikation wird die Eigenschaft $i^2 = (0, 1) \cdot (0, 1) = -1$ sowie die binomische Formel verwendet, da

$$(a_1 + b_1i) \cdot (a_2 + b_2i) = a_1a_2 + i^2b_1b_2 + i(a_1b_2 + a_2b_1) = (a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1).$$

(c) Falls $z = (a, b) = a + bi$ und λ eine reelle Zahl ist, so folgt mit $\lambda = (\lambda, 0)$, dass für die skalare Multiplikation einer komplexen Zahl

$$\lambda z = (\lambda, 0)(a, b) = (\lambda a, \lambda b) = \lambda a + \lambda bi$$

gilt.

$$(d) (1 + 2i)(2 - 3i) = (1 \cdot 2 - 2 \cdot (-3)) + i(2 \cdot 2 + 1 \cdot (-3)) = (2 + 6) + i(4 - 3) = 8 + i$$

Weiterhin gelten die folgende Eigenschaften für die oben definierten Rechenoperationen:

- Assoziativgesetze:

$$z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3 = z_1 + z_2 + z_3,$$

$$z_1 \cdot (z_2 \cdot z_3) = (z_1 \cdot z_2) \cdot z_3 = z_1 \cdot z_2 \cdot z_3$$

- Kommutativgesetze:

$$z_1 + z_2 = z_2 + z_1, \quad z_1 \cdot z_2 = z_2 \cdot z_1$$

- Distributivgesetz:

$$(z_1 + z_2) \cdot z_3 = z_1 \cdot z_3 + z_2 \cdot z_3$$

Definition 3 Es sei $z = (a, b) = a + bi$ eine komplexe Zahl. Dann heißt $\bar{z} = (a, -b) = a - bi$ die zu z konjugiert komplexe Zahl.

Bemerkung 4 (a) Es sei $x^2 + px + q = 0$ mit negativer Diskriminante $D < 0$ gegeben. Dann hat die quadratische Gleichung zwei konjugiert komplexe Lösungen z_1, z_2 mit $z_1 = \bar{z}_2$.

(b) Man kann sich leicht davon überzeugen, dass

$$z + \bar{z} = 2\operatorname{Re} z, \quad z - \bar{z} = 2i\operatorname{Im} z$$

und für $z = a + bi \neq 0$ die Zahl $z \cdot \bar{z}$ reell und positiv ist, da wir

$$z \cdot \bar{z} = (a + bi) \cdot (a - bi) = a^2 + b^2 > 0$$

erhalten.

In Analogie zu \mathbb{R} können wir die komplexen Zahlen \mathbb{C} mit der sogenannten Gaußschen Zahlenebene identifizieren.

Dabei bedeutet graphisch $z \rightarrow \bar{z}$ eine Spiegelung an der Abszissenachse $\operatorname{Re} z$, $z \rightarrow -z$ eine Spiegelung am Koordinatenursprung $0 = (0, 0)$ (Nullpunkt), und die Addition von zwei komplexen Zahlen entspricht der Addition der entsprechenden Ortsvektoren in der Ebene.

Weiterhin kann man unter Verwendung von Bemerkung 3(c) und Bemerkung 4 (b) eine Division von komplexen Zahlen einführen.

Falls $z_1 = (a_1, b_1)$ und $z_2 = (a_2, b_2) \neq 0$, so heißt

$$w = \frac{z_1}{z_2}$$

Lösung der Gleichung $z_2 \cdot w = z_1$.

Es gilt

$$\bar{z}_2 \cdot z_2 \cdot w = \bar{z}_2 \cdot z_1.$$

Da die reelle Zahl $\bar{z}_2 \cdot z_2$ positiv ist, folgt schließlich

$$w = \frac{1}{\bar{z}_2 \cdot z_2} \bar{z}_2 \cdot z_1.$$

Bemerkung 5 Man erweitert also den Quotienten mit $\overline{z_2} = a_2 - b_2i$, falls der Nenner durch $z_2 = a_2 + b_2i \neq 0$ gegeben ist. Eine Division durch $z_2 = 0$ ist wie bei den reellen Zahlen nicht erklärt.

Beispiel 1

$$\frac{1+2i}{2+3i} = \frac{(1+2i)(2-3i)}{(2+3i)(2-3i)} = \frac{8+i}{13} = \frac{8}{13} + \frac{1}{13}i$$

Definition 4 Mit $|z|$ bezeichnet man den Abstand von $z = (a, b) \in \mathbb{C}$ vom Nullpunkt in der Gaußschen Zahlenebene, d.h.

$$|z| = \sqrt{a^2 + b^2} = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}.$$

Bemerkung 6 Es seien $z_k = a_k + b_ki$, $k = 1, 2$, gegeben. Dann berechnet sich der Abstand $|z_1 - z_2|$ nach dem Satz von Pythagoras

$$|z_1 - z_2| = \sqrt{(\operatorname{Re}(z_1 - z_2))^2 + (\operatorname{Im}(z_1 - z_2))^2} = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2}.$$

Lemma Es gilt $z \cdot \bar{z} = |z|^2$ und $|-z| = |z|$.

Beweis: Es sei $z = a + bi$ gegeben.

$$z \cdot \bar{z} = (a + bi)(a - bi) = a^2 + b^2 = |z|^2.$$

Weiterhin gilt

$$|-z| = \sqrt{(-a)^2 + (-b)^2} = \sqrt{a^2 + b^2} = |z|.$$

■

Abschließend führen wir mittels Polarkoordinaten eine weitere Darstellung für die komplexen Zahlen ein. Ausgangspunkt ist hierbei die geometrische Darstellung von $z = x + yi \in \mathbb{C}$ in der Gaußschen Zahlenebene.

Dabei bezeichnet $r = |z|$ den Abstand vom Nullpunkt, und für $z \neq 0$ ist durch $\arg z = \varphi$ das Argument von z gegeben. Hierbei ist φ der Winkel zwischen dem Ortsvektor von z und der positiven Abszissenachse $\operatorname{Re} z > 0$.

Definition 5 Für $z \in \mathbb{C} \setminus \{0\}$ bezeichnet (r, φ) die zugehörigen Polarkordinaten.

Bemerkung 7 (a) Es gilt für $z = (x, y) = x + yi$

$$x = r \cos \varphi, \quad y = r \sin \varphi.$$

Damit erhalten wir folgende *trigonometrische* Darstellung für z

$$z = r(\cos \varphi + i \sin \varphi).$$

(b) Für $x \neq 0$ gilt

$$\frac{y}{x} = \frac{r \sin \varphi}{r \cos \varphi} = \tan \varphi,$$

d.h.

$$\varphi = \arctan \frac{y}{x} = \arctan \frac{\operatorname{Im} z}{\operatorname{Re} z} \quad (\text{Quadrantenbeziehung}).$$

Falls $x = 0$, so gilt $\cos \varphi = 0$. Dann verwendet man folgende Vereinbarung für $z = (0, y)$, $y \neq 0$:

$$x = 0, y > 0 \rightarrow \varphi = \frac{\pi}{2}, \quad x = 0, y < 0 \rightarrow \varphi = \frac{3\pi}{2}.$$

(c) Für $z \neq 0$ ist der Winkel bis auf Vielfaches von 2π eindeutig bestimmt. Deshalb verwendet man die Einschränkung $0 \leq \varphi < 2\pi$ für den sogenannten *Hauptwert* von z . In der Literatur findet man manchmal auch $-\pi \leq \varphi < \pi$.

(d) Für $z = 0$ gilt $r = 0$, aber φ ist unbestimmt.

(e) Offensichtlich sind zwei komplexe Zahlen z_1 und z_2 gleich, falls

$$r_1 = |z_1| = |z_2| = r_2, \quad \varphi_1 - \varphi_2 = \arg z_1 - \arg z_2 = 2k\pi, \quad k \in \mathbb{Z}.$$

Wir wollen jetzt diese trigonometrische Darstellung verwenden. Sie erweist sich bei der Multiplikation, Division, beim Potenzieren und Radizieren oft als vorteilhaft (im Gegensatz zur Addition und Subtraktion).

Es seien $z_k = r_k(\cos \varphi_k + i \sin \varphi_k)$, $k = 1, 2$, zwei komplexe Zahlen. Dann erhalten wir unter Verwendung der Additionstheoreme für \sin und \cos

$$\begin{aligned} z_1 \cdot z_2 &= r_1(\cos \varphi_1 + i \sin \varphi_1)r_2(\cos \varphi_2 + i \sin \varphi_2) \\ &= r_1r_2[(\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2)] + i(\cos \varphi_1 \sin \varphi_2 + \sin \varphi_1 \cos \varphi_2) \\ &= r_1r_2[\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)], \end{aligned}$$

d.h. zwei komplexe Zahlen werden multipliziert, indem ihre Beträge multipliziert und ihre Argumente addiert werden.

Analog erhält man für $z_2 \neq 0$

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}[\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)].$$

Beispiel 2 Für

$$z_1 = \sqrt{3} + i = 2\left(\cos \frac{\pi}{6} + i \sin \frac{\pi}{6}\right)$$

und

$$z_2 = 2 + 2i = 2\sqrt{2}\left(\cos \frac{\pi}{4} + i \sin \frac{\pi}{4}\right)$$

erhalten wir

$$z_1 \cdot z_2 = 4\sqrt{2}\left(\cos \frac{5\pi}{12} + i \sin \frac{5\pi}{12}\right)$$

und

$$\frac{z_1}{z_2} = \frac{\sqrt{2}}{2}\left[\cos\left(\frac{-\pi}{12}\right) + i \sin\left(\frac{-\pi}{12}\right)\right] = \frac{\sqrt{2}}{2}\left(\cos \frac{23\pi}{12} + i \sin \frac{23\pi}{12}\right),$$

wobei wir Bemerkung 7(e) verwendeten.

Als nächstes beschäftigen wir uns mit Potenzen von komplexen Zahlen.

Definition 6 Es sei $z \in \mathbb{C}$. Dann definieren wir

$$z^k = \begin{cases} \underbrace{z \cdot \dots \cdot z}_k : & k = 1, 2, 3, \dots \\ 1 : & k = 0, z \neq 0 \\ \underbrace{\frac{1}{z} \cdot \dots \cdot \frac{1}{z}}_{-k} : & k = -1, -2, -3, \dots, z \neq 0. \end{cases}$$

Folgerung 1 Es sei $k \in \mathbb{Z}$ und $z = r(\cos \varphi + i \sin \varphi)$ mit $r \neq 0$. Dann gilt einerseits

$$z^k = r^k(\cos k\varphi + i \sin k\varphi)$$

und andererseits

$$z^k = r^k(\cos \varphi + i \sin \varphi)^k.$$

Daraus folgt wegen $r \neq 0$ die *Formel von Moivre*

$$\cos k\varphi + i \sin k\varphi = (\cos \varphi + i \sin \varphi)^k.$$

Bemerkung 8 (a) In der Literatur verwendet man oft auch folgende Schreibweise (Eulersche Formel)

$$e^{i\varphi} = \cos \varphi + i \sin \varphi.$$

Damit gilt insbesondere

$$e^{i(\varphi+2k\pi)} = e^{i\varphi}, \quad k \in \mathbb{Z}.$$

(b) Es seien $k, l \in \mathbb{Z}$ und $z, z_1, z_2 \in \mathbb{C} \setminus \{0\}$. Dann gelten folgende Rechenregeln für das Potenzieren:

$$z^k \cdot z^l = z^{k+l}, \quad z_1^k \cdot z_2^k = (z_1 \cdot z_2)^k, \quad (z^k)^l = z^{k \cdot l}.$$

Wir beschäftigen uns abschließend mit dem Radizieren von komplexen Zahlen. Dazu seien die komplexe Zahl $z = r(\cos \varphi + i \sin \varphi)$ sowie $n \in \mathbb{N}$ gegeben. Gesucht ist eine komplexe Zahl $z_0 = r_0(\cos \varphi_0 + i \sin \varphi_0)$, so dass $z_0^n = z$, d.h. $z_0 = \sqrt[n]{z}$ gilt. Daraus folgt aber

$$z_0^n = r_0^n(\cos n\varphi_0 + i \sin n\varphi_0) = r(\cos \varphi + i \sin \varphi)$$

und somit

$$z_0^n = z \iff r_0^n = r, \quad n\varphi_0 = \varphi + 2k\pi, \quad k \in \mathbb{Z}.$$

Wegen $r \geq 0$ erhalten wir schließlich

$$r_0 = \sqrt[n]{r}, \quad \varphi_0 = \frac{\varphi + 2k\pi}{n}, \quad k \in \mathbb{Z}.$$

Damit sind alle Werte durch

$$z_0 = \sqrt[n]{r} \left(\cos \frac{\varphi + 2k\pi}{n} + i \sin \frac{\varphi + 2k\pi}{n} \right), \quad k = 0, 1, \dots, n-1,$$

gegeben.

Satz 1 Sei $n \in \mathbb{N}$. Die n -ten Wurzeln aus einer komplexen Zahl $z = r(\cos \varphi + i \sin \varphi) \neq 0$ sind durch die n verschiedenen komplexen Zahlen

$$z_k = \sqrt[n]{r} \left(\cos \frac{\varphi + 2k\pi}{n} + i \sin \frac{\varphi + 2k\pi}{n} \right), \quad k = 0, 1, \dots, n-1,$$

gegeben.

Beispiel 3 Es sei $z = \sqrt[3]{i}$. Dann gilt wegen $i = 1(\cos \frac{\pi}{2} + i \sin \frac{\pi}{2})$

$$\left(\sqrt[3]{i} \right)_k = 1 \cdot \left(\cos \frac{\pi/2 + 2k\pi}{3} + i \sin \frac{\pi/2 + 2k\pi}{3} \right), \quad k = 0, 1, 2.$$

Bemerkung 9 Es seien $z, z_1, z_2 \in \mathbb{C} \setminus \{0\}$ und $m, n \in \mathbb{N}$. Dann gelten folgende Rechenregeln für das Radizieren:

$$\sqrt[n]{z_1} \cdot \sqrt[n]{z_2} = \sqrt[n]{z_1 \cdot z_2}, \quad \frac{\sqrt[n]{z_1}}{\sqrt[n]{z_2}} = \sqrt[n]{\frac{z_1}{z_2}}, \quad \sqrt[m]{\sqrt[n]{z}} = \sqrt[m \cdot n]{z}.$$

Beispiel 4 Sei $n \in \mathbb{N}$. Man bestimme alle komplexen Lösungen (*Einheitswurzeln*) von

$$z^n = 1.$$

Die Zahl 1 hat die Polarkoordinaten $r = 1$ und $\varphi = 0$. Somit erhalten wir

$$\left(\sqrt[n]{1} \right)_k = 1 \left(\cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n} \right) = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}, \quad k = 0, 1, \dots, n-1.$$

Somit bilden die Einheitswurzeln ein regelmäßiges n -Eck auf dem Rand des Einheitskreises, beginnend mit 1 ($k = 0$). Nach Bemerkung 8(a) gilt

$$\left(\sqrt[n]{1} \right)_k = e^{i \frac{2k\pi}{n}}.$$

Für $n = 4$ erhalten wir für $z^4 = 1$ die Lösungsmenge $\{1, -1, i, -i\}$, für $n = 3$ gilt

$$\left(\sqrt[3]{1} \right)_k = \cos \frac{2k\pi}{3} + i \sin \frac{2k\pi}{3}$$

2 Integralrechnung im \mathbb{R}^1

2.1 Das bestimmte Riemannsche Integral

Wir betrachten als Motivation folgendes Problem. Gegeben sei eine Fläche F mit einer beliebigen Berandung. Wie kann man dann den Flächeninhalt $|F|$ dieser Fläche bestimmen?

Wir beginnen zunächst mit einem Spezialfall, wo die Fläche F durch die Geraden $x = a$, $x = b$ und die Funktion $f : [a, b] \rightarrow \mathbb{R}$ begrenzt ist.

Die Idee besteht dann darin, diese Fläche durch Rechtecke auszuschöpfen.

Wir wählen eine geeignete Zerlegung Z des Intervalls $I = [a, b]$, d.h. es sei

$$a = x_0 < x_1 < x_2 \cdots < x_n < x_{n+1} = b.$$

Somit ist die Zerlegung Z durch die Zerlegungspunkte $\{x_j\}_{j=0}^{n+1}$ definiert. Weiterhin definieren wir

$$I_j = [x_j, x_{j+1}], \quad |I_j| = x_{j+1} - x_j \quad \text{für } j = 0, 1, \dots, n.$$

Falls die Funktion f beschränkt ist, d.h. es existieren reelle Zahlen m, M mit $m \leq f(x) \leq M$ für alle $x \in [a, b]$, so ist anschaulich klar, dass

$$-\infty < m(b-a) \leq \underbrace{\sum_{j=0}^n \inf_{x \in I_j} f(x) |I_j|}_{U(f,Z)} \leq |F| \leq \underbrace{\sum_{j=0}^n \sup_{x \in I_j} f(x) |I_j|}_{O(f,Z)} \leq M(b-a) < +\infty.$$

Dabei bezeichnet $U(f, Z)$ die Untersumme und $O(f, Z)$ die Obersumme für die Funktion f bezüglich der Zerlegung Z .

Daraus folgt, dass i.a. $|F|$ besser (auf jeden Fall nicht schlechter) approximiert wird, je *feiner* die Zerlegung Z ist. Sei $i \in \mathbb{N}$. Dann betrachten wir $O(f, Z_i)$ und $U(f, Z_i)$ für Zerlegungen

$Z_i = \{x_j^{(i)}\}_{j=0}^{n_i+1}$ mit

$$d_i = \max_{j=0,1,\dots,n_i} |I_j^{(i)}| = \max_{j=0,1,\dots,n_i} |x_{j+1}^{(i)} - x_j^{(i)}| \rightarrow 0 \text{ für } i \rightarrow \infty, \quad (1)$$

wobei die entsprechenden Intervalle $I_j^{(i)}$ durch

$$I_j^{(i)} = [x_j^{(i)}, x_{j+1}^{(i)}], \quad j = 0, 1, \dots, n_i,$$

gegeben sind.

Definition 1 Sei $-\infty < a < b < \infty$. Die Funktion f sei beschränkt auf $[a, b]$. Falls für jede Folge von Zerlegungen $\{Z_i\}_{i=1}^{\infty}$ mit (1) die beiden Grenzwerte

$$\begin{aligned} \lim_{i \rightarrow \infty} U(f, Z_i) &= \lim_{i \rightarrow \infty} \sum_{j=0}^{n_i} \inf_{x \in I_j^{(i)}} f(x) |I_j^{(i)}|, \\ \lim_{i \rightarrow \infty} O(f, Z_i) &= \lim_{i \rightarrow \infty} \sum_{j=0}^{n_i} \sup_{x \in I_j^{(i)}} f(x) |I_j^{(i)}| \end{aligned}$$

existieren und gleich sind, dann heißt die Funktion $f(x)$ auf dem Intervall $[a, b]$ integrierbar.

Das bestimmte Riemannsche Integral ist dann durch

$$\int_a^b f(x) dx := \lim_{i \rightarrow \infty} U(f, Z_i) = \lim_{i \rightarrow \infty} O(f, Z_i) \quad (2)$$

definiert.

Bemerkung 1 (a) Es ist also zu beachten, dass das bestimmte Riemannsche Integral unabhängig von einer konkreten Folge von Zerlegungen $\{Z_i\}_{i=1}^{\infty}$ mit der Eigenschaft (1) eingeführt wurde.

(b) Nicht jede beschränkte Funktion ist auf einem endlichen Intervall $[a, b]$ integrierbar. Man betrachte z.B. die Funktion (siehe Vorkurs, Abschnitt 3. Funktionen)

$$f(x) = \begin{cases} -1 & : x \in \mathbb{Q} \cap [a, b] \\ +1 & : x \in \mathbb{R} \setminus \mathbb{Q} \cap [a, b]. \end{cases}$$

Diese Funktion ist zwar in \mathbb{R} beschränkt ($|f| = 1$), aber es gilt für jede Zerlegung Z_i

$$O(f, Z_i) = 1 \cdot (b - a) \neq (-1) \cdot (b - a) = U(f, Z_i).$$

Wir geben jetzt eine hinreichend große Klasse von Funktionen an, die integrierbar sind, und für die meisten Untersuchungen ausreichend sind.

Satz 1 Es sei $f(x)$ stetig auf dem endlichen Intervall $[a, b]$. Dann existiert

$$\int_a^b f(x) dx.$$

Bemerkung 2 Anschaulich ist diese Aussage leicht einzusehen. Falls (1) gilt, so folgt aus der Stetigkeit, dass

$$\max_{x \in I_j^{(i)}} f(x), \quad \min_{x \in I_j^{(i)}} f(x)$$

existieren und

$$\max_{x \in I_j^{(i)}} f(x) - \min_{x \in I_j^{(i)}} f(x) \rightarrow 0 \quad \text{für } i \rightarrow \infty$$

gilt. Damit gibt es für alle $\varepsilon > 0$ eine $i_0(\varepsilon)$, so dass $O(f, Z_i) - U(f, Z_i) \leq \varepsilon(b-a)$, falls $i \geq i_0(\varepsilon)$.

Nach den Ergebnissen aus dem Vorkurs, siehe Abschnitt 3. Funktionen, ist jede dort eingeführte elementare Funktion auf ihrem Definitionsbereich stetig und somit auf entsprechenden endlichen Intervallen integrierbar.

Wir geben jetzt einfache Eigenschaften des bestimmten Integrals an, die aus Definition 1 folgen.

Satz 2 *Es seien $f(x)$ und $g(x)$ auf $[a, b]$ integrierbar.*

(a) *Falls $\lambda, \mu \in \mathbb{R}$, so ist auch $(\lambda f + \mu g)(x)$ auf $[a, b]$ integrierbar, und es gilt*

$$\int_a^b (\lambda f + \mu g)(x) dx = \lambda \int_a^b f(x) dx + \mu \int_a^b g(x) dx.$$

(b) *Es sei $a < c < b$. Dann ist*

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Folgerung 1 *Jede auf $[a, b]$ stückweise stetige Funktion ist integrierbar.*

Beweis Die Funktion f ist auf $[a, b]$ stückweise stetig, falls f auf $[a, b]$ höchstens endlich viele Unstetigkeitsstellen hat, die alle endliche Sprungstellen sind (siehe Vorkurs). Falls f eine endliche Sprungstelle in $x = c$ mit $a < c < b$ hat, so verwende man $[a, c]$ und $[c, b]$ in Satz 1 und Satz 2(b). ■

Für die weiteren Untersuchungen verwenden wir folgende *Vereinbarungen*:

$$\int_a^a f(x) dx =: 0, \quad \int_b^a f(x) dx =: - \int_a^b f(x) dx,$$

falls $-\infty < a < b < \infty$ gilt und $f(x)$ auf $[a, b]$ integrierbar ist.

Beispiel 1 (Seminar)

Man zeige unter Verwendung der Definition 1, dass

$$\int_0^1 e^x dx = e - 1$$

ist, wobei man eine äquidistante Zerlegung verwende!

Wir geben weitere Eigenschaften des Riemannsches Integrals an, die ebenfalls aus Definition 1 folgen.

Satz 3 *Es seien $f(x)$ und $g(x)$ auf $[a, b]$ integrierbar.*

(a) *Falls $f(x) \leq g(x)$ für alle $x \in [a, b]$ gilt, so ist*

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

(b) *$|f(x)|$ ist integrierbar, und es gilt*

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

(c) $\int_a^b 1 dx = b - a$.

(d) *Ist $m \leq f(x) \leq M$ für alle $x \in [a, b]$, so erhalten wir*

$$m(b - a) \leq \int_a^b f(x) dx \leq M(b - a).$$

Bemerkung 3 (a) Nach Aussage (b) gilt: Falls f integrierbar ist, so auch $|f|$. Man beachte aber, dass die Umkehrung i.a. nicht gilt. Dazu nehme man die Funktion aus Bemerkung 1

$$f(x) = \begin{cases} -1 & : x \in \mathbb{Q} \cap [a, b] \\ +1 & : x \in \mathbb{R} \setminus \mathbb{Q} \cap [a, b], \end{cases}$$

die nicht integrierbar ist. Offensichtlich ist aber nach Satz 3(c) die Funktion $|f| = 1$ auf $[a, b]$ integrierbar.

(b) Bisher sind wir nur in der Lage, das bestimmte Integral

$$\int_a^b f(x) dx$$

mittels Definition 1 zu berechnen. Das ist jedoch ziemlich aufwendig, wie schon das einfache Beispiel 1 zeigt. Deshalb sind wir an einer anderen Methode interessiert, die auf dem Zusammenhang zwischen Integration und Differentiation beruht und zum Hauptsatz der Differential- und Integralrechnung führt.

Satz 4 (Hauptsatz der Differential- und Integralrechnung) Die Funktion $f(x)$ sei auf $[a, b]$ stetig. Dann ist

$$F(x) = \int_a^x f(y)dy, \quad a \leq x \leq b, \quad (3)$$

in (a, b) differenzierbar, und es gilt $F'(x) = f(x)$. F heißt Stammfunktion von f .

Bemerkung 4 (a) Die Stammfunktion $F(x)$ wird definiert als bestimmtes Integral von f als Funktion von der oberen Grenze x mit $a \leq x \leq b$, d.h. auf dem Intervall $I_x = [a, x]$. Diese Definition ist sinnvoll, da f auf $[a, b]$ stetig und somit auf allen Intervallen I_x , $a \leq x \leq b$, integrierbar ist.

(b) Die Aussage des Hauptsatzes der Differential- und Integralrechnung ist also, dass das bestimmte Integral einer auf $[a, b]$ stetigen Funktion f als Funktion der oberen Grenze

$$F(x) = \int_a^x f(y)dy$$

Stammfunktion von f ist, d.h. es gilt (zuerst Integration, dann Differentiation)

$$\frac{d}{dx}F(x) = \frac{d}{dx} \int_a^x f(y)dy = f(x).$$

Die Ableitung des bestimmten Integrals der Funktion f nach der oberen Grenze x ergibt also den Wert des Integranden an der Stelle x , d.h. $f(x)$.

Unter Verwendung des letzten Satzes können wir folgende Regel für die Berechnung von bestimmten Integralen mittels Stammfunktion herleiten.

Satz 5 Es seien die Funktionen f und g stetig auf $[a, b]$. Falls f differenzierbar und $f'(x) = g(x)$ ist, so gilt

$$\int_a^b g(x)dx = f(b) - f(a) := f(x) \Big|_a^b. \quad (4)$$

Bemerkung 5 (a) Unter Berücksichtigung von Satz 4 und Satz 5, d.h. (zuerst Differentiation, dann Integration)

$$f(b) - f(a) = \int_a^b f'(x)dx,$$

kann man die Differentiation als Umkehrung der Integration auffassen. Weiterhin liefert der Satz 5 für uns ein bequemes und praktikables Mittel zur Berechnung von bestimmten Riemannschen Integralen mittels Stammfunktionen, d.h. wir sind nicht mehr auf die Methode der Obersummen und Untersummen angewiesen (Definition 1).

(b) Falls $F(x)$ Stammfunktion von $f(x)$ ist, so gilt das auch für die Funktion $F(x) + c$, $c \in \mathbb{R}$. Somit ist $\{F(x) + c : c \in \mathbb{R}\}$ die Menge aller Stammfunktionen einer stetigen Funktion $f(x)$.

Wir verwenden im folgenden als Schreibweise für die Stammfunktionen *das unbestimmte Integral* $\int f(x)dx$.

(c) Nach den Ergebnissen aus dem Vorkurs, Abschnitt 4. Differenzierbare Funktionen, gilt also

$$\begin{aligned}\int x^\alpha dx &= \frac{x^{\alpha+1}}{\alpha+1} + c, \quad \alpha \neq -1, \\ \int \frac{1}{x} dx &= \ln|x| + c, \quad x \neq 0, \\ \int e^x dx &= e^x + c, \\ \int \frac{dx}{1+x^2} &= \arctan x + c\end{aligned}$$

Beispiel 2 (a) Zunächst eine Bemerkung zur Stammfunktion von $\frac{1}{x}$. Wir erhalten für $x > 0$

$$(\ln|x|)' = (\ln x)' = \frac{1}{x}$$

und für $x < 0$ wegen $-x > 0$

$$(\ln|x|)' = (\ln(-x))' = \frac{1}{-x} \cdot (-1) = \frac{1}{x}.$$

(b) Es gilt

$$\int_a^b x^n dx = \frac{x^{n+1}}{n+1} \Big|_a^b = \frac{b^{n+1} - a^{n+1}}{n+1},$$

falls $n \neq -1$.

(c) Analog erhält man

$$\int_0^{\pi/2} \cos x dx = \sin x \Big|_0^{\pi/2} = 1.$$

Folgerung 2 (Mittelwertsatz der Integralrechnung) *Es sei $f(x)$ stetig auf $[a, b]$. Dann existiert ein $\xi \in (a, b)$, so dass*

$$\int_a^b f(x) dx = f(\xi)(b-a). \quad (5)$$

Beweis Es sei $F(x)$ die Stammfunktion von $f(x)$. Unter Verwendung des Mittelwertsatzes der Differentialrechnung, siehe Vorkurs, Abschnitt 4. Differenzierbare Funktionen, existiert ein $\xi \in (a, b)$ mit

$$\int_a^b f(x) dx = F(b) - F(a) = F'(\xi)(b-a) = f(\xi)(b-a).$$

■

(Teig ausrollen)

2.2 Integrationsverfahren

Bisher sind wir nur in der Lage, solche Integrale zu berechnen, die sich aus einfachen Grundintegralen zusammensetzen, um Satz 5 anwenden zu können, d.h. es muss die Stammfunktion bekannt sein. Wir wollen im folgenden kompliziertere Integrale bestimmen, die sich mit geeignete Umformungen auf Grundintegrale oder bereits bekannte Integrale zurückführen lassen. Wir verweisen jedoch darauf, dass nicht jedes Integral zu bekannten Stammfunktionen führt. So lässt sich das bestimmte Integral

$$\int_0^x \frac{\sin t}{t} dt$$

nicht durch elementare Funktionen darstellen, sondern definiert eine neue Funktion, die man *Integralsinus* nennt. Es gilt

$$\int_0^x \frac{\sin t}{t} dt = \text{Si}(x) = x - \frac{x^3}{3 \cdot 3!} + \frac{x^5}{5 \cdot 5!} \mp \dots$$

Andererseits gibt es viele Formelsammlungen, in denen man bezüglich der Stammfunktion nachschlagen kann. Unser Ziel besteht jetzt darin, einige wichtige Standardtechniken darzustellen. Wir beginnen mit der *partiellen Integration*.

Satz 6 *Es seien $f'(x)$ und $g'(x)$ auf $[a, b]$ stetig. Dann gilt*

$$\int_a^b f'(x)g(x)dx = f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x)dx. \quad (6)$$

Beweis Wir erhalten einerseits nach Satz 5

$$\int_a^b (f \cdot g)'(x)dx = f(x)g(x)|_a^b = f(b)g(b) - f(a)g(a)$$

sowie nach der Differentiationsregel für Produkte, siehe Vorkurs, Abschnitt 4. Differenzierbare Funktionen,

$$\int_a^b (f \cdot g)'(x)dx = \int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx. \quad \blacksquare$$

Beispiel 3 (a) Wir berechnen

$$\int_a^b x \cdot \cos x dx.$$

Wir setzen $g(x) = x$ und $f'(x) = \cos x$. Somit gilt $g'(x) = 1$ und $f(x) = \sin x$. Wir erhalten

$$\int_a^b x \cdot \cos x dx = x \sin x|_a^b - \int_a^b \sin x dx = x \sin x|_a^b + \cos x|_a^b.$$

(b) Analog kann man Integrale der Form

$$\int x^k \sin x dx, \int x^k \cos x dx, \int x^k e^x dx$$

berechnen, falls $k \in \mathbb{N}$ gilt. In diesem Falle verwende man mehrmals das Verfahren aus (a).

(c) Es sei $0 < a < b < \infty$ und $\alpha \geq 0$. Wir berechnen

$$\int_a^b x^\alpha \ln x dx.$$

Wir setzen $f'(x) = x^\alpha$ und $g(x) = \ln x$. Dann gilt

$$f(x) = \frac{x^{\alpha+1}}{\alpha+1}, \quad g'(x) = \frac{1}{x}$$

und somit

$$\begin{aligned} \int_a^b x^\alpha \ln x dx &= \int_a^b \left(\frac{x^{\alpha+1}}{\alpha+1} \right)' \ln x dx = \frac{x^{\alpha+1}}{\alpha+1} \ln x \Big|_a^b - \int_a^b \frac{x^{\alpha+1}}{\alpha+1} \frac{1}{x} dx \\ &= \left(\frac{x^{\alpha+1}}{\alpha+1} \ln x - \frac{x^{\alpha+1}}{(\alpha+1)^2} \right) \Big|_a^b. \end{aligned}$$

Insbesondere erhalten wir für $\alpha = 0$

$$\int_a^b \ln x dx = (x \ln x - x) \Big|_a^b.$$

(d) Gegeben sei das unbestimmte Integral

$$\int \cos^2 t dt.$$

Wir setzen $f'(t) = g(t) = \cos t$. Damit gilt $f(t) = \sin t$ und $g'(t) = -\sin t$. Unter Verwendung von

$$\sin^2 t + \cos^2 t = 1$$

folgt

$$\begin{aligned} \int \cos^2 t dt &= \int \cos t \cos t dt = \sin t \cos t + \int \sin^2 t dt \\ &= \sin t \cos t + \int (1 - \cos^2 t) dt = \sin t \cos t + t - \int \cos^2 t dt + C, \end{aligned}$$

d.h.

$$\int \cos^2 t dt = \frac{1}{2}(\sin t \cos t + t) + C.$$

Insbesondere gilt

$$\int_0^{\pi/2} \cos^2 t dt = \frac{1}{2}(\sin t \cos t + t) \Big|_0^{\pi/2} = \frac{\pi}{4}.$$

Die Idee besteht also darin, auf beiden Seiten das gleiche Integral, jedoch mit unterschiedlichem Vorzeichen, zu haben.

Analog erhält man

$$\begin{aligned}\int \cosh^2 t dt &= \frac{1}{2}(\sinh t \cosh t + t) + C \\ \int \sinh^2 t dt &= \frac{1}{2}(\sinh t \cosh t - t) + C.\end{aligned}$$

Als nächstes verwenden wir zur Berechnung von Integralen die Methode der *Variablensubstitution*.

Satz 7 Es sei $g(x)$ stetig differenzierbar auf $[a, b]$ sowie $f(x)$ stetig auf $g([a, b])$. Dann gilt

$$\int_a^b f(g(u)) \cdot g'(u) du = \int_{g(a)}^{g(b)} f(x) dx. \quad (7)$$

Beweis Es sei F Stammfunktion von f : $F'(x) = f(x)$. Für $x = g(u)$ erhalten wir dann mittels Kettenregel, siehe Vorkurs, Abschnitt 4. Differenzierbare Funktionen,

$$f(x) = F'(x) = (F(g(u)))' = F'(g(u)) \cdot g'(u) = f(g(u)) \cdot g'(u)$$

und damit

$$\int_{g(a)}^{g(b)} f(x) dx = \int_a^b f(g(u)) g'(u) du.$$

■

Beispiel 4 (a) Wir untersuchen

$$\int_a^b e^{\sin u} \cos u du.$$

Es gilt $f(x) = e^x$. Setzen wir $x = \sin u = g(u)$, so erhalten wir wegen $g'(u) = \cos u$, dass $dx = g'(u) du = \cos u du$. Daraus folgt aber

$$\int_a^b e^{\sin u} \cos u du = \int_{\sin a}^{\sin b} e^x dx = e^x \Big|_{\sin a}^{\sin b} = e^{\sin b} - e^{\sin a}.$$

(b) Es sei $f(x) \neq 0$. Für Integrale der Form

$$\int \frac{f'(x)}{f(x)} dx$$

erhält mittels Substitution $z = f(x)$, d.h. $dz = f'(x) dx$,

$$\int \frac{f'(x)}{f(x)} dx = \int \frac{dz}{z} = \ln |z| + c = \ln |f(x)| + c.$$

Eine einfaches Anwendungsbeispiel ist

$$\int \cot x dx = \int \frac{\cos x}{\sin x} dx = \ln |\sin x| + C.$$

(c) Falls als Integrand $f(ax + b)$ auftritt, so ergibt sich mit $y = g(x) = ax + b$ und $dy = a dx$

$$\int f(ax + b) dx = \frac{1}{a} \int f(y) dy = \frac{1}{a} F(y) + c = \frac{1}{a} F(ax + b) + C,$$

wobei F die Stammfunktion von f ist.

Merkregel:

Es gilt

$$\int_a^b f(g(u))g'(u)du \quad \rightleftharpoons \quad \int_{g(a)}^{g(b)} f(x)dx,$$

falls die Struktur vorliegt, indem man $g(u) = x$ setzt, und somit $g'(u)du = dx$ erhält.

Man erhält umgekehrt

$$\int_c^d f(x)dx \quad \rightleftharpoons \quad \int_{g^{-1}(c)=\alpha}^{g^{-1}(d)=\beta} f(g(u))g'(u)du,$$

indem man $x = g(u)$, $u = g^{-1}(x)$ und $dx = g'(u)du$ setzt. Dafür ist aber notwendig, dass auch die inverse Funktion g^{-1} auf $[\alpha, \beta]$ existiert (ausreichend z.B. $g'(u) \neq 0$ auf $[\alpha, \beta]$, siehe Vorkurs, Abschnitt 4. Differenzierbare Funktionen).

Bisher haben wir nur die erste Methode verwendet. Eine größere Bedeutung hat jedoch das zweite Verfahren der *Variablensubstitution*. Wir weisen darauf hin, dass es leider keine allgemeine Regel für die Integration der elementaren Funktionen **nicht** gibt. Eine gewisse Routine in der Integration erhält man durch Üben. Wir wollen jetzt zeigen, wie man die Stammfunktionen für

$$\int \sqrt{1-x^2} dx, |x| \leq 1, \quad \int \sqrt{x^2-1} dx, |x| \geq 1, \quad \int \sqrt{1+x^2} dx, x \in \mathbb{R},$$

die man in den Formelsammlungen findet, mittels bestimmter (jedoch verschiedener) Ansätze erhält. Wir bemerken, dass dagegen die Ableitungen dieser Funktionen sehr ähnlich sind. Das zeigt noch einmal, dass das Integrieren gegenüber dem Differenzieren viel komplizierter ist.

Substitution trigonometrischer und hyperbolischer Funktionen

Die Substitution $x = \sin z$

Wir starten mit dem Integral

$$\int \sqrt{1-x^2} dx, |x| \leq 1.$$

In diesem Fall verwendet man die Substitution $x = \sin z$. Nach den Resultaten aus dem Vorkurs, siehe Abschnitt 3. Funktionen, ist $\sin z$ auf $[-\pi/2, +\pi/2]$ streng monoton wachsend und besitzt dort die Umkehrfunktion $z = \arcsin x$. Wir erhalten

$$dx = \cos z dz, \quad \sqrt{1-x^2} = \sqrt{1-\sin^2 z} = \cos z.$$

Unter Verwendung von Beispiel 3(d) folgt

$$\begin{aligned}
 \int \sqrt{1-x^2} dx &= \int \cos z \cos z dz = \frac{\cos z \sin z + z}{2} + C \\
 &= \frac{\cos z \cdot x + \arcsin x}{2} + C \\
 &= \frac{\sqrt{1-\sin^2 z} \cdot x + \arcsin x}{2} + C \\
 &= \frac{\sqrt{1-x^2} \cdot x + \arcsin x}{2} + C.
 \end{aligned} \tag{8}$$

Beispiel 5 Wir berechnen den Flächeninhalt $|F|$ eines Kreises vom Radius r in Nullpunkt-Lage.

Zunächst haben wir $x^2 + y^2 = r^2$. Daraus folgt, dass für die obere Kreishälfte $y = +\sqrt{r^2 - x^2}$, $|x| \leq r$ gilt. Wir substituieren $x = r \sin z$ und erhalten $dx = r \cos z dz$. Damit gilt

$$x = 0 \iff z = 0, \quad x = r \iff z = \pi/2.$$

Aus Symmetriegründen erhalten wir für den Flächeninhalt, wieder unter Verwendung von Beispiel 3(d),

$$\begin{aligned}
 |F| &= 2 \int_{-r}^r \sqrt{r^2 - x^2} dx = 4 \int_0^r \sqrt{r^2 - x^2} dx \\
 &= 4 \int_0^{\pi/2} \sqrt{r^2 - r^2 \sin^2 z} \cdot r \cos z dz = 4 \int_0^{\pi/2} r^2 \cos^2 z dz \\
 &= 4r^2 \cdot \frac{1}{2} (\sin z \cos z + z) \Big|_0^{\pi/2} = 4r^2 \frac{\pi}{4} = \pi r^2.
 \end{aligned} \tag{9}$$

Abschließend bemerken wir, dass auch Integrale der allgemeineren Klasse

$$\int f(x, \sqrt{a^2 - x^2}) dx, \quad |x| \leq a,$$

mittels der Substitution $x = a \sin z$ berechnet werden können (siehe Formelsammlungen).

Für die obige Substitution war die Beziehung

$$\sin^2 z + \cos^2 z = 1 \implies \sqrt{1 - \sin^2 z} = \cos z$$

entscheidend.

Wir betrachten jetzt die aus dem Vorkurs, Abschnitt 3. Funktionen, bekannten *Hyperbelfunktionen* $\sinh x$ und $\cosh x$, und deren Umkehrfunktionen $\operatorname{arsinh} x$ und $\operatorname{arcosh} x$.

Die Substitution $x = \sinh z$.

Wir betrachten jetzt Integrale der Form

$$\int \sqrt{x^2 + 1} dx, x \in \mathbb{R}.$$

Mit der Substitution $x = \sinh z \iff z = \operatorname{arsinh} x = \ln(x + \sqrt{x^2 + 1})$ (Vorkurs, Abschnitt 3) erhalten wir nach dem Vorkurs, Abschnitt 4. Differenzierbare Funktionen

$$dx = \cosh z dz, \sqrt{x^2 + 1} = \sqrt{\sinh^2 z + 1} = \cosh z.$$

Damit folgt mittels partieller Integration, analog zu Beispiel 3(d),

$$\int \sqrt{x^2 + 1} dx = \int \cosh z \cosh z dz = \frac{1}{2}(\sinh z \cosh z + z) + C.$$

Wegen $\cosh z = \sqrt{\sinh^2 z + 1} = \sqrt{x^2 + 1}$ ergibt sich durch Zurücksubstituieren

$$\int \sqrt{x^2 + 1} dx = \frac{1}{2}(x \cosh z + \operatorname{arsinh} x) + C = \frac{1}{2}(x \sqrt{x^2 + 1} + \ln(x + \sqrt{x^2 + 1})) + C.$$

Weiterhin kann man mit der Substitution $x = a \sinh z$ auch Integrale der allgemeineren Klasse

$$\int f(x, \sqrt{x^2 + a^2}) dx$$

berechnen (siehe Formelsammlungen).

Die Substitution $x = \cosh z$.

Wir wollen Integrale der Form

$$\int \sqrt{x^2 - 1} dx, 1 \leq x < +\infty,$$

berechnen. In diesem Fall verwenden wir die Substitution

$$x = \cosh z \iff z = \operatorname{arcosh} x = \ln(x + \sqrt{x^2 - 1})$$

und erhalten mit den Ergebnissen aus dem Vorkurs, Abschnitt 4. Differenzierbare Funktionen

$$dx = \sinh z dz, \sqrt{x^2 - 1} = \sqrt{\cosh^2 z - 1} = \sinh z.$$

Analog zu oben erhalten wir damit mittels partieller Integration und durch Zurücksubstituieren

$$\begin{aligned} \int \sqrt{x^2 - 1} dx &= \int \sinh z \sinh z dz = \frac{1}{2}(\sinh z \cosh z - z) + C \\ &= \frac{1}{2}(x \sqrt{x^2 - 1} - \operatorname{arcosh} x) + C \\ &= \frac{1}{2}(x \sqrt{x^2 - 1} - \ln(x + \sqrt{x^2 - 1})) + C. \end{aligned}$$

Abschließend wollen wir bemerken, dass es noch viele andere Substitutionen für die Integration von speziellen Klassen von Funktionen gibt (z.B. für trigonometrische Ausdrücke), die man in Formelsammlungen finden kann. Wir wollen hier nur noch die Integration von rationalen Funktionen betrachten.

2.3 Integration gebrochen-rationaler Funktionen

Nach Abschnitt 3. Funktionen aus dem Vorkurs ist eine *gebrochen-rationale Funktion* $f(x)$ durch

$$f(x) = \frac{P(x)}{Q(x)}$$

gegeben, wobei $P(x)$ und $Q(x)$ ganze rationale Funktionen (Polynome) bezüglich der Variablen x sind. Ein Polynom $P(x)$ hat den *Grad* n ($\text{Grad}(P) = n$), falls

$$P(x) = \sum_{j=0}^n a_j x^j, \quad a_n \neq 0,$$

gilt. Unter Verwendung der Polynomdivision lässt sich jede unecht-gebrochene rationale Funktion als Summe eines Polynoms und einer echt-gebrochenen rationalen Funktion darstellen, d.h. es gilt

$$f(x) = \frac{P(x)}{Q(x)} = R(x) + \frac{S(x)}{Q(x)}, \quad (10)$$

wobei $R(x)$ als sogenannter *ganzer Teil* ein Polynom und der Rest $S(x)$ entweder eine *echt-gebrochene rationale Funktion* mit

$$\text{Grad}(S) < \text{Grad}(Q)$$

ist, oder es gilt $S(x) \equiv 0$. Dazu verwendet man den *euklidischen Algorithmus*.

Satz 8 *Es seien $P(x)$ und $Q(x)$ reelle Polynome mit $Q(x) \not\equiv 0$. Dann existieren eindeutig bestimmte Polynome $S(x)$ und $R(x)$ mit*

$$P(x) = R(x)Q(x) + S(x),$$

wobei für $S(x)$ entweder $S(x) \equiv 0$ oder $\text{Grad}(S) < \text{Grad}(Q)$ gilt.

Wir geben ein Beispiel für die Anwendung des euklidischen Algorithmus.

Beispiel 6 Wir betrachten $P(x) = x^4 + 1$ und $Q(x) = x - 1$. Dann kann man folgendes Rechenschema verwenden.

$$\begin{aligned} (x^4 + 1) : (x - 1) &= x^3 + x^2 + x + 1 \\ -(x^4 - x^3) & \\ + (x^3 + 1) & \\ -(x^3 - x^2) & \\ + (x^2 + 1) & \end{aligned}$$

$$\begin{aligned}
& -(x^2 - x) \\
& \quad + (x + 1) \\
& \quad \quad - (x - 1) \\
& \quad \quad \quad 2
\end{aligned}$$

Damit erhalten wir

$$(x^4 + 1) = (x^3 + x^2 + x + 1)(x - 1) + 2,$$

d.h.

$$\frac{x^4 + 1}{x - 1} = x^3 + x^2 + x + 1 + \frac{2}{x - 1}.$$

Es gilt also $R(x) = x^3 + x^2 + x + 1$ und $S(x) = 2$. Das Polynom $S(x)$ bezeichnet also den *Rest* bei der Division von $P(x)$ durch $Q(x)$. Falls $S(x) \equiv 0$ ist, so heißt $Q(x)$ *Teiler* von $P(x)$, d.h. es existiert $R(x)$ mit $P(x) = Q(x) \cdot R(x)$.

Bemerkung 6 Aus Satz 8 folgt nach Division durch $Q(x)$ die gewünschte Darstellung (10). Unser Ziel besteht darin, rationale Funktionen $f(x)$ zu integrieren. Somit reduziert sich der allgemeine Fall $\text{Grad}(P) \geq \text{Grad}(Q)$ nach Anwendung des euklidischen Algorithmus auf den Fall der Integration von echt-gebrochenen rationalen Funktionen, da die Integration von Polynomen bekannt ist.

Deshalb wenden wir uns jetzt der *Partialbruchzerlegung echt-gebrochener rationalen Funktionen* zu, d.h. wir setzen voraus, dass für

$$f(x) = \frac{P(x)}{Q(x)}$$

$\text{Grad}(P) < \text{Grad}(Q)$ gilt. (Ansonsten ist die Partialbruchzerlegung **nicht** möglich!)

Das Ziel der Partialbruchzerlegung besteht darin, diese rationale Funktion $f(x)$ als Summe von einfachen Brüchen, ihren *Partialbrüchen*, darzustellen.

Wir werden im folgenden sehen, dass es nur zwei Typen von Partialbrüchen gibt.

- Typ (I): $\frac{A}{(x - x_0)^k}$, $k = 1, 2, \dots$, für reelle Nullstellen,
- Typ (II): $\frac{Bx + C}{(x^2 + bx + c)^k}$, $k = 1, 2, \dots$, für komplexe Nullstellen,

wobei die auftretenden Koeffizienten A, B, C, b und c alle *reell* sind. Ein wichtiges Hilfsmittel, um diese Aussage beweisen zu können, stellt der *Fundamentalsatz der Algebra* dar.

Satz 9 (Fundamentalsatz der Algebra) *Es sei $z \in \mathbb{C}$. Dann besitzt jedes Polynom n -ten Grades*

$$P_n(z) = \sum_{j=0}^n a_j z^j, \quad a_j \in \mathbb{R}, \quad j = 0, \dots, n, \quad a_n \neq 0, \quad n \in \mathbb{N},$$

bei geeigneter Zählweise genau n komplexe Nullstellen, d.h. es existieren genau n eindeutig bestimmte Zahlen $z_1, \dots, z_n \in \mathbb{C}$, mit

$$\sum_{j=0}^n a_j z^j = a_n (z - z_1) \cdot (z - z_2) \cdot \dots \cdot (z - z_n). \quad (11)$$

Bemerkung 7 (a) Dieser Satz wurde von Gauss im Alter von 21 Jahren in seiner Dissertation 1798 bewiesen.

(b) Bei geeigneter Zählweise heißt, dass z.B. $z_1 = z_2$ zugelassen ist. Es werden also mehrfache Nullstellen entsprechend ihrer *algebraischen Vielfachheit* gezählt. Somit lässt sich (11) auch schreiben als

$$\sum_{j=0}^n a_j z^j = a_n (z - w_1)^{n_1} \cdot \dots \cdot (z - w_m)^{n_m}, \quad (12)$$

wobei die Nullstelle w_i die Vielfachheit n_i , $i = 1, \dots, m$, hat, und $n_1 + \dots + n_m = n$ gilt.

Beispiel 7 (a) Es ist z.B.

$$z^7 - 2z^6 + z^5 = z^5(z^2 - 2z + 1) = z^5(z - 1)^2.$$

Damit hat die Nullstelle $z_1 = 0$ die Vielfachheit $n_1 = 5$ und die Nullstelle $z_2 = 1$ die Vielfachheit $n_2 = 2$.

(b) Für $z^2 + 1$ erhalten wir die konjugiert komplexen Nullstellen

$$z_1 = i, \quad z_2 = -i$$

mit Vielfachheit $n_1 = n_2 = 1$.

Aus Satz 9 lassen sich folgende Resultate ableiten. Dabei verwenden wir folgenden Zusammenhang. Falls w_1 mit $\operatorname{Im} w_1 > 0$ eine komplexe Nullstelle ist, so auch \bar{w}_1 mit $\operatorname{Im} w_1 < 0$. Somit erhalten wir nach der dritten binomischen Formel für reelles x

$$(x - w_1)(x - \bar{w}_1) = \left((x - \operatorname{Re} w_1) - i \operatorname{Im} w_1 \right) \cdot \left((x - \operatorname{Re} w_1) + i \operatorname{Im} w_1 \right) = (x - \operatorname{Re} w_1)^2 + (\operatorname{Im} w_1)^2.$$

Folgerung 3 (a) *Ist w eine komplexe Nullstelle von $P_n(z)$, so auch \bar{w} . Die Vielfachheiten von w und \bar{w} stimmen überein.*

(b) Sind v_1, \dots, v_k die reellen Nullstellen und w_1, \dots, w_l die komplexen Nullstellen von $P_n(z)$ mit der Eigenschaft $\text{Im } w_j > 0$, $j = 1, \dots, l$, so gilt für alle **reellen** x

$$P_n(x) = a_n(x - v_1) \cdots (x - v_k) \cdot [(x - \text{Re } w_1)^2 + (\text{Im } w_1)^2] \cdots [(x - \text{Re } w_l)^2 + (\text{Im } w_l)^2],$$

d.h. für reelle Polynome existiert eine reelle Darstellung.

Unter Verwendung dieses Ergebnisses lässt sich jetzt die Partialbruchzerlegung einer echt-gebrochenen rationalen Funktion

$$f(x) = \frac{P(x)}{Q(x)}, \quad \text{Grad}(Q) = n, \quad n \in \mathbb{N},$$

bestimmen.

Beispiel 8 Fall a):

Alle Nullstellen von $Q(x)$ sind reell und einfach, d.h. wir haben

$$Q(x) = a_n(x - x_1)(x - x_2) \cdots (x - x_n).$$

Dann verwenden wir den Ansatz

$$\frac{P(x)}{Q(x)} = \frac{A_1}{x - x_1} + \cdots + \frac{A_n}{x - x_n},$$

wobei die reellen Koeffizienten A_1, \dots, A_n geeignet gewählt werden. Wir erhalten also den Typ (I) mit $k = 1$.

Beispiel (a):

$$\frac{x - 1}{(x + 1)(x - 2)}. \tag{13}$$

In diesem Fall hat $Q(x) = (x + 1)(x - 2)$ die einfachen reellen Nullstellen $x_1 = -1$ und $x_2 = 2$.

Mit dem Ansatz

$$\frac{x - 1}{(x + 1)(x - 2)} = \frac{A_1}{x + 1} + \frac{A_2}{x - 2}$$

erhalten wir nach Multiplikation beider Seiten mit $(x + 1)(x - 2)$, wobei $x \neq -1$ und $x \neq 2$ gilt,

$$x - 1 = A_1(x - 2) + A_2(x + 1).$$

Für $x \rightarrow -1$ bzw. $x \rightarrow 2$ erhalten wir

$$-2 = A_1(-3),$$

$$1 = A_2 \cdot 3.$$

Nach dieser Grenzwertmethode erhalten wir

$$A_1 = \frac{2}{3}, \quad A_2 = \frac{1}{3}.$$

Durch Einsetzen dieser Koeffizienten in den obigen Ansatz überzeugt man sich leicht, dass die Koeffizienten richtig bestimmt worden sind. Wir erhalten also für (13) folgende Darstellung mit Partialbrüchen vom Typ I, wobei $k = 1$, ist.

$$\boxed{\frac{x-1}{(x+1)(x-2)} = \frac{\frac{2}{3}}{x+1} + \frac{\frac{1}{3}}{x-2}}. \quad (14)$$

Fall b):

Alle Nullstellen von $Q(x)$ sind reell. Wir nehmen an, dass die verschiedenen Nullstellen x_1, \dots, x_l die Vielfachheiten n_1, \dots, n_l mit $n_1 + \dots + n_l = n$ haben. Somit gilt

$$Q(x) = a_n(x - x_1)^{n_1}(x - x_2)^{n_2} \dots (x - x_l)^{n_l}.$$

Dann verwenden wir den Ansatz

$$\frac{P(x)}{Q(x)} = \frac{A_{11}}{x - x_1} + \frac{A_{12}}{(x - x_1)^2} + \dots + \frac{A_{1n_1}}{(x - x_1)^{n_1}} + \dots + \frac{A_{l1}}{x - x_l} + \frac{A_{l2}}{(x - x_l)^2} + \dots + \frac{A_{ln_l}}{(x - x_l)^{n_l}},$$

wobei die reellen Koeffizienten geeignet gewählt werden. Wir erhalten also den Typ (I) mit $k = 1, \dots, n_i$, $i = 1, \dots, l$.

Beispiel (b):

$$\frac{x - 1}{x^2(x - 2)}. \quad (15)$$

In diesem Fall hat $Q(x) = x^2(x - 1)$ die reelle Nullstelle $x_1 = 0$ mit Vielfachheit $n_1 = 2$ und die einfache reelle Nullstelle $x_2 = 2$. Der Ansatz lautet also

$$\frac{x - 1}{x^2(x - 2)} = \frac{A_{11}}{x} + \frac{A_{12}}{x^2} + \frac{A_{21}}{x - 2}.$$

Wir multiplizieren beide Seiten mit $x^2(x - 2)$, wobei $x \neq 0$ und $x \neq 2$ gilt. Damit erhalten wir

$$x - 1 = A_{11}x(x - 2) + A_{12}(x - 2) + A_{21}x^2.$$

Für $x \rightarrow 0$ bzw. $x \rightarrow 2$ erhalten wir

$$\begin{aligned} -1 &= A_{12}(-2), \\ 1 &= A_{21}4. \end{aligned}$$

Es gilt damit

$$A_{12} = \frac{1}{2}, \quad A_{21} = \frac{1}{4}.$$

Somit folgt

$$x - 1 = \frac{1}{2}(x - 2) + A_{11}x(x - 2) + \frac{1}{4}x^2 = \frac{1}{2}x - 1 + A_{11}x^2 - 2A_{11}x + \frac{1}{4}x^2.$$

Durch Koeffizientenvergleich bei der Potenz x^2 erhalten wir $0 = A_{11} + 1/4$, d.h.

$$A_{11} = -\frac{1}{4}.$$

Durch Einsetzen dieser Koeffizienten in den obigen Ansatz überzeugt man sich ebenfalls leicht, dass die Koeffizienten richtig bestimmt worden sind. Wir erhalten also folgende Partialbruchzerlegung vom Typ I:

$$\boxed{\frac{x-1}{x^2(x-2)} = \frac{-\frac{1}{4}}{x} + \frac{\frac{1}{2}}{x^2} + \frac{\frac{1}{4}}{x-2}.} \quad (16)$$

Fall c):

$Q(x)$ besitzt auch einfache komplexe Nullstellen. Falls $\alpha = c + di$, $d \neq 0$, eine derartige Nullstelle ist, so ergibt sich aus dem Fundamentalsatz der Algebra und deren Folgerungen, dass auch $\bar{\alpha} = c - di$ eine einfache komplexe Nullstelle ist. Es gilt

$$(x - \alpha)(x - \bar{\alpha}) = (x - c)^2 + d^2.$$

Dann verwenden wir den Ansatz

$$\frac{P(x)}{Q(x)} = \underbrace{\frac{A_{11}}{x - x_1} + \dots}_{\text{reelle Nullstellen}} + \frac{Bx + C}{(x - c)^2 + d^2} + \dots,$$

wobei die reellen Koeffizienten B und C geeignet gewählt werden. Wir haben also neben dem Typ (I) auch den Typ (II) mit $k = 1$.

Beispiel (c):

$$\frac{4}{x^3 + 4x}. \quad (17)$$

$Q(x) = x^3 + 4x = x(x^2 + 4)$ hat die einfache reelle Nullstelle $x_1 = 0$ und die beiden einfachen komplexen Nullstellen $\alpha = 2i$ und $\bar{\alpha} = -2i$. Es gilt

$$(x - \bar{\alpha})(x - \alpha) = x^2 + 4.$$

Wir verwenden den Ansatz

$$\frac{4}{x^3 + 4x} = \frac{4}{x(x^2 + 4)} = \frac{A}{x} + \frac{Bx + C}{x^2 + 4}.$$

Wir multiplizieren beide Seiten mit $x(x^2 + 4)$, wobei $x \neq 0$ und $x \neq \pm 2i$ gilt. Damit erhalten wir

$$4 = A(x^2 + 4) + (Bx + C)x.$$

Für $x \rightarrow 0$ erhalten wir

$$4 = 4A, \quad \implies A = 1.$$

Es gilt damit

$$4 = x^2 + 4 + Bx^2 + Cx.$$

Ein Koeffizientenvergleich bei x ergibt $0 = C$ sowie bei x^2 , dass $0 = 1 + B$ gelten muss. Daraus folgt also

$$C = 0, \quad B = -1.$$

Durch Einsetzen dieser Koeffizienten in den obigen Ansatz überzeugt man sich ebenfalls leicht, dass die Koeffizienten richtig bestimmt worden sind. Wir erhalten also folgende Partialbruchzerlegung vom Typ I und vom Typ II:

$$\boxed{\frac{4}{x^3+4x} = \frac{1}{x} + \frac{-x}{x^2+4}} \quad (18)$$

Fall d): $Q(x)$ besitzt auch mehrfache komplexe Nullstellen. Falls $\alpha = c + di$, $d \neq 0$, eine l -fache komplexe Nullstelle ist, so ergibt sich aus dem Fundamentalsatz der Algebra und deren Folgerungen, dass auch $\bar{\alpha} = c - di$ eine l -fache komplexe Nullstelle ist. Es gilt

$$(x - \alpha)^l (x - \bar{\alpha})^l = ((x - c)^2 + d^2)^l.$$

Dann verwenden wir den Ansatz

$$\frac{P(x)}{Q(x)} = \underbrace{\frac{A_{11}}{x - x_1} + \dots}_{\text{reelle Nullstellen}} + \frac{B_1x + C_1}{(x - c)^2 + d^2} + \frac{B_2x + C_2}{((x - c)^2 + d^2)^2} + \dots + \frac{B_lx + C_l}{((x - c)^2 + d^2)^l} + \dots,$$

wobei die reellen Koeffizienten B_i und C_i , $i = 1, \dots, l$, geeignet gewählt werden. Wir haben also neben dem Typ (I) auch den Typ (II) mit $k = 1, \dots, k = l$.

Beispiel (d):

$$\frac{x^2 - 1}{(x^2 + 1)^2}. \quad (19)$$

$Q(x) = (x^2 + 1)^2 = ((x - i)(x + i))^2$ hat die beiden doppelten komplexen Nullstellen $\alpha = i$ und $\bar{\alpha} = -i$. Es gilt

$$((x - \bar{\alpha})(x - \alpha))^2 = (x^2 + 1)^2.$$

Wir verwenden den Ansatz

$$\frac{x^2 - 1}{(x^2 + 1)^2} = \frac{B_1x + C_1}{x^2 + 1} + \frac{B_2x + C_2}{(x^2 + 1)^2}.$$

Wir multiplizieren beide Seiten mit $(x^2 + 1)^2$, wobei $x \neq \pm i$ gilt. Damit erhalten wir

$$x^2 - 1 = (B_1x + C_1)(x^2 + 1) + (B_2x + C_2).$$

Für $x \rightarrow i$ folgt durch Vergleich von Real- und Imaginärteil

$$-2 = B_2i + C_2, \quad \implies B_2 = 0, C_2 = -2,$$

also

$$x^2 - 1 = (B_1x + C_1)(x^2 + 1) - 2 = B_1x^3 + B_1x + C_1x^2 + C_1 - 2.$$

Ein Koeffizientenvergleich bei x^3 ergibt $B_1 = 0$ sowie bei x^2 , dass $C_1 = 1$ gelten muss. Daraus folgt also für die restlichen Koeffizienten

$$C_1 = 1, \quad B_1 = 0.$$

Durch Einsetzen dieser Koeffizienten in den obigen Ansatz überzeugt man sich ebenfalls leicht, dass die Koeffizienten richtig bestimmt worden sind. Wir erhalten also folgende Partialbruchzerlegung vom Typ II:

$$\boxed{\frac{x^2-1}{(x^2+1)^2} = \frac{1}{x^2+1} + \frac{-2}{(x^2+1)^2}.} \quad (20)$$

Allgemein gilt folgender Satz.

Satz 10 *Jede echt-gebrochene rationale Funktion*

$$S(x) = \frac{P(x)}{Q(x)}$$

lässt sich eindeutig als Summe von Partialbrüchen vom Typ (I) und (II) darstellen.

Bemerkung 8 Nach den obigen Überlegungen treten Partialbrüche vom Typ I bei einfachen und mehrfachen reellen Nullstellen sowie Partialbrüche vom Typ II bei einfachen oder mehrfachen komplexen Nullstellen von $Q(x)$ auf.

Die gerade erzielten Ergebnisse benutzen wir jetzt zur Integration von beliebigen rationalen Funktionen der Form

$$\int f(x)dx = \int \frac{P(x)}{Q(x)}dx.$$

Dabei verwenden wir folgenden Algorithmus.

1. Schritt: Wir spalten den ganzen Teil der rationalen Funktion $f(x)$ mittels euklidischen Algorithmus ab.

Wir erhalten

$$\int f(x)dx = \int \frac{P(x)}{Q(x)}dx = \int R(x)dx + \int \frac{S(x)}{Q(x)}dx,$$

wobei der Integrand im ersten Summanden ein Polynom und der im zweiten Summand ein echter Bruch ist. Zur Integration des zweiten Summanden verwenden wir Satz 10, wonach wir nur

noch Partialbrüche vom Typ (I) und vom Typ (II) zu integrieren brauchen.

2. *Schritt:* Wir integrieren zunächst das entstandene Polynom $R(x)$ und anschließend alle auftretenden Partialbrüche. Dabei verwenden wir folgende Integrationsformeln.

Integrale nach Partialbruchzerlegung

$$\int \frac{dx}{(x-\xi)^m} = \begin{cases} -\frac{1}{m-1} \cdot \frac{1}{(x-\xi)^{m-1}} + C : & m = 2, 3, \dots \\ \ln|x-\xi| + C : & m = 1 \end{cases}$$

$$\int \frac{dx}{x^2 + ax + b} = \frac{2}{\sqrt{4b - a^2}} \arctan \frac{2x + a}{\sqrt{4b - a^2}} + C$$

$$\int \frac{dx}{(x^2 + ax + b)^m} = \frac{2x + a}{(m-1)(4b - a^2)(x^2 + ax + b)^{m-1}} + \frac{2(2m-3)}{(m-1)(4b - a^2)} \int \frac{dx}{(x^2 + ax + b)^{m-1}} + C, \quad m = 2, 3, \dots$$

$$\int \frac{\alpha x + \beta}{x^2 + ax + b} dx = \frac{\alpha}{2} \ln|x^2 + ax + b| + \left(\beta - \frac{\alpha a}{2}\right) \int \frac{dx}{x^2 + ax + b} + C$$

$$\int \frac{\alpha x + \beta}{(x^2 + ax + b)^m} dx = -\frac{\alpha}{2(m-1)(x^2 + ax + b)^{m-1}} + \left(\beta - \frac{\alpha a}{2}\right) \int \frac{dx}{(x^2 + ax + b)^m} + C, \quad m = 2, 3, \dots$$

Beispiel 9 Wir integrieren die rationalen Funktionen aus Beispiel 8.

Beispiel 8(a)

Wir betrachten

$$\int \frac{x-1}{(x+1)(x-2)} dx.$$

Es liegt ein echter Bruch vor. Bei der Integration erhalten wir also nach (14)

$$\begin{aligned} \int \frac{x-1}{(x+1)(x-2)} dx &= \frac{2}{3} \int \frac{dx}{x+1} + \frac{1}{3} \int \frac{dx}{x-2} \\ &= \frac{2}{3} \ln|x+1| + \frac{1}{3} \ln|x-2| + c. \end{aligned}$$

Beispiel 8(b)

Wir berechnen

$$\int \frac{x-1}{x^2(x-2)} dx.$$

Es liegt wieder ein echter Bruch vor. Bei der Integration erhalten wir also nach (16)

$$\begin{aligned} \int \frac{x-1}{x^2(x-2)} dx &= \frac{-1}{4} \int \frac{dx}{x} + \frac{1}{2} \int \frac{dx}{x^2} + \frac{1}{4} \int \frac{dx}{x-2} \\ &= \frac{-1}{4} \ln|x| - \frac{1}{2x} + \frac{1}{4} \ln|x-2| + c. \end{aligned}$$

Beispiel 8(c)

Wir berechnen

$$\int \frac{4}{x^3+4x} dx.$$

Es liegt wieder ein echter Bruch vor. Bei der Integration erhalten wir also nach (18)

$$\begin{aligned} \int \frac{4}{x^3+4x} dx &= \int \frac{dx}{x} + \int \frac{-x}{x^2+4} dx \\ &= \ln|x| - \frac{1}{2} \ln(x^2+4) + c. \end{aligned}$$

Beispiel 8(d)

Wir betrachten das Integral

$$\int \frac{x^2-1}{(x^2+1)^2} dx.$$

Es liegt wieder ein echter Bruch vor. Verwenden wir (20), so erhalten wir

$$\begin{aligned} \int \frac{x^2-1}{(x^2+1)^2} dx &= -2 \int \frac{dx}{(x^2+1)^2} + \int \frac{dx}{x^2+1} \\ &= -2I_1 + \arctan x + c_1. \end{aligned}$$

Für das Integral I_1 gilt mit $m=2$, $a=0$, $b=1$

$$\begin{aligned} I_1 &= \int \frac{dx}{(x^2+1)^2} \\ &= \frac{1}{2} \cdot \frac{x}{x^2+1} + \frac{1}{2} \int \frac{dx}{x^2+1} \\ &= \frac{1}{2} \cdot \frac{x}{x^2+1} + \frac{1}{2} \arctan x + c_2. \end{aligned}$$

Somit folgt

$$\begin{aligned} \int \frac{x^2-1}{(x^2+1)^2} dx &= -\frac{2x}{2(x^2+1)} - \arctan x + \arctan x + c \\ &= -\frac{x}{x^2+1} + c. \end{aligned}$$

Die Probe ergibt nach der Quotientenregel

$$\left(\frac{-x}{x^2+1}\right)' = \frac{x^2-1}{(x^2+1)^2}.$$

Durch geeignete Substitutionen kann man auch rationale Funktionen

- $R(e^x)$ bezüglich e^x ,
- $R(\sin x, \cos x)$, $R(\sin x, \cos x, \tan x, \cot x)$ sowie
- $R(\sin x, \cos x, \sin 2x, \cos 2x, \dots, \sin nx, \cos nx)$

berechnen. Entsprechende Formeln findet man in Nachschlagewerken.

3 Taylorreihen und Potenzreihen

3.1 Taylorreihen

Gegeben sei eine Funktion f , die im Intervall $[a, b]$ n -mal stetig differenzierbar ist. Weiterhin sei $x_0 \in (a, b)$ ein fixierter Punkt. Falls $x \in U(x_0) \subset (a, b)$ gilt, so bezeichnet man mit

$$P_n(x, x_0) = f(x_0) + f'(x_0)(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

das n -te *Taylorpolynom* der Funktion f , $n \in \mathbb{N}_0$, und mit

$$R_n(x, x_0) = f(x) - P_n(x, x_0)$$

das zugehörige *Restglied*.

Vereinbarung: In diesem Abschnitt verwenden wir (ähnlich wie bei binomischen Formeln) die Definition

$$x^0 = 1, \quad \text{für alle } x \in \mathbb{R}.$$

Damit lassen sich viele der folgenden Formeln einfacher aufschreiben. Das hat also nichts mit dem unbestimmten Ausdruck

$$0^0$$

zu tun!

In Zukunft verwenden wir die Schreibweise $P_n(x)$ bzw. $R_n(x)$, falls der Punkt x_0 fest gewählt ist und keine Verwechslungen auftreten können. Falls also das n -te Taylorpolynom $P_n(x)$ die gegebene Funktion f approximieren soll, so entspricht $R_n(x)$ dem *Approximationsfehler*. Das folgende Resultat gibt eine Darstellung des Restgliedes.

Satz 1 Die Funktion f besitze in $U(x_0)$ stetige Ableitungen bis zur Ordnung $n + 1$. Weiterhin sei $x \in U(x_0)$ mit $x < x_0$. Dann existiert ein $\xi \in (x, x_0)$ mit der Eigenschaft

$$R_n(x) = f^{(n+1)}(\xi) \frac{(x - x_0)^{n+1}}{(n + 1)!} \quad (1)$$

(*Restglied von Lagrange*).

Bemerkung 1 (a) Aus Gleichung (1) folgt

$$\lim_{x \rightarrow x_0} \frac{R_n(x)}{(x - x_0)^n} = 0, \quad (2)$$

d.h. für alle $\varepsilon > 0$ existiert ein $\delta(\varepsilon) > 0$ mit

$$|R_n(x)| < \varepsilon |x - x_0|^n \quad \text{falls} \quad |x - x_0| < \delta(\varepsilon).$$

Man sagt dann, dass das Polynom $P_n(x)$ die gegebene Funktion $f(x)$ mit der *Ordnung* n in $U(x_0)$ approximiert.

(b) Der Punkt $x = x_0$ gibt an, **wo** die Funktion in eine Taylorreihe entwickelt wird.

(c) Man verwendet Taylorpolynome für die näherungsweise Berechnung von Funktionswerten $f(x)$ mit einer vorgegebenen Genauigkeit (Tabellen, Taschenrechner mit bestimmter Genauigkeit). Weiterhin besteht der Vorteil darin, dass sich die neuen Glieder *rekursiv* aus den alten berechnen lassen, was für die Programmierung sehr günstig ist.

(d) Für den Spezialfall $n = 0$ erhalten wir

$$R_0(x) = f(x) - f(x_0) = f'(\xi)(x - x_0),$$

d.h. den Mittelwertsatz der Differentialrechnung, siehe Vorkurs, Abschnitt 4. Differenzierbare Funktionen.

(e) Für $n = 1$ erhalten wir aus Satz 1 die bereits bekannte *Approximationseigenschaft der Tangente* (siehe Vorkurs, Abschnitt 4. Differenzierbare Funktionen), d.h. es gilt

$$t(x) = P_1(x) = f(x_0) + f'(x_0)(x - x_0).$$

(f) Falls für $x \in U(x_0)$ die Beziehung $x > x_0$ gilt, so existiert analog ein $\xi \in (x_0, x)$ mit der obigen Eigenschaft (1).

(g) Nach Satz 1 erhalten wir somit folgende Darstellung für die Funktion f .

$$f(x) = P_n(x) + R_n(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}. \quad (3)$$

Beispiel 1 (a) Man berechne für $x \in [-1, 1]$ die Funktion $\sin x$ mit einer Genauigkeit von 10^{-2} , wobei $x_0 = 0$ verwendet wird!

Es gilt zunächst wegen

$$(\sin x)' = \cos x, \quad (\sin x)'' = -\sin x, \quad (\sin x)''' = -\cos x, \quad \sin^{(4)} x = \sin x,$$

dass

$$\sin^{(2k)} x|_{x=0} = 0, \quad \sin^{(2k+1)} x|_{x=0} = \begin{cases} +1 & : k = 0, 2, 4, \dots \\ -1 & : k = 1, 3, 5, \dots \end{cases}.$$

Die Struktur ist also

$$(-1)^j \frac{x^{2j+1}}{(2j+1)!}.$$

Somit erhalten wir

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} \mp \dots + \frac{\sin^{(n+1)}(\xi)}{(n+1)!} x^{n+1},$$

und damit

$$|R_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!} \leq \frac{1}{(n+1)!} < 10^{-2},$$

falls $n+1 = 5$ gilt. Die Näherungsformel ist also

$$\sin x = x - \frac{x^3}{3!} + R_4(x), \quad |R_4(x)| < 10^{-2},$$

d.h.

$$\sin x \approx x - \frac{x^3}{3!}.$$

(b) Es sei f beliebig oft differenzierbar. Dann gilt

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k,$$

falls $R_n(x) \rightarrow 0$ für $n \rightarrow \infty$ erfüllt ist. Somit ist für diese *Potenzreihen*-Darstellung ausreichend, dass die Bedingung

$$\frac{|f^{(n+1)}(\xi)|}{(n+1)!} |x - x_0|^{n+1} \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

für alle $\xi \in [x_0, x]$ bzw. alle $\xi \in [x, x_0]$ erfüllt ist. Ist das der Fall, so erhält man die *Taylorreihe* für die unendlich oft differenzierbare Funktion f . Insbesondere gilt für $x_0 = 0$

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k.$$

(c) Wir zeigen, dass die Taylorreihe der Funktion $\sin x$ für alle reellen x durch die Reihe

$$\sin x = \sum_{j=0}^{\infty} (-1)^j \frac{x^{2j+1}}{(2j+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} \mp \dots$$

gegeben ist. Wir setzen $x_0 = 0$ und beweisen, dass für eine beliebige reelle Zahl x die obige Bedingung erfüllt ist. Es gilt zunächst

$$|R_n(x)| = \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} \right| |x|^{n+1} \leq \frac{|x|^{n+1}}{(n+1)!}.$$

Definieren wir $S_n = \frac{|x|^{n+1}}{(n+1)!}$, so erhalten wir

$$\frac{S_{n+1}}{S_n} = \frac{\frac{|x|^{n+2}}{(n+2)!}}{\frac{|x|^{n+1}}{(n+1)!}} = \frac{|x|}{n+2} \leq \frac{1}{2},$$

falls $n/2 + 1 \geq |x|$ gilt. Somit haben wir $S_n(x) \rightarrow 0$ für $n \rightarrow \infty$ für jede beliebige reelle Zahl x . Daraus folgt aber die Bedingung an das Restglied $R_n(x)$.

Analog zeigt man, dass für reelle x

$$e^x = \sum_{j=0}^{\infty} \frac{x^j}{j!},$$

$$\cos x = \sum_{j=0}^{\infty} (-1)^j \frac{x^{2j}}{(2j)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} \mp \dots$$

gilt. Insbesondere folgt daraus

$$e = \sum_{j=0}^{\infty} \frac{1}{j!}.$$

Ohne die obige Vereinbarung müssten wir z.B.

$$e^x = 1 + \sum_{k=1}^{\infty} \frac{x^k}{k!}$$

schreiben!

Weiterhin erhalten wir

$$\ln(1+x) = \sum_{j=1}^{\infty} (-1)^{j+1} \frac{x^j}{j} = x - \frac{x^2}{2} + \frac{x^3}{3} \mp \dots,$$

falls $|x| < 1$ ist. Diese Reihe konvergiert sogar für $x = 1$. Man erhält in diesem Falle als Wert für die alternierende Reihe (siehe Vorkurs, Abschnitt 2. Folgen und Reihen)

$$\ln 2 = \sum_{j=1}^{\infty} (-1)^{j+1} \frac{1}{j}.$$

Für $x = -1$ liegt bestimmte Divergenz vor, siehe harmonische Reihe in Vorkurs, Abschnitt 2. Folgen und Reihen. Das entspricht $\ln(1+x) \rightarrow -\infty$, falls $x \downarrow -1$.

Der Vorteil dieser Taylorreihen besteht darin, dass zur Berechnung des Wertes $f(x)$ nur einfache algebraische Operationen (Addition/Subtraktion, Multiplikation/Division) benötigt werden.

Somit gelten für $x \approx 0$ z.B. als Näherungen

$$e^x \approx 1 + x, \quad \sin x \approx x, \quad \cos x \approx 1 - \frac{x^2}{2}, \quad \ln(1+x) \approx x,$$

die in vielen Fällen für die praktische Anwendungen schon ausreichend sind. Die Taylorreihen von weiteren elementaren Funktionen findet man in Formelsammlungen.

3.2 Potenzreihen

Wir wollen uns jetzt etwas ausführlicher mit den Potenzreihen beschäftigen, die u.a. bei der Taylorentwicklung von unendlich oft differenzierbaren Funktionen entstehen. Wir wollen jedoch

darauf hinweisen, dass sich nicht jede unendlich oft differenzierbare Funktion in eine Potenzreihe entwickeln lässt (nicht-elementare Gegenbeispiele).

Es seien x_0 ein fester und x ein beliebiger Punkt des \mathbb{R} , und a_j , $j \in \mathbb{N}_0$, seien reelle Zahlen. Dann nennt man die Funktion

$$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots = \sum_{j=0}^{\infty} a_j(x - x_0)^j$$

Potenzreihe bezüglich des Punktes x_0 . Auch hier verwenden wir die Vereinbarung

$$(x - x_0)^0 = 1.$$

Wir sind daran interessiert, für welche $x \in \mathbb{R}$ diese Potenzreihe konvergiert. Zunächst ist offensichtlich, dass $|x - x_0|$ und $|a_j|$ zueinander indirekt proportional sind. Wenden wir das Wurzelkriterium für Reihen (siehe Vorkurs, Abschnitt 2. Folgen und Reihen) an, so erhalten wir wegen

$$\sqrt[j]{|a_j||x - x_0|^j} = \sqrt[j]{|a_j|}|x - x_0|,$$

dass die obige Reihe für alle $x \in \mathbb{R}$ (absolut) konvergiert, falls ein $q < 1$ existiert mit

$$\sqrt[j]{|a_j||x - x_0|} \leq q \iff |x - x_0| \leq \frac{q}{\sqrt[j]{|a_j|}} \text{ für alle } j. \quad (4)$$

Andererseits liegt offensichtlich Divergenz vor für alle reellen x mit

$$\sqrt[j]{|a_j||x - x_0|} \geq Q \iff |x - x_0| \geq \frac{Q}{\sqrt[j]{|a_j|}} \text{ für alle } j, \quad (5)$$

falls $Q > 1$ ist.

Analoge Resultate kann man mittels des Quotientenkriteriums, siehe ebenfalls Vorkurs, Abschnitt 2. Folgen und Reihen, herleiten. Dann muss im Falle der (absoluten) Konvergenz

$$|x - x_0| \leq \frac{q}{\frac{|a_{j+1}|}{|a_j|}} \text{ für alle } j \quad (6)$$

gelten, wobei $q < 1$ ist. Im Falle der Divergenz erhalten wir

$$|x - x_0| \geq \frac{Q}{\frac{|a_{j+1}|}{|a_j|}} \text{ für alle } j \quad (7)$$

falls $Q > 1$ ist.

Nach den früheren Bemerkungen reicht es wieder aus, falls ein j_0 existiert, so dass (4) – (7) für alle $j \geq j_0$ richtig ist.

Definition 1 Die größte Zahl $R \geq 0$ mit der Eigenschaft, dass für $|x - x_0| < R$ die Potenzreihe (absolut) konvergiert und für $|x - x_0| > R$ Divergenz vorliegt, bezeichnet man als Konvergenzradius der Potenzreihe.

Unter Verwendung von (4) – (7) erhalten wir sofort folgendes Ergebnis.

Satz 2 Der Konvergenzradius R ist durch

$$\frac{1}{R} = \lim_{j \rightarrow \infty} \sqrt[j]{|a_j|} \quad \text{oder} \quad \frac{1}{R} = \lim_{j \rightarrow \infty} \frac{|a_{j+1}|}{|a_j|}$$

gegeben, falls dieser (endliche oder unendliche) Grenzwert existiert.

Dabei verwenden wir folgende Vereinbarungen:

$$R = 0, \quad \text{falls} \quad \lim_{j \rightarrow \infty} \sqrt[j]{|a_j|} = \infty$$

und

$$R = \infty, \quad \text{falls} \quad \lim_{j \rightarrow \infty} \sqrt[j]{|a_j|} = 0.$$

(analog für das Quotientenkriterium).

Für $|x - x_0| = R$ ist das Verhalten unbestimmt. Oft gehört $|x - x_0| = R$ nicht mehr zum Konvergenzbereich. Es muss also das Verhalten der Reihe für $x - x_0 = R$ und $x - x_0 = -R$ extra untersucht werden. Wir verweisen auf die nachfolgenden Beispiele.

Beispiel 2 (a) Wir betrachten (mit der üblichen Vereinbarung)

$$f(x) = \sum_{j=0}^{\infty} x^j.$$

Wir haben $x_0 = 0$. Wegen $a_j \equiv 1$ und $\sqrt[j]{|a_j|} = 1$ erhalten wir

$$\lim_{j \rightarrow \infty} \sqrt[j]{|a_j|} = \lim_{j \rightarrow \infty} 1 = 1$$

und somit $R = 1$. Damit konvergiert die Potenzreihe für alle reellen Zahlen x mit $|x| < 1$. Es gilt in diesem Falle (analog zur geometrischen Reihe, siehe Vorkurs, Abschnitt 2)

$$f(x) = \frac{1}{1-x}.$$

Setzen wir $x = 1$, so erhalten wir $\sum_{j=0}^{\infty} 1^j = \infty$ (bestimmte Divergenz). Falls $x = -1$ ist, so gilt für die Partialsummen $S_{2N+1} = 0$ und $S_{2N} = 1$, $N \in \mathbb{N}_0$. Es liegt also ebenfalls keine Konvergenz

vor (Reihe ist unbestimmt divergent).

(b) Wir untersuchen

$$f(x) = \sum_{j=1}^{\infty} j^j x^j.$$

Somit ist $a_j = j^j$. Wegen $\sqrt[j]{|a_j|} = j$ folgt

$$\lim_{j \rightarrow \infty} \sqrt[j]{|a_j|} = \lim_{j \rightarrow \infty} j = \infty.$$

Wir erhalten damit entsprechend der Vereinbarung $R = 0$, d.h. die Reihe konvergiert für **kein** reelles $x \neq 0$.

(c) Für

$$f(x) = \sum_{j=0}^{\infty} \frac{x^j}{j!}$$

ergibt sich wegen $a_j = \frac{1}{j!}$ der Grenzwert $\lim_{j \rightarrow \infty} \frac{|a_{j+1}|}{|a_j|} = \lim_{j \rightarrow \infty} \frac{j!}{(j+1)!} = \lim_{j \rightarrow \infty} \frac{1}{j+1} = 0$. Somit erhalten wir entsprechend der Vereinbarung $R = \infty$, d.h. die Reihe konvergiert für **alle** reellen Zahlen x . Es gilt $f(x) = e^x$.

Für die Addition, Multiplikation, Differentiation und Integration von Potenzreihen gelten folgende Eigenschaften.

Satz 3 *Es seien $\sum_{j=0}^{\infty} a_j(x-x_0)^j$ und $\sum_{j=0}^{\infty} b_j(x-x_0)^j$ zwei Potenzreihen mit Konvergenzradius R_a bzw. R_b .*

(a) *Sind λ und μ reelle Zahlen, und ist R der Konvergenzradius von*

$$\lambda \sum_{j=0}^{\infty} a_j(x-x_0)^j + \mu \sum_{j=0}^{\infty} b_j(x-x_0)^j = \sum_{j=0}^{\infty} (\lambda a_j + \mu b_j)(x-x_0)^j,$$

so gilt $R \geq \min\{R_a, R_b\}$.

(b) *Ist R der Konvergenzradius der Reihe*

$$\left(\sum_{j=0}^{\infty} a_j(x-x_0)^j \right) \cdot \left(\sum_{j=0}^{\infty} b_j(x-x_0)^j \right) = \sum_{j=0}^{\infty} (a_0 b_j + a_1 b_{j-1} + \dots + a_j b_0)(x-x_0)^j,$$

so gilt $R \geq \min(R_a, R_b)$.

(c) *Die Potenzreihe $f(x) = \sum_{j=0}^{\infty} a_j(x-x_0)^j$ ist für $|x-x_0| < R_a$ beliebig oft differenzierbar. Es gilt*

$$f^{(k)}(x) = \sum_{j=k}^{\infty} j(j-1)(j-2) \cdots (j-k+1) a_j (x-x_0)^{j-k}, \quad k = 1, 2, \dots$$

(gliedweise Differentiation), wobei sämtliche Reihen $f^{(k)}(x)$ einen Konvergenzradius $R_k \geq R_a$ haben. Insbesondere gilt

$$f^{(k)}(x_0) = a_k k!$$

(d) Liegt das Intervall $[a, b]$ innerhalb des Konvergenzradius R_a , so ist die Potenzreihe $f(x) = \sum_{j=0}^{\infty} a_j (x - x_0)^j$ im Intervall $[a, b]$ integrierbar. Es gilt

$$\int_a^b f(x) dx = \sum_{j=0}^{\infty} \int_a^b a_j (x - x_0)^j dx,$$

(gliedweise Integration)

Beispiel 3 Unter Verwendung der Potenzreihe der Sinusfunktion und der gliedweisen Integration betrachten wir das Integral der Funktion $\frac{\sin t}{t}$ (Integralsinus). Wir erhalten

$$\begin{aligned} \text{Si}(x) &= \int_0^x \frac{\sin t}{t} dt = \int_0^x \left(1 - \frac{t^2}{3!} + \frac{t^4}{5!} - \frac{t^6}{7!} \pm \dots\right) dt \\ &= x - \frac{x^3}{3 \cdot 3!} + \frac{x^5}{5 \cdot 5!} - \frac{x^7}{7 \cdot 7!} \pm \dots \end{aligned}$$

Es gilt nun folgender Zusammenhang zwischen Taylorreihe und Potenzreihe einer unendlich oft differenzierbaren Funktion (*Identitätssatz*).

Satz 4 Falls sich eine beliebig oft differenzierbare Funktion als Potenzreihe

$$f(x) = \sum_{j=0}^{\infty} a_j (x - x_0)^j$$

darstellen lässt, so gilt für die Koeffizienten

$$a_j = \frac{f^{(j)}(x_0)}{j!}.$$

Wir werden später Potenzreihen bei der Lösung von speziellen Differentialgleichungen verwenden (*Potenzreihenansatz*).

4 Lineare Algebra

4.1 Lineare Gleichungssysteme

Als Einführung betrachten wir folgende 2 Gleichungen mit 2 Unbekannten.

$$\begin{aligned}x_1 + x_2 &= 4 \\2x_1 + 3x_2 &= 9.\end{aligned}$$

In diesem Fall existiert genau ein Lösungspaar (x_1, x_2) , das beiden Gleichungen genügt.

Durch äquivalente Umformungen (Multiplikation der 1. Gleichung mit 2 und anschließende Subtraktion von Gleichung 2) erhält man

$$\begin{aligned}2x_1 + 2x_2 &= 8 \\ & - \\2x_1 + 3x_2 &= 9.\end{aligned}$$

$x_2 = 1$ und damit aus $2x_1 + 2 = 8$, dass $x_1 = 3$ gilt. Somit ist die *eindeutig bestimmte Lösung* gegeben durch $(x_1, x_2) = (3, 1)$. Wir betrachten jetzt das lineare Gleichungssystem

$$\begin{aligned}x_1 + 3x_2 &= 3 \\x_1 + 3x_2 &= 6.\end{aligned}$$

Dieses Gleichungssystem ist *nicht lösbar*, da $3 \neq 6$ gilt.

Abschließend betrachten wir das lineare Gleichungssystem

$$\begin{aligned}x_1 + x_2 &= 1 \\2x_1 + 2x_2 &= 2.\end{aligned}$$

In diesem Falle sind beide Gleichungen äquivalent. Ist also $x_1 = t$ ein beliebiger reeller Parameter, so erhalten wir $x_2 = 1 - x_1 = 1 - t$. Es existieren also *unendlich viele Lösungen* (x_1, x_2) , die durch $(t, 1 - t)$, $t \in \mathbb{R}$, gegeben sind.

Insgesamt erhalten wir, dass ein lineares Gleichungssystem entweder *keine, genau eine* oder *unendlich viele* Lösung(en) besitzen kann.

Durch den sogenannten *Gaußschen Algorithmus* kann man das Lösen von linearen Gleichungssystemen (GLS) formalisieren. Wir beschreiben das an einem konkreten Beispiel.

Beispiel 1 Gegeben sei das GLS, bestehend aus 2 Gleichungen für 3 Unbekannte

$$\begin{aligned}2x_1 - x_2 + x_3 &= 0 \\2x_1 - x_2 - x_3 &= 1.\end{aligned}$$

Es gilt:

Die *elementaren Umformungen*

- Multiplikation einer Gleichung mit einer reellen Zahl $a \neq 0$,
- Addition (bzw. Subtraktion) einer Gleichung zu (bzw. von) einer anderen Gleichung,
- Vertauschen zweier Gleichungen

verändern die Lösungsmenge eines GLS **nicht**. Wir schreiben jetzt das GLS formal als

$$\begin{array}{cccc}2 & -1 & 1 & 0 \\2 & -1 & -1 & 1\end{array}$$

Dabei bezeichnet \searrow die sogenannte *Hauptdiagonale*. Nun verwenden wir diese elementaren Umformungen, um *Trapezgestalt* zu erreichen, d.h. unterhalb der Hauptdiagonale sollen nur noch *Nullen* stehen. In unserem konkreten Beispiel erhalten wir aus

$$\begin{array}{cccc} 2 & -1 & 1 & 0 \\ \boxed{2} & -1 & -1 & 1 \\ 2 & -1 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array}$$

d.h. es gilt

$$-2x_3 = 1 \implies x_3 = -\frac{1}{2}.$$

Damit bleibt nach Einsetzen dieses Wertes für x_3 in die erste Zeile eine Gleichung für die Unbekannten x_1 und x_2 , die durch

$$2x_1 - x_2 + x_3 = 0 \implies 2x_1 - x_2 = -x_3 = \frac{1}{2}$$

gegeben ist. Wählen wir für $x_2 = t$, $t \in \mathbb{R}$ freier Parameter, so folgt schließlich

$$x_1 = \frac{1}{4} + \frac{t}{2}.$$

Die Lösungsmenge unseres GLS ist also durch

$$\{x = (x_1, x_2, x_3) = (\frac{t}{2} + \frac{1}{4}, t, -\frac{1}{2}) : t \in \mathbb{R}\}$$

gegeben. Ausgehend von diesem Beispiel geben wir jetzt das Schema für den *allgemeinen Lösungsalgorithmus* an.

Es sei das GLS mit m Gleichungen und n Unbekannten

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \dots\dots\dots & = & \dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array} \tag{1}$$

gegeben. Dabei gelte **immer** $m \leq n$.

Unsere Strategie besteht darin, dass wir durch elementare Umformungen eine Trapezgestalt erreichen wollen. Dabei verwenden wir wieder die formale Schreibweise.

O.B.d.A. nehmen wir an, dass $a_{11} = 1$ gilt. Sonst dividieren wir die erste Zeile durch a_{11} , falls

$a_{11} \neq 0$ gilt. Im anderen Fall führen wir einen Zeilentausch und gegebenenfalls eine Division durch. Wie oben verwenden wir das Schema

$$\begin{array}{cccccc} \boxed{1} & a_{12} & \cdots & a_{1n} & b_1 & \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m & \end{array} \quad (2)$$

Nun addieren wir das $(-a_{i1})$ -fache der 1. Zeile zur i -ten Zeile für $i = 2, \dots, m$, um Nullen in der ersten Spalte unterhalb der Hauptdiagonale zu haben.

$$\begin{array}{cccccc} 1 & a_{12} & \cdots & a_{1n} & b_1 & \\ 0 & a'_{22} & \cdots & a'_{2n} & b'_2 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & a'_{m2} & \cdots & a'_{mn} & b'_m & \end{array} \quad (3)$$

Jetzt wenden wir auf

$$\begin{array}{cccc} a'_{22} & \cdots & a'_{2n} & b'_2 \\ \cdots & \cdots & \cdots & \cdots \\ a'_{m2} & \cdots & a'_{mn} & b'_m \end{array} \quad (4)$$

die gleiche Prozedur (2) wie vorher auf das Ausgangsschema an (o.B.d.A. $a'_{22} = 1$.) Dabei lassen wir triviale Zeilen (d.h. Zeilen mit nur Nullen) weg, da sie keine Informationen geben.

Falls bei der Prozedur eine Zeile der Form

$$0 \quad \cdots \quad 0 \quad b$$

mit $b \neq 0$ entsteht, so beenden wir das Verfahren; in diesem Fall ist das GLS **nicht lösbar**, da in diesem Fall

$$0x_1 + \cdots + 0x_n = b \neq 0$$

gelten müsste. Das GLS besitzt somit *keine Lösung*.

Ansonsten erreichen wir nach endlich vielen Schritten *Trapezgestalt*, d.h. wir haben folgende Situation

$$\begin{array}{ccccccc} 1 & \tilde{a}_{12} & \cdots & \tilde{a}_{1l} & \cdots & \tilde{a}_{1n} & \tilde{b}_1 \\ 0 & \tilde{a}_{22} & \cdots & \tilde{a}_{2l} & \cdots & \tilde{a}_{2n} & \tilde{b}_2 \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \tilde{a}_{ll} & \cdots & \tilde{a}_{ln} & \tilde{b}_l \end{array} \quad (5)$$

Es können nun die folgenden zwei Fälle auftreten:

Fall (a): Auf der Hauptdiagonale stehen **nur** Einsen.

Fall (b): Auf der Hauptdiagonale stehen **auch** Nullen. Dann müssen **alle** Zahlen in den entsprechenden Spalten unterhalb der Hauptdiagonale nach dem Algorithmus ebenfalls Nullen sein (sonst ist noch Zeilentausch möglich).

Es gelte $l = n$ (Fall I):

Falls (a) erfüllt ist, so können wir *sukzessiv* die Lösung des GLS berechnen: Wir setzen

$$x_n = \tilde{b}_n \implies x_{n-1} \cdots \implies x_1.$$

Somit erhalten wir eine *eindeutig bestimmte Lösung*.

Falls (b) gilt, so setzen wir $x_k = t_k \in \mathbb{R}$ (freie Parameter) für $\tilde{a}_{kk} = 0$ und verfahren anschließend wie im Fall (a). Wir erhalten also *unendlich viele Lösungen* des GLS.

Es gelte $l < n$ (Fall II):

Dann setzen wir $x_i = t_i$ für $i = l + 1, \dots, n$ (freie Parameter). Anschließend verfahren wir wie im Fall I, um sukzessiv x_i , $i = 1, \dots, l$, zu berechnen. Wir erhalten also ebenfalls *unendlich viele Lösungen* des GLS.

Abschließend veranschaulichen wir uns diesen Algorithmus noch einmal an einem Beispiel.

Beispiel 2 Wir betrachten das GLS bestehend aus 4 Gleichungen für 5 Unbekannte.

$$\begin{aligned} 2x_1 - x_2 + x_3 - x_4 + x_5 &= 0 \\ -2x_1 + x_2 - 2x_3 - x_4 + 2x_5 &= -1 \\ 4x_1 - 2x_2 + x_3 - x_4 - x_5 &= 2 \\ 2x_1 - x_2 - x_3 - 2x_4 + x_5 &= 1 \end{aligned} \tag{6}$$

Damit erhalten wir

$$\begin{array}{cccccc} \boxed{2} & -1 & 1 & -1 & 1 & 0 \\ -2 & 1 & -2 & -1 & 2 & -1 \\ 4 & -2 & 1 & -1 & -1 & 2 \\ 2 & -1 & -1 & -2 & 1 & 1 \end{array}$$

Wir multiplizieren jetzt die 1. Zeile mit $1/2$. Damit gilt

$$\begin{array}{cccccc} 1 & -1/2 & 1/2 & -1/2 & 1/2 & 0 \\ -2 & 1 & -2 & -1 & 2 & -1 \\ 4 & -2 & 1 & -1 & -1 & 2 \\ 2 & -1 & -1 & -2 & 1 & 1 \end{array}$$

Nun addieren wir das 2-fache der 1. Zeile zur 2., das (-4)-fache zur 3. und das (-2)-fache zur letzten Zeile. Damit erhalten wir

$$\begin{array}{cccccc} 1 & -1/2 & 1/2 & -1/2 & 1/2 & 0 \\ 0 & 0 & -1 & -2 & 3 & -1 \\ 0 & 0 & \boxed{-1} & 1 & -3 & 2 \\ 0 & 0 & -2 & -1 & 0 & 1 \end{array}$$

Wir multiplizieren jetzt die 3. Zeile mit (-1). Es folgt

$$\begin{array}{cccccc} 1 & -1/2 & 1/2 & -1/2 & 1/2 & 0 \\ 0 & 0 & -1 & -2 & 3 & -1 \\ 0 & 0 & \boxed{1} & -1 & 3 & -2 \\ 0 & 0 & -2 & -1 & 0 & 1 \end{array}$$

Nun addieren wir das 2-fache der 3. Zeile zur letzten Zeile.

$$\begin{array}{cccccc} 1 & -1/2 & 1/2 & -1/2 & 1/2 & 0 \\ 0 & 0 & -1 & -2 & 3 & -1 \\ 0 & 0 & 1 & -1 & 3 & -2 \\ 0 & 0 & 0 & \boxed{-3} & 6 & -3 \end{array}$$

Damit haben wir die Trapezform erreicht. Abschließend multiplizieren wir die 4. Zeile noch mit (-1/3).

$$\begin{array}{cccccc} 1 & -1/2 & 1/2 & -1/2 & 1/2 & 0 \\ 0 & \boxed{0} & -1 & -2 & 3 & -1 \\ 0 & 0 & 1 & -1 & 3 & -2 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}$$

Wir wählen jetzt $x_5 = t$, wobei t ein freier reeller Parameter ist. Daraus folgt aus der 4. Zeile wegen $x_4 - 2x_5 = 1$, dass $x_4 = 2t + 1$ ist und damit aus der 3. Zeile wegen $x_3 - x_4 + 3x_5 = -2$

$$\begin{aligned} x_3 &= -2 + x_4 - 3x_5 \\ &= -2 + 2t + 1 - 3t = -1 - t. \end{aligned}$$

Da der Koeffizient auf der Hauptdiagonale in der 2. Zeile 0 ist, wählen wir $x_2 = u$, wobei u ein freier reeller Parameter ist. Damit folgt aus der 1. Zeile

$$\begin{aligned} x_1 - \frac{1}{2}u &= -\frac{1}{2}x_3 + \frac{1}{2}x_4 - \frac{1}{2}x_5 \\ &= \frac{1}{2} + \frac{1}{2}t + t + \frac{1}{2} - \frac{1}{2}t = 1 + t \end{aligned}$$

und damit $x_1 = \frac{1}{2}u + 1 + t$. Somit ist die Lösungsmenge des GLS (6) gegeben durch

$$\{(x_1, x_2, x_3, x_4, x_5) = (1 + \frac{u}{2} + t, u, -1 - t, 1 + 2t, t) : u, t \in \mathbb{R}\}.$$

Somit besitzt die Lösung zwei freie Parameter. Wir können also die Lösung schreiben als

$$\underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}}_x = \begin{pmatrix} 1 + \frac{u}{2} + t \\ u \\ -1 - t \\ 1 + 2t \\ t \end{pmatrix} = \underbrace{\begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}}_{\tilde{x}_0} + u \cdot \underbrace{\begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\tilde{x}_1} + t \cdot \underbrace{\begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \\ 1 \end{pmatrix}}_{\tilde{x}_2}, \quad (7)$$

$u, t \in \mathbb{R}$ beliebig, oder

$$x = \tilde{x}_0 + u\tilde{x}_1 + t\tilde{x}_2, \quad t, u \in \mathbb{R}.$$

4.2 Matrizen

Wir verwenden jetzt die obige symbolische Schreibweise und führen den Begriff der *Matrix* ein.

Definition 1 (a) *Eine Matrix ist ein rechteckiges Schema von reellen Zahlen, die Elemente der Matrix heißen. Eine (m, n) -Matrix besteht aus m Zeilen, n Spalten und somit $m \cdot n$ reellen Zahlen.*

(b) *Eine Matrix heißt quadratisch, falls $m = n$ gilt.*

(c) *Eine $(m, 1)$ -Matrix nennt man auch Spaltenvektor und eine $(1, n)$ -Matrix Zeilenvektor.*

(d) *Falls $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_n) \in \mathbb{R}^n$, so bezeichnet*

$$\langle x, y \rangle = \langle y, x \rangle = \sum_{i=1}^n x_i \cdot y_i \in \mathbb{R} \quad (8)$$

das Skalarprodukt von x und y .

Im folgenden verwenden wir als Schreibweise für eine (m, n) -Matrix A :

$$A = (a_{ij}) = (a_{ij})_{(m,n)} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}. \quad (9)$$

Dabei bezeichnet a_{ij} das allgemeine Element in der i -ten Zeile und j -ten Spalte, wobei für die Indizes $1 \leq i \leq m, 1 \leq j \leq n$ gilt. Diese Matrix besitzt also m Zeilenvektoren $a_i \in \mathbb{R}^n$

$$a_i = (a_{i1}, \dots, a_{in}), \quad i = 1, \dots, m,$$

und n Spaltenvektoren $a^j \in \mathbb{R}^m$

$$a^j = \begin{pmatrix} a_{1j} \\ \dots \\ a_{mj} \end{pmatrix}, \quad j = 1, \dots, n.$$

Weiterhin sind zwei (m, n) -Matrizen A und B gleich ($A = B$), falls

$$a_{ij} = b_{ij}, \quad i = 1, \dots, m, \quad j = 1, \dots, n$$

gilt.

Für $x \in \mathbb{R}^n$ verwenden wir hierbei auch die Schreibweise als Spaltenvektor

$$x = (x_1, \dots, x_n), \quad x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}.$$

Falls $x \in \mathbb{R}^n$ und A die (m, n) -Matrix aus (9) ist, so definieren wir $Ax \in \mathbb{R}^m$ durch

$$Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}$$

Wir erhalten also für die i -te Zeile

$$(Ax)_i = \sum_{k=1}^n a_{ik}x_k = \langle a_i, x \rangle, \quad i = 1, \dots, m.$$

Damit ergibt sich für das GLS (1) folgende Matrixschreibweise mittels der (m, n) -Matrix A

$$A = (a_{ij}) : \quad A : x \in \mathbb{R}^n \rightarrow Ax \in \mathbb{R}^m, \quad Ax = b,$$

wobei

$$b = \begin{pmatrix} b_1 \\ \dots \\ b_m \end{pmatrix}.$$

Ein GLS heißt *homogen*, falls die rechte Seite b der Nullvektor ist, d.h. $b_i = 0$ für $i = 1, \dots, m$.

Ansonsten sprechen wir von einem *inhomogenen* GLS.

Das Skalarprodukt in der Definition 1 besitzt folgende Eigenschaften:

- (Symmetrie) $\langle x, y \rangle = \langle y, x \rangle$,
- (Linearität) $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle, \quad \lambda, \mu \in \mathbb{R}$,

- (Positivität) $\langle x, x \rangle \geq 0$ und $\langle x, x \rangle = 0$ genau dann, wenn $x = 0$ gilt.

Verwendet man die Linearität des Skalarproduktes (oder die Linearität des GLS), so folgt

$$\begin{aligned} \left(A(\lambda x + \mu y) \right)_i &= \langle a_i, \lambda x \rangle + \langle a_i, \mu y \rangle \\ &= \lambda \langle a_i, x \rangle + \mu \langle a_i, y \rangle = \lambda (Ax)_i + \mu (Ay)_i \end{aligned}$$

und somit

$$A(\lambda x + \mu y) = \lambda Ax + \mu Ay,$$

d.h. A ist eine *lineare Abbildung*. Damit erhält man leicht folgende Aussage bzgl. der Lösungsstruktur eines GLS. Wir werden später, z.B. bei den *linearen* Differentialgleichung, eine analoge Struktur haben.

Satz 1 Die allgemeine Lösung x_{inh} eines inhomogenen GLS $Ax = b$ ist die Summe aus der allgemeinen Lösung x_{hom} des zugehörigen homogenen GLS $Ax = 0$ und einer speziellen Lösung x_s des inhomogenen GLS. $\boxed{x_{inh} = x_{hom} + x_s}$

Beweis: Es seien x und y zwei beliebige Lösungen des inhomogenen GLS. Dann folgt aus $Ax = b$ und $Ay = b$ wegen der Linearität sofort

$$0 = b - b = Ax - Ay = A(x - y),$$

d.h. $x - y$ ist Lösung des zugehörigen homogenen GLS. ■

Beispiel 3 Im Beispiel 2 im Abschnitt 4.1 gilt

$$x_{inh} := x, \quad x_s := \tilde{x}_0, \quad x_{hom} := u\tilde{x}_1 + t\tilde{x}_2.$$

Wir führen jetzt für Matrizen eine *Addition*, *Subtraktion* und eine *Multiplikation* (Hintereinanderausführung) ein.

Es seien $A = (a_{ij})$ und $B = (b_{ij})$ zwei (m, n) -Matrizen und λ eine reelle Zahl. Dann wird durch

$$A + B = (a_{ij} + b_{ij})$$

die Addition der Matrizen A und B und durch

$$\lambda A = (\lambda \cdot a_{ij})$$

die skalare Multiplikation der Matrix A *elementeweise* definiert.

Es gilt nach dieser Definition

$$(\lambda A + \mu B)x = \lambda Ax + \mu Bx$$

für alle $x \in \mathbb{R}^n$ und alle reellen Zahlen λ, μ .

Beispiel 4 Wir betrachten die beiden Matrizen

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & -1 \\ -1 & -2 & 3 \end{pmatrix}.$$

Dann gilt

$$A + B = \begin{pmatrix} 1 & 2 & 2 \\ -1 & 0 & 3 \end{pmatrix}, \quad (-3)A = \begin{pmatrix} -3 & -6 & -9 \\ 0 & -6 & 0 \end{pmatrix}$$

Weiterhin bezeichnen wir mit $0 = (0)$ die *Null-Matrix* und mit $-A = (-a_{ij})$ die zu $A = (a_{ij})$ *entgegengesetzte Matrix*.

Damit gilt

$$A + 0 = 0 + A = A, \quad A + (-A) = -A + A = 0.$$

Deshalb bezeichnet man die Null-Matrix als das *neutrale Element* und $-A$ als das *inverse Element* zu A bezüglich der **Addition**. Die Gültigkeit der entsprechenden Assoziativ-, Kommutativ- und Distributivgesetze ist leicht nachweisbar, da sie elementweise auf den entsprechenden Eigenschaften der reellen Zahlen beruhen.

Falls $A = (a_{ij})$ eine (m, n) -Matrix ist, so erhält man die zu A *transponierte Matrix* A^T durch Vertauschen der Zeilen und Spalten von A . Somit ist $A^T = (a_{ji})$ eine (n, m) -Matrix.

Beispiel 5

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Eine Matrix heißt *symmetrisch*, falls $A = A^T$ gilt. Sie ist also stets quadratisch und geht bei Spiegelung an der Hauptdiagonale in sich über.

Beispiel 6

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix} = A^T.$$

Es seien A und B zwei $(2, 2)$ -Matrizen. Wir führen jetzt das *Matrizenprodukt* AB dieser beiden Matrizen ein.

Es gelte

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

Dann definieren wir die Matrix $AB = C$ durch

$$\begin{aligned} c_{11} &= a_{11}b_{11} + a_{12}b_{21}, & c_{12} &= a_{11}b_{12} + a_{12}b_{22} \\ c_{21} &= a_{21}b_{11} + a_{22}b_{21}, & c_{22} &= a_{21}b_{12} + a_{22}b_{22}, \end{aligned}$$

d.h. es gilt unter Verwendung der Zeilenvektoren a_i und der Spaltenvektoren b^j

$$c_{ij} = \langle a_i, b^j \rangle, \quad i, j = 1, 2,$$

(Vorzeile \times Nachspalte).

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} c_{11} & \boxed{c_{12}} \\ c_{21} & c_{22} \end{pmatrix}.$$

Es lässt sich zeigen, dass

$$A(Bx) = (AB)x$$

für alle $x \in \mathbb{R}^2$ gilt. Somit entspricht das Matrizenprodukt AB der Hintereinanderausführung $A \circ B$.

Die Definition des Matrizenproduktes und diese Eigenschaft kann man auf beliebige (n, n) -Matrizen A und B mit $n \geq 3$ verallgemeinern.

Beispiel 7 (a) Gegeben seien die Matrizen

$$A = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Dann gilt

$$AE = EA = A,$$

d.h. die Matrix E ist das *neutrale Element* bezüglich des **Matrizenproduktes** und wird als *Einheitsmatrix* bezeichnet. Sie entspricht der *identischen Abbildung* ($Ex = x$ für alle $x \in \mathbb{R}^n$, hier $n = 3$).

(b) Im allgemeinen gilt $AB \neq BA$, d.h. die Reihenfolge der Anwendung von A und B ist wichtig!

So erhalten wir z.B. für

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 0 \\ 1 & 1 \end{pmatrix},$$

dass

$$AB = \begin{pmatrix} 6 & 2 \\ 5 & 1 \end{pmatrix} \neq BA = \begin{pmatrix} 4 & 8 \\ 2 & 3 \end{pmatrix}.$$

Definition 2 Es sei A eine (n, n) -Matrix. Falls eine (n, n) -Matrix B existiert, so dass

$$AB = BA = E$$

gilt, so heißt die Matrix B die zu A inverse Matrix (bezüglich der **Matrizenmultiplikation**).

Wir setzen dann $B = A^{-1}$. In diesem Falle nennt man A invertierbar oder regulär.

Bemerkung 1 (a) Hierbei ist natürlich E die (n, n) -Einheitsmatrix. Falls A regulär ist, so ist A^{-1} eindeutig bestimmt.

Es gelte

$$AB = BA = E, \quad AC = CA = E.$$

Wir zeigen, dass dann $B = C$ folgt. Das folgt aus

$$B = EB = (CA)B = C(AB) = CE = C.$$

(b) Es gilt wegen $AA^{-1} = A^{-1}A = E$

$$A(A^{-1}x) = A^{-1}(Ax) = Ex = x \quad \text{für alle } x \in \mathbb{R}^n.$$

Wir erhalten also

$$x \xrightarrow{A^{-1}} A^{-1}x \xrightarrow{A} x$$

und

$$x \xrightarrow{A} Ax \xrightarrow{A^{-1}} x$$

Satz 2 Es sei A eine reguläre (n, n) -Matrix.

(a) A ist invertierbar genau dann, wenn A^{-1} invertierbar ist.

(b) $(A^{-1})^{-1} = A$.

(c) Ist außerdem B eine reguläre (n, n) -Matrix, so ist AB ebenfalls eine reguläre (n, n) -Matrix, und es gilt

$$(AB)^{-1} = B^{-1}A^{-1}.$$

(d) Das GLS $Ax = b$ hat für jedes $b \in \mathbb{R}^n$ genau eine Lösung $x \in \mathbb{R}^n$. Sie ist gegeben durch

$$x = A^{-1}b.$$

Der Beweis dieser Eigenschaften ist leicht. Wir zeigen nur (c). Wegen

$$B^{-1}A^{-1}AB = B^{-1}(A^{-1}A)B = B^{-1}EB = B^{-1}B = E$$

und der Eindeutigkeit der inversen Matrix ergibt sich die Aussage (c).

Satz 2(d) unterstreicht, dass die Existenz und die Berechnung der inversen Matrix wichtig ist für die Lösung von GLS.

Wir beschäftigen uns deshalb jetzt mit zwei Verfahren für die Berechnung der inversen Matrix (falls diese überhaupt existiert).

1. Verfahren (Anwendung des Gauß-Algorithmus)

Es sei A eine (n, n) -Matrix. Wir suchen also eine (n, n) -Matrix, die wir X nennen wollen, so dass $AX = E$ gilt. Bezeichnen wir mit

$$e^i = \begin{pmatrix} e_{1i} \\ \dots \\ e_{ii} \\ \dots \\ e_{ni} \end{pmatrix} = \begin{pmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{pmatrix}, \quad x^i = \begin{pmatrix} x_{1i} \\ \dots \\ x_{ii} \\ \dots \\ x_{ni} \end{pmatrix}$$

den i -ten Spaltenvektor von E bzw. X , so muss also

$$Ax^i = e^i, \quad i = 1, \dots, n,$$

gelten. Wir haben also n GLS mit gleicher Koeffizientenmatrix A zu lösen. Wir können also den Gauß-Algorithmus simultan auf diese n GLS anwenden. Wir erhalten somit die inverse Matrix X aus $AX = E$, falls die Matrix A durch äquivalente Umformungen in die Einheitsmatrix E überführt werden kann. In diesem Falle gilt

$$AX = E \iff EX = A^{-1} \iff X = A^{-1},$$

d.h.

$$\begin{array}{cccc|cccc} a_{11} & \dots & a_{1n} & | & 1 & \dots & 0 & \\ \dots & \dots & \dots & | & \dots & \dots & \dots & \\ a_{n1} & \dots & a_{nn} & | & 0 & \dots & 1 & \end{array}$$

wird in

$$\begin{array}{cccc|cccc} 1 & \dots & 0 & | & x_{11} & \dots & x_{1n} & \\ \dots & \dots & \dots & | & \dots & \dots & \dots & \\ 0 & \dots & 1 & | & x_{n1} & \dots & x_{nn} & \end{array}$$

umgewandelt. Falls E nicht erreichbar ist, so ist A nicht regulär. A heißt dann *singulär*.

Beispiel 8 Wir untersuchen, ob die Matrix

$$B = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 1 & -1 \\ -2 & -1 & 2 \end{pmatrix}$$

regulär ist und berechnen in diesem Fall die inverse Matrix B^{-1} . Wir erhalten

$$\begin{array}{ccc|ccc} 2 & 1 & -1 & | & 1 & 0 & 0 & \boxed{1} & 1 & -1 & | & 0 & 1 & 0 \\ \boxed{1} & 1 & -1 & | & 0 & 1 & 0 & & 2 & 1 & -1 & | & 1 & 0 & 0 \\ -2 & -1 & 2 & | & 0 & 0 & 1 & & -2 & -1 & 2 & | & 0 & 0 & 1 \\ \\ 1 & 1 & -1 & | & 0 & 1 & 0 & & 1 & 1 & -1 & | & 0 & 1 & 0 \\ 0 & -1 & 1 & | & 1 & -2 & 0 & & 0 & \boxed{1} & 0 & | & 0 & 2 & 1 \\ 0 & \boxed{1} & 0 & | & 0 & 2 & 1 & & 0 & -1 & 1 & | & 1 & -2 & 0 \\ \\ 1 & 1 & -1 & | & 0 & 1 & 0 & & 1 & 0 & -1 & | & 0 & -1 & -1 \\ 0 & \boxed{1} & 0 & | & 0 & 2 & 1 & & 0 & 1 & 0 & | & 0 & 2 & 1 \\ 0 & 0 & 1 & | & 1 & 0 & 1 & & 0 & 0 & \boxed{1} & | & 1 & 0 & 1 \\ \\ 1 & 0 & 0 & | & 1 & -1 & 0 & & & & & & & & \\ 0 & 1 & 0 & | & 0 & 2 & 1 & & \implies B^{-1} = & \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \\ 0 & 0 & 1 & | & 1 & 0 & 1 & & & & & & & & \end{array}$$

Man kann sich leicht davon überzeugen, dass $BB^{-1} = B^{-1}B = E$ gilt. Dieses Verfahren ist immer anwendbar. Der Nachteil besteht jedoch darin, dass man kein einfaches Kriterium hat, ob überhaupt die inverse Matrix existiert. Es kann also passieren, dass man erst nach großem Rechenaufwand erkennt, dass die Matrix singularär ist.

2. Verfahren (Cramersche Regel am Beispiel von $(2, 2)$ -Matrizen)

Gegeben sei die allgemeine $(2, 2)$ -Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Gesucht ist die inverse $(2, 2)$ -Matrix

$$A^{-1} = B = \begin{pmatrix} u & v \\ w & z \end{pmatrix}.$$

Wir wollen jetzt die Elemente u, v, w, z berechnen. Wir erhalten aus

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} u & v \\ w & z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

die beiden GLS

$$au + \boxed{bw} = 1$$

$$cu + \boxed{dw} = 0$$

und

$$av + bz = 0$$

$$cv + dz = 1$$

Wir suchen jetzt nach *notwendigen Bedingungen* für die Lösbarkeit dieser GLS.

Multiplizieren wir im 1. GLS die 1. Gleichung mit d und die zweite mit b , so erhalten wir

$$adu + bdw = d$$

$$bcu + bdw = 0.$$

Subtrahieren wir nun die zweite Gleichung von der ersten, so folgt

$$adu - bcu = (ad - bc)u = d. \quad (10)$$

Analog erhält man aus dem 1. GLS, indem man die 1. Gleichung mit c und die zweite mit a multipliziert wegen

$$\boxed{au} + bw = 1$$

$$\boxed{cu} + dw = 0.$$

$$bcw - adw = (bc - ad)w = c. \quad (11)$$

Wir nehmen an, dass $ad - bc = 0$ gelten. Dann würde aus (10) und (11) die Beziehung $c = d = 0$ folgen. Damit ergibt sich jedoch ein Widerspruch zu $cv + dz = 1$. Die notwendige Bedingung für die Lösbarkeit lautet also

$$ad - bc \neq 0. \quad (12)$$

Falls (12) erfüllt ist, so erhalten wir aus (10) und (11)

$$u = \frac{d}{ad - bc}, \quad w = \frac{-c}{ad - bc}.$$

Analoge Untersuchungen für das 2. GLS liefern

$$v = \frac{-b}{ad - bc}, \quad z = \frac{a}{ad - bc}.$$

Wir haben somit folgende Aussage hergeleitet.

Satz 3 Es sei die $(2, 2)$ -Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

gegeben.

(a) Die inverse Matrix A^{-1} existiert genau dann, wenn (12) gilt. Wir erhalten dann

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad (13)$$

(b) Das GLS $Ax = q$, $q \in \mathbb{R}^2$, hat genau dann eine eindeutig bestimmte Lösung

$$x = A^{-1}q,$$

falls (12) gilt.

Bemerkung 2 Das Kriterium $ad - bc \neq 0$ entspricht bei der Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

der Differenz aus der Multiplikation der Elemente auf der Hauptdiagonale und der Elemente auf der *Nebendiagonale* von A . Diese Größe führt jetzt zum Begriff der *Determinante* einer (n, n) -Matrix.

Als Anwendung von Satz 3 betrachten wir folgendes Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 &= 4 \\ 2x_1 + 3x_2 &= 5. \end{aligned}$$

In diesem Fall haben wir

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, \quad q = \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

Damit folgt nach (a)

$$A^{-1} = \frac{1}{-1} \begin{pmatrix} 3 & -2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$$

sowie nach (b)

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A^{-1}q = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \end{pmatrix}.$$

4.3 Determinanten

Definition 3 (a) Es sei die (2, 2)-Matrix A gegeben durch

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Dann heißt die Zahl $ad - bc$ die Determinante zweiter Ordnung von A . Wir schreiben

$$\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \quad (14)$$

(b) Es sei die (3, 3)-Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

gegeben. Dann definieren wir die Determinante dritter Ordnung von A durch

$$\begin{aligned} \det A = & +a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & -a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \end{aligned} \quad (15)$$

Beispiel 9 (a) Man verwendet auch die Schreibweise

$$\det A = |A|.$$

Im Fall $n = 3$, siehe Definition 3 (b), erhalten wir die sogenannte *Sarrus'sche Regel*. Diese entspricht folgender Merkmregel.

$$\begin{array}{ccccc} & & & & a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} & a_{11} & a_{12} & & \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} & & \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} & & \\ & & & & & a_{11} & a_{12} & a_{13} \\ & & & & & a_{21} & a_{22} & a_{23} \end{array} \quad \text{oder}$$

(b) Wir berechnen die Determinante zweiter Ordnung

$$\det A = \begin{vmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{vmatrix}. \quad (16)$$

Wir erhalten $\det A = \cos \varphi \cdot r \cos \varphi - (-r \sin \varphi \cdot \sin \varphi) = r \cos^2 \varphi + r \sin^2 \varphi = r$.

(c) Wir berechnen die Determinante dritter Ordnung

$$\det B = \begin{vmatrix} \cos \varphi \sin \vartheta & -r \sin \varphi \sin \vartheta & r \cos \varphi \cos \vartheta \\ \sin \varphi \sin \vartheta & r \cos \varphi \sin \vartheta & r \sin \varphi \cos \vartheta \\ \cos \vartheta & 0 & -r \sin \vartheta \end{vmatrix}. \quad (17)$$

In diesem Fall erhalten wir

$$\begin{aligned}
 \det B &= \cos \varphi \sin \vartheta \cdot r \cos \varphi \sin \vartheta \cdot (-r \sin \vartheta) \\
 &\quad + \sin \varphi \sin \vartheta \cdot 0 \cdot r \cos \varphi \cos \vartheta \\
 &\quad + \cos \vartheta \cdot (-r \sin \varphi \sin \vartheta) \cdot r \sin \varphi \cos \vartheta \\
 &\quad - \cos \vartheta \cdot r \cos \varphi \sin \vartheta \cdot r \cos \varphi \cos \vartheta \\
 &\quad - \cos \varphi \sin \vartheta \cdot 0 \cdot r \sin \varphi \cos \vartheta \\
 &\quad - \sin \varphi \sin \vartheta \cdot (-r \sin \varphi \sin \vartheta) \cdot (-r \sin \vartheta) \\
 &= -r^2 \sin^2 \vartheta \sin \vartheta \cos^2 \varphi - r^2 \sin \vartheta \cos^2 \vartheta \sin^2 \varphi \\
 &\quad - r^2 \sin \vartheta \cos^2 \vartheta \cos^2 \varphi - r^2 \sin^2 \vartheta \sin \vartheta \sin^2 \varphi \\
 &= -r^2 \left(\sin \vartheta \cos^2 \vartheta (\sin^2 \varphi + \cos^2 \varphi) + \sin^2 \vartheta \sin \vartheta (\sin^2 \varphi + \cos^2 \varphi) \right) \\
 &= -r^2 \sin \vartheta (\cos^2 \vartheta + \sin^2 \vartheta) = -r^2 \sin \vartheta.
 \end{aligned} \tag{18}$$

Die Determinante einer (n, n) -Matrix lässt sich allgemein stets mittels des Gaußschen Algorithmus bestimmen. Dazu verwendet man insbesondere die im folgenden Lemma angegebenen Eigenschaften (d) – (h) für die äquivalenten Umformungen.

Wir geben nun ohne Beweis einige Rechenregeln für Determinanten an. (Es ist jedoch nicht schwer, diese Eigenschaften der Determinante für $n = 2, 3$ nachzuvollziehen, insbesondere folgt (b) aus (a) und (f) aus (e).) Wir sagen, dass eine (n, n) -Matrix Dreiecksgestalt hat, falls unterhalb der Hauptdiagonale nur Nullen stehen.

Lemma 1 *Es seien A und B zwei (n, n) -Matrizen.*

- (a) $\det(AB) = \det A \cdot \det B$.
- (b) $\det E = 1$. Gilt $\det A \neq 0$, so existiert A^{-1} und $\det A^{-1} = \frac{1}{\det A}$.
- (c) $\det A = \det A^T$.
- (d) Die Determinante einer Dreiecksmatrix ist gleich dem Produkt ihrer Diagonalelemente.
- (e) Beim Vertauschen von zwei Zeilen (Spalten) ändert sich das Vorzeichen der Determinante.
- (f) Hat eine Matrix A zwei gleiche Zeilen (Spalten), so gilt $\det A = 0$.
- (g) Der Wert einer Determinante ändert sich nicht, wenn man zu einer Zeile (Spalte) ein Vielfaches einer anderen Zeile (Spalte) addiert.
- (h) Bei Multiplikation einer Zeile (Spalte) mit einer reellen Zahl λ ändert sich der Wert der Determinante um das λ -fache.

Beispiel 10 (a) Wir zeigen jetzt an einem Beispiel, wie man diese Rechenregeln zur Berechnung

der Determinante der (4, 4)-Matrix

$$A = \begin{pmatrix} 4 & 0 & 4 & 1 \\ 7 & 1 & 0 & 1 \\ 8 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

verwenden kann. Wir vertauschen zuerst die 1. und 3. Spalte, danach multiplizieren wir die 3. Zeile mit $1/8$ und subtrahieren schließlich die 3. Zeile von der 4. Zeile.

$$\begin{aligned} \det A &= (-1) \cdot \begin{vmatrix} \boxed{4} & 0 & 4 & 1 \\ 0 & \boxed{1} & 7 & 1 \\ 0 & 0 & 8 & 1 \\ 0 & 0 & 1 & 1 \end{vmatrix} = (-1) \cdot 8 \cdot \begin{vmatrix} \boxed{4} & 0 & 4 & 1 \\ 0 & \boxed{1} & 7 & 1 \\ 0 & 0 & \boxed{1} & \frac{1}{8} \\ 0 & 0 & 1 & 1 \end{vmatrix} \\ &= (-1) \cdot 8 \cdot \begin{vmatrix} \boxed{4} & 0 & 4 & 1 \\ 0 & \boxed{1} & 7 & 1 \\ 0 & 0 & \boxed{1} & \frac{1}{8} \\ 0 & 0 & 0 & \boxed{\frac{7}{8}} \end{vmatrix} = (-1) \cdot 8 \cdot 4 \cdot 1 \cdot 1 \cdot \frac{7}{8} = -28. \end{aligned}$$

(b) Schaut man sich den Gaußschen Algorithmus im Beispiel 8 noch einmal an, so erkennt man, dass nach den obigen Rechengesetzen $\det B = 1$ gilt.

Wir bemerken, dass es weitere Verfahren zur Berechnung der Determinante gibt. Zum Beispiel beruht der Entwicklungssatz von Laplace darauf, dass die Bestimmung einer Determinante der Ordnung n auf die Untersuchung von n Determinanten der Ordnung $(n - 1)$ zurückführt wird. Dadurch erreicht man nach endlich vielen Schritten mindestens Determinanten der Ordnung 3. Es sei

$$\det A = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}, \quad n \geq 3.$$

Als *Unterdeterminante* des Elementes a_{ij} , $1 \leq i, j \leq n$, bezeichnet man die Determinante $(n-1)$ -ter Ordnung A^{ij} , die man aus der Ausgangs-Determinante durch Streichung der i -ten Zeile und j -ten Spalte erhält. Setzen wir die

$$A_{ij} = (-1)^{i+j} A^{ij}$$

(**Adjunkte** zum Element a_{ij}) so gilt

$$\det A = \sum_{i=1}^n a_{ij} A_{ij}, \quad j = 1, \dots, n,$$

(Entwicklung nach der j -ten Spalte) sowie

$$\det A = \sum_{j=1}^n a_{ij} A_{ij}, \quad i = 1, \dots, n,$$

(Entwicklung nach der i -ten Zeile). Uns stehen also $2n$ Möglichkeiten zur Verfügung um $\det A$ zu berechnen. Die Idee besteht natürlich darin, eine Variante zu verwenden, wo in der Zeile bzw. Spalte viele Nullen stehen!

Es sei

$$A = \begin{pmatrix} 4 & 4 & 1 \\ 8 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Entwickeln wir jetzt nach der 2. Spalte, so folgt

$$\det A = (-4) \begin{vmatrix} 8 & 1 \\ 1 & 1 \end{vmatrix} = -28.$$

Analog erhält man für

$$B = \begin{pmatrix} 4 & 0 & 4 & 1 \\ 7 & 1 & 0 & 1 \\ 8 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

bei der Entwicklung nach der 2. Spalte

$$\det B = (+1) \det A = -28.$$

Als Verallgemeinerung von Satz 3 erhalten wir die folgenden Aussagen.

Satz 4 *A sei eine (n, n) -Matrix.*

(a) *Die inverse Matrix A^{-1} existiert genau dann, wenn $\det A \neq 0$ erfüllt ist.*

(b) *Das GLS $Ax = b$ besitzt für jede rechte Seite $b \in \mathbb{R}^n$ genau dann eine eindeutig bestimmte Lösung $x \in \mathbb{R}^n$, wenn $\det A \neq 0$ gilt.*

Der Satz 4 liefert ein *notwendiges und hinreichendes* Kriterium für die Lösbarkeit von GLS, jedoch nicht ein Verfahren für die Berechnung der Lösung. Deshalb betrachten wir noch einmal die früheren Untersuchungen für den Fall $n = 2$. Es sei also

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}.$$

Es gilt

$$Ax = q \iff x = A^{-1}q.$$

Nach (13) wissen wir

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Somit erhalten

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} dq_1 - bq_2 \\ -cq_1 + aq_2 \end{pmatrix}.$$

Anschaulich bedeutet das

$$x_1 = \frac{1}{\det A} (+dq_1 - bq_2) = \frac{1}{\det A} \begin{vmatrix} q_1 & b \\ q_2 & d \end{vmatrix} \quad (19)$$

bzw.

$$x_2 = \frac{1}{\det A} (-cq_1 + aq_2) = \frac{1}{\det A} \begin{vmatrix} a & q_1 \\ c & q_2 \end{vmatrix}. \quad (20)$$

Man erhält also die Koordinaten x_i der Lösung x , indem man in der Determinante von A den i -ten Spaltenvektor durch die rechte Seite q des inhomogenen GLS ersetzt und anschließend durch den Wert $\det A$ dividert.

Wir verallgemeinern dieses Ergebnis, d.h. (19) und (20), auf den Fall $n \geq 2$.

Satz 5 (Cramersche Regel) *Es sei A eine reguläre (n, n) -Matrix, d.h. es gilt $\det A \neq 0$. Es sei $b \in \mathbb{R}^n$ mit*

$$b = \begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix}.$$

Dann ist $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ mit

$$x_i = \frac{1}{\det A} \begin{vmatrix} a_{11} & \dots & a_{1i-1} & b_1 & a_{1i+1} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{ni-1} & b_n & a_{ni+1} & \dots & a_{nn} \end{vmatrix}, \quad i = 1, \dots, n, \quad (21)$$

die eindeutig bestimmte Lösung des GLS $Ax = b$.

Bemerkung 3 Die Anwendung der Cramerschen Regel ist besonders sinnvoll für $n = 2, 3$, da man in diesen Fällen ziemlich schnell alle auftretenden Determinanten berechnen kann. Ansonsten sollte man i.a. besser den Gaußschen Algorithmus benutzen. Man beachte, dass die Cramersche Regel nur für reguläre (n, n) -Matrizen anwendbar ist!

Beispiel 11 (a) Wir betrachten folgendes GLS

$$\begin{aligned}x_1 + 2x_2 + 3x_3 &= 1 \\5x_2 + 6x_3 &= 0 \\7x_1 + 8x_2 &= 0.\end{aligned}$$

Wir erhalten

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Nach der Sarrus'schen Regel erhalten wir wegen

$$\begin{array}{ccc} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 7 & 8 & 0 \end{array} \quad \det A = 7 \cdot 8 \cdot 0 = 0 + 0 + 84 - 105 - 48 - 0 = -69.$$

$$\begin{array}{ccc} 1 & 2 & 3 \\ 0 & 5 & 6 \end{array}$$

Wenden wir jetzt Satz 5 an, so ist die eindeutig bestimmte Lösung $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ des obigen GLS mittels Sarrus'scher Regel (jeweils nur ein Term ist verschieden von Null) durch

$$x_1 = -\frac{1}{69} \begin{vmatrix} \boxed{1} & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 8 & 0 \end{vmatrix} = -\frac{1}{69} \cdot (-48) = \frac{48}{69},$$

$$x_2 = -\frac{1}{69} \begin{vmatrix} 1 & \boxed{1} & 3 \\ 0 & 0 & 6 \\ 7 & 0 & 0 \end{vmatrix} = -\frac{1}{69} \cdot 42 = -\frac{42}{69},$$

$$x_3 = -\frac{1}{69} \begin{vmatrix} 1 & 2 & \boxed{1} \\ 0 & 5 & 0 \\ 7 & 8 & 0 \end{vmatrix} = -\frac{1}{69} \cdot (-35) = \frac{35}{69}$$

gegeben.

(b) Ersetzen wir jetzt den obigen Vektor b sukzessiv durch

$$b = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

so würde man mit der Cramerschen Regel ebenfalls die Koeffizienten der inversen Matrix A^{-1} erhalten.

4.4 Basis, Koordinatentransformation

Wie üblich sei

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

die Menge aller geordneten n -Tupel reeller Zahlen.

Die *kanonische Basis* im \mathbb{R}^2 ist gegeben durch die Basisvektoren

$$e^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

bzw. im \mathbb{R}^3 durch

$$e^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e^3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Diese Basen besitzen folgende Eigenschaften:

- Vollständigkeit: Jedes $x \in \mathbb{R}^n$, $n = 2, 3$, lässt sich als Linearkombination der Basisvektoren e^i , $i = 1, \dots, n$, darstellen.
- Minimalität: Falls wir einen Basisvektor entfernen, so ist die Vollständigkeit nicht mehr erfüllt.
- eindeutige Darstellung von $x \in \mathbb{R}^n$, $n = 2, 3$, durch die Koordinaten bzgl. der Basis.
- Orthonormalität (*Orthonormalbasis*): Es gilt

$$|e^i| = 1, \quad \langle e^i, e^j \rangle = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad \delta_{ij} : \text{Kronecker-Symbol.}$$

Wir wollen diese Eigenschaften verallgemeinern.

Definition 4 Es seien $k, n \in \mathbb{N}$. Die Vektoren $v^1, \dots, v^k \in \mathbb{R}^n$ heißen linear unabhängig, falls aus

$$\alpha_1 v^1 + \dots + \alpha_k v^k = 0 \quad \text{im } \mathbb{R}^n \tag{22}$$

die Aussage $\alpha_1 = \dots = \alpha_k = 0$ folgt, d.h. es gibt nur eine triviale Darstellung des Nullelementes $0 \in \mathbb{R}^n$.

Ansonsten heißen die Vektoren linear abhängig.

Die Untersuchung der linearen Unabhängigkeit führt also zur Bestimmung der Lösungen des homogenen GLS (22). Die Vektoren sind linear unabhängig, falls (22) nur die triviale Lösung $\alpha_1 = \dots = \alpha_k = 0$ hat. Falls nichttriviale Lösungen existieren, sind die Vektoren linear abhängig.

Beispiel 12 (a) Die Vektoren

$$w^1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad w^2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad w^3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

sind nach Definition 4 im \mathbb{R}^3 linear unabhängig, da das homogene GLS

$$\alpha_1 \begin{pmatrix} 1 \\ \boxed{1} \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ \boxed{1} \end{pmatrix} + \alpha_3 \begin{pmatrix} \boxed{1} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

(Lösung kann man erraten) nur die triviale Lösung $\alpha_1 = \alpha_2 = \alpha_3 = 0$ hat:

Aus der zweiten Gleichung folgt $\alpha_1 = 0$, somit mit der dritten Gleichung aus $\alpha_1 + \alpha_2 = 0$, dass $\alpha_2 = 0$ und schließlich mit der ersten Gleichung aus $\alpha_1 + \alpha_2 + \alpha_3 = 0$, dass $\alpha_3 = 0$ gelten muss. Hierbei interpretieren wir also die Menge $B = \{w^1, w^2, w^3\}$ als Matrix

$$B = (w^1 \ w^2 \ w^3),$$

d.h. die Elemente der Menge B werden als Spaltenvektoren der Matrix B aufgefasst.

Das gleiche Ergebnis erhält man auch unter Verwendung der Eigenschaften der $(3,3)$ -Determinante

$$\det B = \begin{vmatrix} 1 & 1 & \boxed{1} \\ \boxed{1} & 0 & 0 \\ 1 & \boxed{1} & 0 \end{vmatrix} = 1 \neq 0.$$

Folglich besitzt das homogene GLS eine eindeutige Lösung, die durch die stets existierende triviale Lösung gegeben ist.

(b) Die Vektoren

$$v^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad v^2 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \quad v^3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad v^4 = \begin{pmatrix} 3 \\ 3 \\ 4 \\ 5 \end{pmatrix}$$

sind im \mathbb{R}^4 linear abhängig. Es gilt

$$v^1 + v^2 + v^3 - v^4 = 0,$$

d.h. wir können z. B. $\alpha_1 = \alpha_2 = \alpha_3 = 1$, $\alpha_4 = -1$ wählen.

Man zeige selbst, dass alle folgenden Mengen von Vektoren

$$\{v^1, v^2, v^3\}, \quad \{v^2, v^3, v^4\}, \quad \{v^1, v^3, v^4\}$$

ebenfalls im \mathbb{R}^4 linear unabhängig sind.

(c) Falls $0 \in M$, so ist M stets linear abhängig.

(d) Die lineare Unabhängigkeit von n Vektoren im \mathbb{R}^n garantiert die *Eindeutigkeit der Darstellung* eines Elementes $x \in \mathbb{R}^n$ als *Linearkombination* dieser n Vektoren.

Falls

$$x = \sum_{i=1}^n \alpha_i v^i = \sum_{i=1}^n \beta_i v^i,$$

so folgt wegen

$$0 = \sum_{i=1}^n (\alpha_i - \beta_i) v^i$$

aus der linearen Unabhängigkeit von v^i , $i = 1, \dots, n$, dass

$$\alpha_i = \beta_i, \quad i = 1, \dots, n,$$

gelten muss. Das ist jedoch nicht für linear abhängige Vektoren richtig. So folgt z.B. im \mathbb{R}^4 wegen

$$v^1 + v^2 + v^3 = v^4,$$

siehe (b),

$$0 = 0v^1 + 0v^2 + 0v^3 + 0v^4 = 1v^1 + 1v^2 + 1v^3 - 1v^4,$$

dass die Darstellung des Nullelementes $0 \in \mathbb{R}^4$ durch die Vektoren v^1, v^2, v^3, v^4 nicht eindeutig ist.

Damit ist folgende Definition sinnvoll.

Definition 5 *Es sei die Menge $B = \{v^1, v^2, \dots, v^n\} \subset \mathbb{R}^n$ gegeben.*

(a) *B heißt Basis im \mathbb{R}^n , falls die n Vektoren v^1, \dots, v^n linear unabhängig sind.*

(b) *Es sei B eine Basis in \mathbb{R}^n und $x \in \mathbb{R}^n$. Falls*

$$x = \sum_{i=1}^n \xi_i v^i$$

gilt, so bezeichnet

$$x_B = \begin{pmatrix} \xi_1 \\ \dots \\ \xi_n \end{pmatrix}$$

die Koordinaten von x bezüglich der Basis B .

Wir bemerken, dass im \mathbb{R}^n die maximale Anzahl von linear unabhängigen Vektoren genau n ist. Deshalb nennt man den \mathbb{R}^n auch einen *n-dimensionalen Raum*. Es ist auch üblich die Koordinaten als Zeilenvektor $x_B = (\xi_1, \dots, \xi_n)$ zu schreiben. Im allgemeinen verwendet man bei der Koordinatendarstellung bzgl. der kanonischen Basis B_K die Schreibweise

$$x = x_{B_K} = (\xi_1, \dots, \xi_n).$$

Bisher haben wir uns nur mit der kanonischen Basis beschäftigt. Bei vielen Aufgabenstellungen, z.B. in der Mathematik, Physik und Chemie, ist man jedoch an einer Basis interessiert, die dem konkreten Problem angepasst ist. So stimmt in der Chemie bei Kristalluntersuchungen das Koordinatensystem mit der vorhandenen Molekülstruktur überein.

Um Verwechslungen zu vermeiden, verwenden wir im folgenden für

- Koordinaten — griechische Buchstaben: $\alpha, \beta, \gamma, \dots$,
- Vektoren — lateinische Kleinbuchstaben: a, b, c, \dots ,
- Matrizen — lateinische Großbuchstaben: A, B, C, \dots .

Beispiel 13 Es sei der Vektor $x \in \mathbb{R}^2$ durch

$$x = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

und die zwei Basen

$$B_1 = \left\{ v^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v^2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}, \quad B_2 = \left\{ w^1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}, w^2 = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \right\}$$

bezüglich der **kanonischen** Basis $B_K = \{e^1, e^2\} \subset \mathbb{R}^2$ gegeben. Es gilt also

$$x = x_{B_K} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}.$$

Wir berechnen jetzt die Koordinaten x_{B_1} von x bezüglich der Basis B_1 . Wir erhalten aus

$$x = \alpha_1 v^1 + \alpha_2 v^2 = \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & \boxed{2} \\ 1 & \boxed{1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix},$$

dass $\alpha_2 = 1$ und somit $\alpha_1 = 2$ gilt. Somit ergibt sich

$$x_{B_1} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Wir berechnen jetzt die Koordinaten x_{B_2} von x bezüglich der Basis B_2 . Wir haben

$$x = \beta_1 w^1 + \beta_2 w^2 = \begin{pmatrix} 4 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix},$$

und damit mittels Cramerscher Regel

$$\beta_1 = \frac{1}{13} \begin{vmatrix} 4 & -1 \\ 3 & 3 \end{vmatrix} = \frac{15}{13}, \quad \beta_2 = \frac{1}{13} \begin{vmatrix} 4 & 4 \\ 1 & 3 \end{vmatrix} = \frac{8}{13}.$$

Somit erhalten wir

$$x_{B_2} = \begin{pmatrix} \frac{15}{13} \\ \frac{8}{13} \end{pmatrix}.$$

Interpretieren wir jetzt die Basen B_1 und B_2 wie in den Beispielen 12 und 13 als Matrizen, d.h. wir fassen die Basiselemente als Spaltenvektoren in den entsprechenden Matrizen auf, so erhalten wir

$$B_1 = (v^1 \ v^2), \quad B_2 = (w^1 \ w^2).$$

Damit gilt nach den obigen Untersuchungen für alle $x \in \mathbb{R}^2$

$$x = B_1 x_{B_1} = B_2 x_{B_2} \quad \iff \boxed{B_1 x_{B_1} = B_2 x_{B_2}}.$$

Da beide Matrizen wegen der linearen Unabhängigkeit der Basisvektoren regulär sind, erhalten wir also

$$x_{B_1} = B_1^{-1} B_2 x_{B_2}, \quad x_{B_2} = B_2^{-1} B_1 x_{B_1}.$$

(Aus der linearen Unabhängigkeit folgt die eindeutige Lösbarkeit der homogenen Gleichungssysteme, und damit $\det B_1 \neq 0$ und $\det B_2 \neq 0$. Dieses Ergebnis lässt sich auch für den \mathbb{R}^n nachweisen.)

Satz 6 (Koordinatentransformation) Sind $B_1 = \{v^1, \dots, v^n\}$ und $B_2 = \{w^1, \dots, w^n\}$ zwei Basen des \mathbb{R}^n , so können die Koordinaten x_{B_2} von $x \in \mathbb{R}^n$ bezüglich der Basis B_2 aus dem Koordinatenvektor x_{B_1} bezüglich der Basis B_1 durch

$$x_{B_2} = B_2^{-1} B_1 x_{B_1} \tag{23}$$

berechnet werden. Die Matrix

$$S = B_2^{-1} B_1$$

wird als Übergangsmatrix von der Basis B_1 zur Basis B_2 bezeichnet.

$S^{-1} = B_1^{-1} B_2$ ist dann die Übergangsmatrix von der Basis B_2 zur Basis B_1 und es gilt

$$x_{B_1} = B_1^{-1} B_2 x_{B_2}. \tag{24}$$

Beispiel 14 Wir verwenden die gleichen Basen wie im Beispiel 13. Somit haben wir, unter Verwendung von (13) aus Satz 3,

$$B_1 = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \quad B_1^{-1} = \frac{1}{\det B_1} \begin{pmatrix} 1 & -2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}.$$

Weiterhin erhalten wir entsprechend der Rechenregeln für die Matrizenmultiplikation

$$B_1^{-1}B_2 = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 4 & -1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} -2 & 7 \\ 3 & -4 \end{pmatrix}$$

und damit für

$$x_{B_2} = \begin{pmatrix} \frac{15}{13} \\ \frac{8}{13} \end{pmatrix}$$

nach (24)

$$x_{B_1} = B_1^{-1}B_2x_{B_2} = \begin{pmatrix} -2 & 7 \\ 3 & -4 \end{pmatrix} \begin{pmatrix} \frac{15}{13} \\ \frac{8}{13} \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Spezialfall: Falls $B_1 = B_K$ die kanonische Basis ist, so ist die korrespondierende Matrix B_1 die Einheitsmatrix E . Damit ergibt sich in (24), dass $S = B_2^{-1}E = B_2^{-1}$ und $S^{-1} = B_2$ gilt.

Jetzt befassen wir uns noch kurz mit dem Verhalten von **Matrizen** bezüglich Koordinatentransformationen.

Es sei B_1 die alte Basis und B_2 die neue Basis im \mathbb{R}^n . Dann gilt nach Satz 6 und (24)

$$x_{B_2} = B_2^{-1}B_1x_{B_1} = Sx_{B_1}, \quad S = B_2^{-1}B_1,$$

wobei die Übergangsmatrix S die Basistransformation von B_1 nach B_2 bezeichnet, d.h. wir haben

$$B_2S = B_1 \iff B_2 = B_1S^{-1}.$$

Durch die Matrix A bezüglich einer Basis B_1 (z.B. wird oft als $B_1 = B_K$ die kanonische Basis verwendet) ist eine lineare Abbildung, die wir mit L_A bezeichnen wollen, definiert (im Falle $B_1 = B_K$ ist $L_Ax = Ax$). Beim Übergang zu einer neuen Basis B_2 wird diese Matrix A zu einer Matrix \tilde{A} transformiert, wenn sie die gleiche lineare Abbildung $L_A : x \mapsto L_Ax$ darstellen soll.

Wir haben also für die Abbildung L_A bzgl. der zwei verschiedenen Basen die folgenden Matrixdarstellungen

$$A : x_{B_1} \rightarrow (Ax)_{B_1}, \quad \tilde{A} : x_{B_2} \rightarrow (\tilde{A}x)_{B_2}.$$

Somit folgt mit

$$x = B_1x_{B_1} = B_2x_{B_2}, \quad L_Ax = B_1(Ax)_{B_1} = B_2(\tilde{A}x)_{B_2}$$

und

$$x_{B_2} = Sx_{B_1}, \quad (Ax)_{B_1} = S^{-1}(\tilde{A}x)_{B_2},$$

das folgende Diagramm

$$\begin{array}{l} B_1 : \quad x_{B_1} \xrightarrow{A} (Ax)_{B_1} \\ \\ B_2 : \quad x_{B_2} \xrightarrow{\tilde{A}} (\tilde{A}x)_{B_2}. \end{array}$$

Somit muss für die Matrizen die Beziehungen

$$A = S^{-1}\tilde{A}S, \quad \tilde{A} = SAS^{-1} \quad (25)$$

gelten. Falls $B_1 = B_K$ gilt, d.h. für die zugehörige Matrix ist $B_1 = E$, so erhalten wir wieder $S = B_2^{-1}$ und $S^{-1} = B_2$ in (25).

Für den Vektor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ ist durch

$$|x| = \sqrt{x_1^2 + \dots + x_n^2}$$

dessen *Betrag* (oder Norm) definiert. Damit gilt nach Definition 1 folgender Zusammenhang zwischen dem Betrag $|\cdot|$ und dem Skalarprodukt $\langle \cdot, \cdot \rangle$

$$|x| = \sqrt{x_1^2 + \dots + x_n^2} = \sqrt{\langle x, x \rangle} \iff |x|^2 = x_1^2 + \dots + x_n^2 = \langle x, x \rangle.$$

Falls $|x| = 1$ für $x \in \mathbb{R}^n$ gilt, so heißt x normierter *Einheitsvektor*.

Weiterhin sind zwei Elemente $x, y \in \mathbb{R}^n$ *orthogonal* (*senkrecht*) zueinander ($x \perp y$), falls die Bedingung $\langle x, y \rangle = 0$ erfüllt ist. Insbesondere gilt wegen $\langle 0, x \rangle = 0$, dass wir

$$0 \perp x$$

für alle $x \in \mathbb{R}^n$ haben.

Es seien n Vektoren v^1, \dots, v^n im \mathbb{R}^n gegeben, für die

$$\langle v^j, v^i \rangle = \delta_{ji} = \begin{cases} 1 & : \quad j = i \\ 0 & : \quad j \neq i \end{cases}, \quad i, j = 1, \dots, n,$$

gilt. Folglich sind diese n Vektoren im \mathbb{R}^n paarweise orthogonal sowie normierte Einheitsvektoren.

Außerdem sind sie stets linear unabhängig im \mathbb{R}^n : Es gelte $0 = \lambda_1 v^1 + \dots + \lambda_n v^n$. Daraus folgt wegen $\langle 0, v^i \rangle = 0$ und der Linearität des Skalarproduktes

$$0 = \langle 0, v^i \rangle = \left\langle \sum_{j=1}^n \lambda_j v^j, v^i \right\rangle = \sum_{j=1}^n \lambda_j \langle v^j, v^i \rangle = \lambda_j \delta_{ji} = \lambda_i, \quad i = 1, \dots, n,$$

dass alle λ_i gleich Null sind.

Falls $B = \{v^i\}_{i=1}^n$ die obigen Eigenschaft $\langle v^j, v^i \rangle = \delta_{ji}$, $i, j = 1, \dots, n$, erfüllt, so nennt man B eine *Orthonormalbasis* (ONB) im \mathbb{R}^n . Eine ONB besteht also aus n -linear unabhängigen paarweise zueinander orthogonalen Basisvektoren v^1, \dots, v^n . Damit erhalten wir wegen

$$\begin{aligned} B^T B &= \begin{pmatrix} v^1 \\ \dots \\ v^n \end{pmatrix} (v^1 \dots v^n) = \begin{pmatrix} \langle v^1, v^1 \rangle \dots \langle v^1, v^n \rangle \\ \dots \dots \dots \\ \langle v^n, v^1 \rangle \dots \langle v^n, v^n \rangle \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}. \end{aligned}$$

Da aber die inverse Matrix von B eindeutig ist, folgt hieraus

$$B^T = B^{-1}.$$

Es sei wieder B eine ONB im \mathbb{R}^n , und $x \in \mathbb{R}^n$ habe die Koordinaten $x_B = (\alpha_1, \dots, \alpha_n)$. Dann erhält man analog zur linearen Unabhängigkeit folgenden Zusammenhang zwischen den Koordinaten von x bezüglich v^i und dem Skalarprodukt $\langle x, v^i \rangle$. Aus

$$x = \alpha_1 v^1 + \dots + \alpha_n v^n$$

ergibt sich

$$\langle x, v^1 \rangle = \sum_{j=1}^n \alpha_j \langle v^j, v^1 \rangle = \alpha_j \delta_{j1} = \alpha_1,$$

und damit allgemein

$$\langle x, v^i \rangle = \sum_{j=1}^n \alpha_j \langle v^j, v^i \rangle = \alpha_j \delta_{ji} = \alpha_i, \quad i = 1, \dots, n.$$

Wir erhalten somit folgende Darstellung

$$x = \sum_{i=1}^n \langle x, v^i \rangle v^i. \tag{26}$$

Bei vielen Untersuchungen ist es günstig, mit einer ONB zu arbeiten, die dem vorgegebenen Problem angepasst ist, siehe z.B. nächster Abschnitt. Daraus ergibt sich die Frage, ob man eine beliebige Basis im \mathbb{R}^n in eine ONB überführen kann. Zur Lösung dieses Problems verwenden wir das *Schmidtsche Orthonormalisierungsverfahren*.

Gegeben sei die Basis $B_1 = \{v^1, \dots, v^n\}$ im \mathbb{R}^n . Wir müssen also mittels B_1 eine neue Basis $B_2 = \{w^1, \dots, w^n\}$ im \mathbb{R}^n bestimmen, für die gilt

$$\langle w^i, w^j \rangle = \delta_{ij}.$$

Der erste Vektor w^1 wird durch

$$w^1 = \frac{v^1}{|v^1|}$$

definiert. Jetzt bestimmen wir für $\tilde{w}^2 = v^2 + t_1 w^1$ die reelle Zahl t_1 so, dass

$$0 = \langle \tilde{w}^2, w^1 \rangle = \langle v^2, w^1 \rangle + t_1 \langle w^1, w^1 \rangle,$$

d.h. $\tilde{w}^2 \perp w^1$ gilt. Wegen $\langle w^1, w^1 \rangle = 1$, erhalten wir $t_1 = -\langle v^2, w^1 \rangle$, d.h.

$$\tilde{w}^2 = v^2 - \langle v^2, w^1 \rangle w^1.$$

Abschließend normieren wir diesen zu w^1 orthogonalen Vektor \tilde{w}^2 noch und erhalten

$$w^2 = \frac{\tilde{w}^2}{|\tilde{w}^2|} = \frac{v^2 - \langle v^2, w^1 \rangle w^1}{|v^2 - \langle v^2, w^1 \rangle w^1|}.$$

Wir setzen dieses Orthonormalisierungsverfahren mit w^3 fort. Dann folgt für $\tilde{w}^3 = v^3 + s_1 w^1 + s_2 w^2$ aus $\tilde{w}^3 \perp w^1$, $\tilde{w}^3 \perp w^2$, dass $s_1 = -\langle v^3, w^1 \rangle$ und $s_2 = -\langle v^3, w^2 \rangle$ gelten muss. Schließlich ist $w^3 = \frac{\tilde{w}^3}{|\tilde{w}^3|}$. Allgemein definieren wir

$$w^{k+1} = \frac{v^{k+1} - \sum_{i=1}^k \langle v^{k+1}, w^i \rangle w^i}{|v^{k+1} - \sum_{i=1}^k \langle v^{k+1}, w^i \rangle w^i|}, \quad k = 0, 1, \dots, n-1. \quad (27)$$

Nach Konstruktion bilden die so definierten Vektoren w^i , $i = 1, \dots, n$, die gesuchte ONB B_2 im \mathbb{R}^n .

Beispiel 15 Gegeben sei die Basis (selbst überprüfen)

$$v^1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v^2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad v^3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Wir erhalten nach (27) für $k = 0$

$$w^1 = \frac{v^1}{|v^1|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Wegen $\langle v^2, w^1 \rangle = 0$, gilt bereits $v^2 \perp w^1$. Damit folgt mit (27) für $k = 1$

$$w^2 = \frac{v^2 - \langle v^2, w^1 \rangle w^1}{|v^2 - \langle v^2, w^1 \rangle w^1|} = \frac{v^2}{|v^2|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Weiterhin erhalten wir

$$\langle v^3, w^1 \rangle = 0, \quad \langle v^3, w^2 \rangle = \frac{1}{\sqrt{2}}, \quad \langle v^3, w^2 \rangle w^2 = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad v^3 - \langle v^3, w^2 \rangle w^2 = \begin{pmatrix} \frac{1}{2} \\ -1 \\ \frac{1}{2} \end{pmatrix}.$$

Somit gilt für w^3 nach (27) im Falle $k = 2$

$$w^3 = \frac{v^3 - \langle v^3, w^1 \rangle w^1 - \langle v^3, w^2 \rangle w^2}{|v^3 - \langle v^3, w^1 \rangle w^1 - \langle v^3, w^2 \rangle w^2|} = \frac{v^3 - \langle v^3, w^2 \rangle w^2}{|v^3 - \langle v^3, w^2 \rangle w^2|} = \frac{1}{\frac{1}{2}\sqrt{6}} \begin{pmatrix} \frac{1}{2} \\ -1 \\ \frac{1}{2} \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

4.5 Eigenwertprobleme

Wir beschäftigen uns jetzt mit *Eigenwertproblemen*, die vielfältige Anwendungen haben.

Definition 6 *Es sei A eine (n, n) -Matrix. Dann heißt die reelle Zahl λ Eigenwert von A , falls ein $x \in \mathbb{R}^n \setminus \{0\}$ existiert, so dass*

$$Ax = \lambda x$$

gilt. Man nennt dann x Eigenvektor von A zum Eigenwert λ .

Bemerkung 4 (a) Da $x = 0$ (Nullvektor) stets eine triviale Lösung von $Ax = \lambda x$ für alle reellen λ ist, muss $x = 0$ in der Definition 6 ausgeschlossen werden.

(b) Die besondere Eigenschaft eines Eigenvektors x zum Eigenwert λ besteht also darin, dass sich bei Anwendung von A der Wert um das λ -fache ändert.

Beispiel 16 Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Dann ist

$$x^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Eigenvektor zum Eigenwert $\lambda_1 = 0$ und

$$x^2 = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

Eigenvektor zum Eigenwert $\lambda_2 = 2$, denn es gilt

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 4 \\ -4 \end{pmatrix}.$$

Für Eigenvektoren ergibt sich aus der Linearität von A folgende Eigenschaft.

Lemma 2 *Es seien α und β reelle Zahlen. Falls x^1 und x^2 Eigenvektoren von A zum gleichen Eigenwert λ sind, so gilt das auch für $\alpha x^1 + \beta x^2$.*

Wir beschäftigen uns jetzt mit einem Algorithmus für die Lösung des Eigenwertproblems, d.h. wir wollen für eine vorgegebene Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

alle Eigenwerte und die zugehörigen Eigenvektoren bestimmen.

Es sei $0 \neq x \in \mathbb{R}^n$ ein Eigenvektor zum Eigenwert λ . Dann folgt aus $Ax = \lambda x$, dass das homogene GLS, gegeben durch

$$(A - \lambda E)x = 0,$$

eine *nichttriviale* Lösung hat. Daraus ergibt sich aber

$$\det(A - \lambda E) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0. \quad (28)$$

Wir bezeichnen mit

$$P_n(\lambda) = \det(A - \lambda E) = c_0 + c_1\lambda + \cdots + c_{n-1}\lambda^{n-1} + (-1)^n\lambda^n$$

das *charakteristische Polynom* des obigen Eigenwertproblems. Es gilt zunächst $\text{Grad}(P_n) = n$ bezüglich des Parameters λ . Wir müssen also unter Verwendung des Fundamentalsatzes der Algebra, siehe Satz 9 im Abschnitt 2.3, die Nullstellen von $P_n(\lambda)$ bestimmen.

Haben wir $0 \leq k \leq n$ verschiedene (reelle) Nullstellen λ_i , $i = 1, \dots, k$, gefunden, so müssen wir anschließend für jeden Eigenwert λ_i das zugehörige GLS

$$(A - \lambda_i E)x^i = 0, \quad x^i = \begin{pmatrix} x_1^i \\ \cdots \\ x_n^i \end{pmatrix},$$

lösen, um die zugehörigen Eigenvektoren x^i bestimmen zu können.

Beispiel 17 (a) Wir betrachten

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Dann folgt aus

$$P_2(\lambda) = \det(A - \lambda E) = \begin{vmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2,$$

dass $\lambda_1 = \lambda_2 = \lambda = 1$ gilt. Aus

$$(A - \lambda E)x = (A - 1E)x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ergibt sich $x_2 = 0$ bei frei wählbarem Parameter x_1 . Somit ist z.B.

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

ein normierter Eigenvektor zum Eigenwert $\lambda = 1$.

(b) Für die Matrix

$$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

erhalten wir

$$P_2(\lambda) = \det(B - \lambda E) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1$$

und somit die komplexen Nullstellen

$$\lambda_1 = i, \quad \lambda_2 = -i.$$

(In diesem Fall kann man die Eigenvektoren nur in

$$\mathbb{C} \cong \mathbb{R}^2$$

bestimmen. Man würde dann die normierten konjugiert komplexen Eigenvektoren

$$z = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad w = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

erhalten.)

Da wir im weiteren nur an **reellen** Eigenwerten interessiert sind, können wir also nur eine *spezielle Klasse* von Matrizen betrachten.

Bei unseren späteren Anwendungen sind wir insbesondere an Eigenwertproblemen für *reelle symmetrische* (n, n) -Matrizen A interessiert, d.h. es gilt

$$A = A^T.$$

Für derartige Matrizen (*Hermitesche Matrizen*) kann man folgende Eigenschaften beweisen, die für unsere späteren Anwendungen entscheidend sind.

Satz 7 *Es sei A eine (reelle) symmetrische (n, n) -Matrix. Das charakteristische Polynom $P_n(\lambda)$ besitze k verschiedene Nullstellen $\lambda_1, \dots, \lambda_k$ mit den Vielfachheiten n_1, \dots, n_k . Dann gilt:*

- (a) *Alle Eigenwerte λ_i sind reell.*
- (b) *Die Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal zueinander.*
- (c) *Hat der Eigenwert λ_i die Vielfachheit n_i , so existieren genau n_i linear unabhängige Eigenvektoren zum Eigenwert λ_i , d.h. algebraische und geometrische Vielfachheit stimmen überein.*
- (d) *Aus dem Fundamentalsatz der Algebra ergibt sich*

$$n_1 + \dots + n_k = n.$$

Bemerkung 5 Da nach den früheren Untersuchungen orthogonale nichttriviale Vektoren stets linear unabhängig sind, folgt aus Satz 7, dass insgesamt n linear unabhängige Eigenvektoren existieren, aus denen man mit Hilfe des Schmidtschen Orthonormalisierungsverfahren eine ONB konstruieren kann. Wir veranschaulichen dieses Verfahren anhand eines Problems, das in der Chemie auftritt.

Beispiel 18 Wir betrachten die symmetrische Matrix

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Sie wird in der Chemie im Rahmen der Hückel-Theorie zur Darstellung der Molekularorbitale und der zugehörigen Energieniveaus für π -Elektronensysteme des Cyclopropenyl-Radikals verwendet.

Wir erhalten für

$$|A - \lambda E| = \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = -\lambda^3 + 3\lambda + 2 = -(\lambda + 1)^2(\lambda - 2) = 0$$

stellen. Wir erhalten hier

$$w^1 \qquad \qquad \qquad w^2 \qquad \qquad \qquad w^3$$

gleich groß gleich groß oder Null verschiedene Größe
 gleiches Vorzeichen verschiedene Vorzeichen verschiedene Vorzeichen

Abschließend wollen wir eine gegebene symmetrische (n, n) -Matrix A , die bezüglich der kanonischen Basis $B_1 = B_K$ gegeben ist, in eine Matrix Λ bezüglich einer ONB B_2 der Eigenvektoren der Matrix A transformieren, siehe auch (25).

In diesem Falle ist die korrespondierende Matrix B_1 wieder die Einheitsmatrix E . Nach Satz 7 existiert für die Matrix A eine ONB aus n Eigenvektoren $\{x^1, \dots, x^n\}$ mit $Ax^i = \lambda_i x^i$. Definieren wir die Basis B_2 durch $B_2 = \{x^1, \dots, x^n\}$, so gilt für die zugehörige Matrix $B_2 = (x^1 \cdots x^n)$. Damit erhalten wir nach Konstruktion

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}$$

Nach (25) ergibt sich mit $\tilde{A} = \Lambda$ die Aussage

$$\Lambda = B_2^{-1} A B_2 = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Andererseits folgt, wie bereits früher gezeigt, aus der Tatsache, dass die Eigenvektoren eine ONB bilden, die Beziehung

$$\begin{aligned} B_2^T B_2 &= \begin{pmatrix} x^1 \\ \cdots \\ x^n \end{pmatrix} (x^1 \cdots x^n) = \begin{pmatrix} \langle x^1, x^1 \rangle \cdots \langle x^1, x^n \rangle \\ \cdots \cdots \cdots \\ \langle x^n, x^1 \rangle \cdots \langle x^n, x^n \rangle \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}, \end{aligned}$$

und damit wieder

$$B_2^T = B_2^{-1}, \quad (29)$$

falls entsprechend unser Festlegung die Spaltenvektoren der Matrix B_2 eine ONB bilden.

Somit erhalten wir

$$\Lambda = B_2^T A B_2.$$

Eine Matrix U , für die (29) erfüllt ist, nennt man eine *orthogonale* Matrix.

Beispiel 19 Wir wenden jetzt unsere Untersuchungen auf das Beispiel 18 an. Dann ist die orthogonale Matrix B_2 , die die Matrix A diagonalisiert, durch

$$B_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & \sqrt{3} & 1 \\ \sqrt{2} & 0 & -2 \\ \sqrt{2} & -\sqrt{3} & 1 \end{pmatrix}$$

gegeben. Man kann sich leicht selbst davon überzeugen, dass folgendes gilt

$$B_2^T B_2 = B_2 B_2^T = E, \quad \Lambda = B_2^T A B_2 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

4.6 Orthogonale Transformationen im \mathbb{R}^2 und \mathbb{R}^3 , Symmetriegruppen

Wir beschäftigen uns jetzt mit speziellen Abbildungen im \mathbb{R}^2 und \mathbb{R}^3 .

Drehung in der Anschauungsebene \mathbb{R}^2

Wir untersuchen nur den Fall der Drehung um den Nullpunkt $0 \in \mathbb{R}^2$ mit einem Drehwinkel $\varphi \in [0, 2\pi)$, der entgegen dem Uhrzeigersinn gemessen wird. Allgemeine Drehungen kann man als Hintereinanderausführung einer Verschiebung in Nullpunktlage, einer Drehung im Nullpunkt und einer anschließenden Rückverschiebung auffassen.

Gegeben sei der *Drehwinkel* φ und ein beliebiger Punkt $x \in \mathbb{R}^2$. Dann seien seine kartesischen und Polarkoordinaten, siehe Vorkurs, Abschnitt 1: Allgemeine Grundlagen,

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = r \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}.$$

Bezeichnen wir mit A_φ die Drehung um den Nullpunkt mit dem Winkel φ , so erhalten wir für $y = A_\varphi x$

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = r \begin{pmatrix} \cos(\alpha + \varphi) \\ \sin(\alpha + \varphi) \end{pmatrix}.$$

Unter Verwendung der Additionstheorem für die sin- und cos-Funktionen folgt daraus

$$\begin{aligned} y_1 = r \cos(\alpha + \varphi) &= r \cos \alpha \cos \varphi - r \sin \alpha \sin \varphi \\ &= x_1 \cos \varphi - x_2 \sin \varphi \end{aligned}$$

und

$$\begin{aligned} y_2 = r \sin(\alpha + \varphi) &= r \cos \alpha \sin \varphi + r \sin \alpha \sin \varphi \\ &= x_1 \sin \varphi + x_2 \cos \varphi, \end{aligned}$$

d.h.

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Die Drehung um den Winkel φ wird also bezüglich der kanonischen Basis durch die $(2, 2)$ -Matrix

$$A_\varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}, \quad x \mapsto y = A_\varphi x, \quad (30)$$

beschrieben. Weiterhin folgt aus $\sin(-x) = -\sin x$ und $\cos(-x) = \cos x$ für die inverse Matrix

$$A_\varphi^{-1} = A_{-\varphi} = \begin{pmatrix} \cos(-\varphi) & -\sin(-\varphi) \\ \sin(-\varphi) & \cos(-\varphi) \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}$$

wegen

$$A_\varphi^T = A_\varphi^{-1},$$

dass die Drehmatrix orthogonal ist. Es ist leicht zu überprüfen, dass

$$A_{\varphi_1 + \varphi_2} = A_{\varphi_1} A_{\varphi_2} = A_{\varphi_2} A_{\varphi_1} = A_{\varphi_2 + \varphi_1}$$

erfüllt ist (kommutative Matrizen).

Jede Drehung ist *Winkel- und Längen-invariant*, d.h. wegen der äquivalenten Definition des Skalarproduktes in \mathbb{R}^2 durch $\langle x, y \rangle = |x||y| \cos \alpha$ erhalten wir

$$\langle A_\varphi x^1, A_\varphi x^2 \rangle = \langle x^1, x^2 \rangle.$$

Somit geht bei der Anwendung von A_φ eine ONB in eine ONB über. Weiterhin ergibt sich

$$\det A_\varphi = \cos^2 \varphi + \sin^2 \varphi = 1.$$

Beispiel 20 Gegeben sei die kanonische Basis

$$B_1 = B_K = \{e^1, e^2\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \cong E.$$

(\cong : entsprechend unserer Interpretation)

Bei einer Drehung um den Winkel φ erhalten wir als neue Basis

$$B_2 = \{e_\varphi^1, e_\varphi^2\} = \left\{ \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} \right\} \cong A_\varphi$$

Die Übergangsmatrix ist also $S = A_\varphi^T = A_{-\varphi}$ (siehe 4.4), da $S := B_2^{-1}B_1 = A_\varphi^{-1}E = A_\varphi^{-1} = A_\varphi^T$. Unter Verwendung von Satz 6 (Koordinatentransformation) ergibt sich für die Koordinaten von $x \in \mathbb{R}^2$ bezüglich der Basis B_1 bzw. B_2

$$x_{B_2} = Sx_{B_1} = A_\varphi^T x_{B_1}.$$

Gegeben sei jetzt der Drehwinkel $\varphi = \pi/4$ und $x = x_{B_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Damit erhalten wir wegen

$$\sin \frac{\pi}{4} = \cos \frac{\pi}{4} = \frac{\sqrt{2}}{2}$$

$$x_{B_2} = A_\varphi^T x_{B_1} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}.$$

Bemerkung 6 (a) Für jede orthogonale Matrix D gilt $\det D = \pm 1$. Das folgt aus

$$1 = \det E = \det(DD^{-1}) = \det(DD^T) = \det D \cdot \det D^T = (\det D)^2.$$

(b) Es sei

$$D = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

eine orthogonale $(2, 2)$ -Matrix mit $\det D = 1$. Dann kann man zeigen, dass dann D eine Drehung im \mathbb{R}^2 beschreibt.

Wir erhalten

$$\begin{aligned} DD^T &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}^2 + a_{12}^2 & a_{11}a_{21} + a_{12}a_{22} \\ a_{11}a_{21} + a_{12}a_{22} & a_{21}^2 + a_{22}^2 \end{pmatrix} = \begin{pmatrix} \boxed{1} & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

und

$$\begin{aligned} D^T D &= \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}^2 + a_{21}^2 & a_{11}a_{12} + a_{21}a_{22} \\ a_{12}a_{11} + a_{22}a_{21} & a_{12}^2 + a_{22}^2 \end{pmatrix} = \begin{pmatrix} \boxed{1} & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Damit erhalten wir folgende Gleichungen

$$1 = a_{11}^2 + a_{12}^2 = a_{11}^2 + a_{21}^2 \quad (31)$$

$$1 = a_{21}^2 + a_{22}^2 = a_{12}^2 + a_{22}^2 \quad (32)$$

sowie

$$0 = a_{11}a_{21} + a_{12}a_{22}. \quad (33)$$

Aus (31) ergibt sich

$$|a_{12}| = |a_{21}| \quad (34)$$

und aus (31), (32) und (33), dass

$$|a_{ij}| \leq 1, \quad i, j = 1, 2. \quad (35)$$

Verwenden wir (31), so existiert genau ein $\varphi \in [0, 2\pi)$ mit

$$a_{11} = \cos \varphi, \quad a_{21} = \sin \varphi.$$

Aus (35) und

$$1 = \det D = a_{11}a_{22} - a_{12}a_{21}$$

folgt

$$a_{12}a_{21} \leq 0$$

und damit wegen (34)

$$a_{12} = -a_{21} = -\sin \varphi.$$

Mit (33) erhalten wir schließlich

$$0 = \cos \varphi \sin \varphi - a_{22} \sin \varphi \implies a_{22} = \cos \varphi.$$

Damit lässt sich die Matrix D darstellen als

$$D = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}, \quad \varphi \in [0, 2\pi).$$

Drehung im Anschauungsraum \mathbb{R}^3

Hierbei erfolgt eine Drehung A_φ mit dem Winkel $0 \leq \varphi < 2\pi$ um eine Gerade, die *Drehachse* genannt wird.

Wir nehmen z.B. an, dass die Drehachse durch die z-Achse gegeben ist, die durch den Vektor

$$e^3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

erzeugt wird. Folglich bleiben alle Vektoren auf der z-Achse bei der Anwendung von A_φ unverändert. Es gilt also

$$A_\varphi e^3 = e^3.$$

Somit ist e^3 ein Eigenvektor der Drehmatrix A_φ zum Eigenwert $\lambda = 1$, und die Drehung wirkt sich nur auf die (x, y) -Ebene aus, die durch die Vektoren

$$e^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

erzeugt wird. Somit hat man in der (x, y) -Ebene die gleiche Situation wie bereits vorher bei der Drehung im \mathbb{R}^2 . Daraus folgt für

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

dass

$$A_\varphi x = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

gilt. Da diese Drehung ebenfalls Winkel- und Längen-Invarianz hat, ergibt sich analog zum \mathbb{R}^2

$$\langle A_\varphi x, A_\varphi y \rangle = \langle x, y \rangle, \det A_\varphi = 1, A_\varphi^{-1} = A_\varphi^T.$$

Insbesondere ist A_φ ebenfalls eine orthogonale Matrix, die eine ONB in eine ONB überführt.

Vektorprodukt in \mathbb{R}^3

Gegeben seien zwei Vektoren

$$a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

bezüglich der kanonischen Basis $B_K = \{e^1, e^2, e^3\}$.

Definition 7 Das Vektorprodukt $a \times b$ zweier Vektoren a und b aus dem \mathbb{R}^3 ist ein Vektor mit folgenden Eigenschaften:

- (i) $a \times b \in \mathbb{R}^3$ steht senkrecht auf a und b ,
- (ii)

$$|a \times b| = |a||b| \sin \varphi, \tag{36}$$

wobei φ den Winkel zwischen a und b mit $0 \leq \varphi \leq \pi$ bezeichnet,

- (iii) die Vektoren a , b und $a \times b$ bilden in dieser Reihenfolge ein Rechtssystem.

Bemerkung 7 (a) Das Vektorprodukt ordnet zwei Vektoren wieder einen Vektor zu.

(b) Das Rechtssystem kann man mittels der „Rechten-Hand-Regel“ bestimmen: der ausgestreckte Daumen zeigt in Richtung von a , der ausgestreckte Zeigefinger in Richtung von b und der nach oben ausgestreckte Mittelfinger in Richtung von $a \times b$. Man beachte aber, dass eine Rechtssystem **nicht** aus orthogonalen Vektoren bestehen muss (siehe Spatprodukt, später)! (c)

Es gelte jetzt $a \neq 0$, $b \neq 0$ und $a \neq \lambda b$. Somit sind a und b in \mathbb{R}^3 linear unabhängig. Betrachtet man das von a und b gebildete Parallelogramm, so entspricht der Betrag des Vektorproduktes (oder *äußeren Produktes*) der Vektoren a und b dem Flächeninhalt dieses Parallelogramms.

(d) Wir erinnern daran, dass beim Skalarprodukt (oder *inneren Produkt*) den zwei Vektoren a und b die skalare Größe $\langle a, b \rangle$ zugeordnet. Es lässt sich zeigen, dass das in Definition 1(d) eingeführte Skalarprodukt

$$\langle a, b \rangle = a_1 b_1 \langle e_1, e_1 \rangle + a_2 b_2 \langle e_2, e_2 \rangle + a_3 b_3 \langle e_3, e_3 \rangle$$

sich auch folgendermaßen berechnen lässt

$$\langle a, b \rangle = a_1 b_1 + a_2 b_2 + a_3 b_3 = |a||b| \cos \varphi. \quad (37)$$

Beispiel 21 Das Vektorprodukt wird z.B. bei der Berechnung des Drehmomentes \vec{M} , gegeben durch

$$\vec{M} = \vec{r} \times \vec{K},$$

sowie des Drehimpulses \vec{l} bezüglich des Nullpunktes

$$\vec{l} = \vec{r} \times \vec{p} = \vec{r} \times m\vec{v}$$

verwendet.

Kraft \vec{K} im Punkt B , $\vec{r} = \overrightarrow{AB}$

bzw. Masseteilchen mit Masse m , Geschwindigkeit \vec{v} und Impuls $\vec{p} = m\vec{v}$ im Punkte P , $\vec{r} = \overrightarrow{OP}$

Wir wollen jetzt wie im Falle des Skalarproduktes eine *Koordinatendarstellung* des Vektorproduktes angeben. Aus Definition 7 folgt unmittelbar

$$e^i \times e^i = 0, \quad i = 1, 2, 3,$$

$$e^1 \times e^2 = e^3, \quad e^2 \times e^3 = e^1, \quad e^3 \times e^1 = e^2.$$

Weiterhin gilt für $a, b, c \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$

$$a \times b = -(b \times a),$$

$$a \times (b + c) = a \times b + a \times c,$$

$$a \times (\lambda b) = \lambda(a \times b).$$

Unter Verwendung dieser Rechenregeln kann man folgende Aussage herleiten. Dabei sollen die Vektoren $a, b \in \mathbb{R}^3$ bezüglich der kanonischen Basis die Darstellung

$$a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

haben.

Satz 8 Die Koordinatendarstellung des Vektorproduktes $a \times b$ lautet

$$a \times b = (a_2b_3 - a_3b_2)e^1 + (a_3b_1 - a_1b_3)e^2 + (a_1b_2 - a_2b_1)e^3. \quad (38)$$

Man kann insbesondere folgende Merkmregel verwenden, die formal durch

$$a \times b \simeq \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ e^1 & e^2 & e^3 \end{vmatrix}$$

gegeben ist. Man beachte, dass es sich hierbei jedoch nicht um eine Determinante handelt.

Spatprodukt im \mathbb{R}^3

Das Spatprodukt (oder *gemischtes Produkt*) der drei Vektoren $a, b, c \in \mathbb{R}^3$, bezeichnet mit $\langle a \times b, c \rangle$, ist die reelle Zahl, die dem Volumen des von den drei Vektoren aufgespannten Parallelepipeds (Paralleleflachs) ist und positiv ist, falls a, b und c ein Rechtssystem bilden. Im anderen Fall erhält das Produkt ein negatives Vorzeichen. Verwenden wir Satz 8 und die Definition des Skalarproduktes, so lässt sich das Spatprodukt auch durch

$$\langle a \times b, c \rangle \equiv \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

berechnen, falls die drei Vektoren bezüglich der kanonischen Basis die Darstellung

$$a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \quad c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

haben. Insbesondere gilt dann unter Verwendung der bekannten Determinantengesetze

$$\langle a \times b, c \rangle = \langle b \times c, a \rangle = \langle c \times a, b \rangle = -\langle a \times c, b \rangle = -\langle c \times b, a \rangle = -\langle b \times a, c \rangle.$$

Unter Verwendung dieser Resultate kann man jetzt (wie im \mathbb{R}^2) analoge Untersuchungen für orthogonale Matrizen A im \mathbb{R}^3 mit $\det A = 1$ vornehmen und zeigen, dass dann die Matrix A eine Drehung im \mathbb{R}^3 beschreibt.

Es sei die orthogonale Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

bezüglich der kanonischen Basis $B_1 = B_K = \{e^1, e^2, e^3\}$ in \mathbb{R}^3 mit $\det A = 1$ gegeben. Unser Ziel besteht darin, aus den Koeffizienten von A die Drehachse und den Drehwinkel zu erhalten.

Als erstes bestimmen wir die Drehachse, d.h. einen nichttrivialen Eigenvektor x^3 zum Eigenwert $\lambda = 1$ (oben e^3). Die Existenz folgt unter Verwendung der Orthogonalität von A und Determinanteneigenschaften (Lemma 1 im Abschnitt 4.3)

$$\begin{aligned} \det(A - 1E) &= \det(A - E) = \det(A(E - A^{-1})) = \det(A(E - A^T)) = \det A \cdot \det(E - A^T) \\ &= 1 \cdot (-1)^3 \det(A^T - E) = -\det(A - E)^T = -\det(A - E), \end{aligned}$$

d.h. $\det(A - 1E) = 0$. Somit ist $\lambda = 1$ ein Eigenwert von A , und es existiert ein nichttrivialer Vektor $x^3 \in \mathbb{R}^3$ mit $|x^3| = 1$.

Mittels Schmidtschen Orthonormalisierungsverfahren bilden wir eine ONB $B_2 = \{x^1, x^2, x^3\}$. Verwendet man die Transformationsformeln für Matrizen aus Abschnitt 4.4, so erhalten wir mit $S^{-1} = S^T = (x^1 \ x^2 \ x^3) \cong B_2$ aus A die neue Matrix $\tilde{A} = SAS^T$ bezüglich der ONB B_2 mit

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} & 0 \\ \tilde{a}_{21} & \tilde{a}_{22} & 0 \\ 0 & 0 & \boxed{1} \end{pmatrix},$$

$\det \tilde{A} = \det A \cdot \det(SS^T) = \det A \cdot \det E = \det A = 1$ und

$$\tilde{A}\tilde{A}^T = SAS^T \circ (SAS^T)^T = SAS^T \circ (S^T)^T A^T S^T = E.$$

Jetzt können wir auf die (2,2)-Untermatrix von \tilde{A} , die somit ebenfalls orthogonal ist und deren Determinante gleich 1 ist, dieselben Überlegungen wie oben anwenden und erhalten nach

geeigneter orthogonaler Basistransformation die neue Matrix

$$\hat{A} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (39)$$

Wir bestimmen jetzt den Drehwinkel φ mit $-\pi \leq \varphi < \pi$. Man kann zeigen, dass

$$\cos \varphi = \frac{1}{2}(a_{11} + a_{22} + a_{33} - 1) \quad (40)$$

gilt.

Das folgt aus der Tatsache, dass bei orthogonalen Basistransformationen die sogenannte *Spur* einer Matrix, d.h. die Summe der Elemente der Hauptdiagonale, invariant ist. Somit erhalten wir aus (39) die Gleichung $2 \cos \varphi + 1 = a_{11} + a_{22} + a_{33}$.

Wir müssen jetzt nur noch das Vorzeichen von φ bestimmen. Dazu wählen wir $x \in \mathbb{R}^3$ mit $x \times x^3 \neq 0$. Dann lässt sich zeigen, dass φ das gleiche Vorzeichen wie das Spatprodukt $\langle x \times Ax, x^3 \rangle$ besitzen muss. Gilt z.B. $\langle x \times Ax, x^3 \rangle > 0$, so bilden $\{x, Ax, x^3\}$ ein Rechtssystem (Drehwinkel $\varphi > 0$ entspricht rechtsherum entsprechend x^3 : *Rechtsgewinde*).

Wir behandeln nun den umgekehrten Fall, d.h. die Drehachse und der Drehwinkel φ seien gegeben, und wir suchen die Drehmatrix A_φ bezüglich der kanonischen Basis $B = B_K = \{e^1, e^2, e^3\}$. Dazu wählen wir zuerst den Vektor

$$x^3 = (\xi_3, \eta_3, \zeta_3)$$

mit $\langle x^3, x^3 \rangle = 1$, der die vorgegebene Drehachse beschreibt. Anschließend wählen wir x^1 und x^2 so, dass $B' = \{x^1, x^2, x^3\}$ ebenfalls eine ONB und ein Rechtssystem in \mathbb{R}^3 bilden. Dazu verwenden wir das Schmidtsche Orthonormalisierungsverfahren. Bezüglich B' erhalten wir

$$\tilde{A}_\varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Verwenden wir jetzt die Ergebnisse über orthogonale Basistransformationen von Matrizen wie im Abschnitt 4.4, so erhalten wir mit $S^{-1} = S^T = B^{-1}B' = EB' = B' = (x^1 \ x^2 \ x^3)$

$$\tilde{A}_\varphi = SA_\varphi S^T.$$

Damit gilt für die gesuchte Matrix

$$A_\varphi = S^T \tilde{A}_\varphi S = B' \tilde{A}_\varphi B'^T.$$

Beispiel 22 Gegeben sei die Drehachse durch den Vektor

$$x = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

und der Drehwinkel $\varphi = \pi/6$. Dann erhalten wir als orthonormierte Vektoren zunächst

$$x^3 = \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ -\frac{1}{2}\sqrt{2} \\ 0 \end{pmatrix}, \quad x^1 = \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \\ 0 \end{pmatrix}.$$

Verwenden wir die Eigenschaften des Vektorproduktes, so kann man den dritten Vektor durch

$$x^2 = x^3 \times x^1 = \begin{vmatrix} e^1 & e^2 & e^3 \\ \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} & 0 \\ \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0 \end{vmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

definieren (oder auch erraten). Dann bilden x^3 , x^1 und x^2 und damit auch x^1 , x^2 und x^3 ein Rechtssystem. Wir setzen jetzt $B' = \{x^1, x^2, x^3\}$. Wegen

$$\cos \frac{\pi}{6} = \frac{1}{2}\sqrt{3}, \quad \sin \frac{\pi}{6} = \frac{1}{2}$$

und $A_\varphi = S^T \tilde{A}_\varphi S = B' \tilde{A}_\varphi B'^T$ erhalten wir folglich

$$\begin{aligned} A_\varphi &= \begin{pmatrix} \frac{1}{2}\sqrt{2} & 0 & \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} & 0 & -\frac{1}{2}\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2}\sqrt{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} + \frac{1}{4}\sqrt{3} & -\frac{1}{2} + \frac{1}{4}\sqrt{3} & -\frac{1}{4}\sqrt{2} \\ -\frac{1}{2} + \frac{1}{4}\sqrt{3} & \frac{1}{2} + \frac{1}{4}\sqrt{3} & -\frac{1}{4}\sqrt{2} \\ \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} & \frac{1}{2}\sqrt{3} \end{pmatrix}. \end{aligned}$$

Spiegelung in der Ebene

Auch hier nehmen wir o.B.d.A. an, dass die *Spiegelungsgerade* durch den Nullpunkt geht (ansonsten zusätzliche Verschiebungen).

Wir sind jetzt an einer Matrix

$$A_S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

interessiert, die eine Spiegelung an einer vorgegebenen Gerade durch den Nullpunkt beschreibt.

Es müssen also folgende Eigenschaften gelten:

1. $A_S(A_S x) = x$ für alle $x \in \mathbb{R}^2$.
2. $|A_S x| = |x|$ für alle $x \in \mathbb{R}^2$.
3. Es existiert eine Gerade g durch den Nullpunkt mit $A_S x = x$ für alle $x \in g$.
4. Es existiert eine Gerade \tilde{g} durch den Nullpunkt mit $\tilde{g} \perp g$ und $A_S x = -x$ für alle $x \in \tilde{g}$.

Aus 3. bzw. 4. folgt, dass $\lambda = 1$ bzw. $\lambda = -1$ ein Eigenwert von A_S ist!

Es sei die Richtung der Geraden g durch den normierten Einheitsvektor

$$w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

definiert. Dann ist nach der 3. Eigenschaft w ein Eigenvektor zum Eigenwert $\lambda = 1$, und es muss

$$\det(A_S - E) = (a - 1)(d - 1) - bc = ad - d - a + 1 - bc = \det A_S - a - d + 1 = 0$$

gelten. Weiterhin sei die Gerade \tilde{g} durch den normierten Einheitsvektor

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

definiert. Dann ist nach der 4. Eigenschaft v ein Eigenvektor zum Eigenwert $\lambda = -1$, und es muss

$$\det(A_S + E) = (a + 1)(d + 1) - bc = ad + d + a + 1 - bc = \det A_S + a + d + 1 = 0$$

gelten. Wir erhalten damit

$$a + d = 0 \implies d = -a,$$

d.h.

$$\det A_S = -1 \tag{41}$$

und wegen $d = -a$

$$-a^2 - cb = -1 \iff a^2 + cb = 1. \tag{42}$$

Da eine Spiegelung langeninvariant ist, folgt weiter aus

$$A_S \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}, \quad A_S \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b \\ -a \end{pmatrix},$$

dass

$$1 = a^2 + c^2 = a^2 + b^2 \implies b^2 = c^2.$$

Somit haben wir $b = \pm c$. Wir nehmen jetzt an, dass $b = -c$ gilt. Dann erhalten wir wegen $bc = -c^2$ aus der letzten Beziehung und (42)

$$1 = a^2 + c^2 = a^2 - c^2.$$

Also muss entweder $c = b = 0$ erfullt sein oder die Annahme $b = -c$ muss falsch sein. In beiden Fallen gilt somit $b = c$. Damit hat A_S folgende Struktur

$$A_S = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}, \quad \det A_S = -(a^2 + b^2) = -1. \quad (43)$$

Wegen $\det A_S = -1$ existiert S^{-1} . Da $a^2 + b^2 = 1$, erhalten wir wegen

$$A_S A_S^T = A_S A_S = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \begin{pmatrix} a & b \\ b & -a \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ab - ab \\ ab - ab & a^2 + b^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E$$

und der Eindeutigkeit der inversen Matrix die Beziehung (bzw. auch aus der 1. Eigenschaft)

$$A_S^T = A_S^{-1}.$$

Somit wird jede Spiegelung im \mathbb{R}^2 durch eine *orthogonale* Matrix A_S mit $\det A_S = -1$ dargestellt. Falls die orthogonale (2,2)-Matrix A mit $\det A = -1$ gegeben ist, so folgt aus

$$\begin{aligned} \det(A \pm 1E) &= \det(A(E \pm A^T)) = \det A \det(E \pm A^T) = \det A \det(A^T \pm E)(-1)^2 \\ &= \det A \det(A \pm E) = -\det(A \pm E), \end{aligned}$$

dass $\lambda = \pm 1$ Eigenwerte sind. Somit wird durch den normierten Eigenvektor w mit $|w| = 1$ zum Eigenwert $\lambda = 1$ die Spiegelgerade g und durch den normierten Eigenvektor v mit $|v| = 1$ und $v \perp w$ zum Eigenwert $\lambda = -1$ die orthogonale Gerade \tilde{g} beschrieben.

Falls umgekehrt die Spiegelgerade g und die orthogonale Gerade \tilde{g} im \mathbb{R}^2 vorgegeben sind, so wahlt man zunachst eine ONB $\{v, w\}$ mit $v \in \tilde{g}$ und $w \in g$. Somit wird die Spiegelung bzgl. dieser ONB durch die Matrix

$$\tilde{A} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

dargestellt. Es gilt

$$-1 = \det \tilde{A} = \det B^T \det A \det B = \det B^{-1} \det A \det B = -1 = \det A$$

Anschließend führt man, analog zu den Drehungen, eine orthogonale Basistransformation durch, um die Matrix A bzgl. der kanonischen Basis zu erhalten.

Ähnliche Überlegungen gelten auch für Spiegelungen im \mathbb{R}^3 . In diesem Falle hat man eine Spiegelebene durch den Nullpunkt, die durch zwei orthonormierte Vektoren $w^1, w^2 \in \mathbb{R}^3$ aufgespannt wird. Die orthogonale Gerade \tilde{g} wird dann durch einen Vektor v erzeugt, so dass $\{v, w^1, w^2\}$ eine ONB in \mathbb{R}^3 darstellt. Da dann w^1 und w^2 zwei Eigenvektoren zum Eigenwert $\lambda = 1$ und v ein Eigenvektor zum Eigenwert $\mu = -1$ sind, gilt für die Matrix

$$\tilde{A}_S = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

bzgl. dieser Basis. Durch eine orthogonale Basistransformation erhält man die Matrix A_S bzgl. der kanonischen Basis in \mathbb{R}^3 . Somit ergibt sich auch hier, dass eine Matrix A , die eine Spiegelung in \mathbb{R}^3 beschreibt, eine orthogonale Matrix ist ($A^T = A^{-1}$) und für ihre Determinante gilt $\det A = -1$.

Symmetriegruppen im \mathbb{R}^3

Für den Chemiker ist das Gebiet der Molekularsymmetrie von großer Bedeutung. Ein wichtiger Bestandteil ist dabei zunächst die Theorie der Drehsymmetrioperationen (Drehungen von Molekülsymmetrieachsen). Dabei werden alle Drehungen, die ein Molekül in sich überführen, zu einer Menge zusammengefasst, die eine Gruppenstruktur besitzt. Man nennt sie deshalb Drehsymmetriegruppe eines Moleküls. Die Symmetrie-Elemente eines Moleküls sind die Drehachsen. Weitere Symmetrie-Elemente eines Moleküls sind Spiegelungsebenen, Drehspiegelachsen und Inversionszentren sowie die korrespondierenden Symmetrioperationen (Spiegelungen, Drehungen, Drehspiegelungen, Inversionen (Punktspiegelungen)). Die Gesamtheit aller Symmetrioperationen (*Decktransformationen*) bilden dann in der *Molekularorbital-Theorie* die Symmetriegruppe eines Moleküls. Man versteht also unter einer Symmetrioperation (Decktransformation) eines Moleküls im \mathbb{R}^3 eine Abbildung

$$A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

mit den folgenden Eigenschaften:

- (i) Es gibt mindestens einen Punkt $x \in \mathbb{R}^3$ mit $Ax = x$ (Fixpunkt).

- (ii) Die Abbildung A ist langen- und winkelinvariant. Somit bleiben die Bindungslangen und Bindungswinkel gleich.
- (iii) Das Molekulkerngerust (in der Gleichgewichtslage) wird durch die Abbildung A in sich uberfuhrt (d.h. *gleiche* Atomkerne konnen vertauscht werden).

Wir wollen diese Decktransformationen hier am Beispiel des Wassermolekuls H_2O beschreiben. Dabei soll sich das Sauerstoff-Atom O im Koordinatenursprung befinden, und die beiden Wasserstoff-Atome H^1, H^2 sollen in der yz -Ebene symmetrisch zur z -Achse angeordnet seien.

Bezuglich der kanonischen Basis im \mathbb{R}^3 verwenden wir folgende Bezeichnungen fur die verschiedenen Decktransformationen.

- E : identische Abbildung:

$$H^1 \rightarrow H^1, \quad H^2 \rightarrow H^2, \quad O \rightarrow O,$$

gegeben durch

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- $C_2(z)$: Drehung um Winkel $\varphi = \pi$ bzgl. der z -Achse:

$$H^1 \rightarrow H^2, \quad H^2 \rightarrow H^1, \quad O \rightarrow O,$$

ist wegen $\cos \pi = -1$ und $\sin \pi = 0$ gegeben durch

$$C_2(z) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- σ_{xz} : Spiegelung an der xz -Ebene:

$$H^1 \rightarrow H^2, \quad H^2 \rightarrow H^1, \quad O \rightarrow O,$$

gegeben durch

$$\sigma_{xz} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- σ_{yz} : Spiegelung an der yz -Ebene:

$$H^1 \rightarrow H^1, \quad H^2 \rightarrow H^2, \quad O \rightarrow O,$$

gegeben durch

$$\sigma_{yz} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Wir erhalten damit, dass die Menge der Decktransformationen

$$\mathcal{D} = \{E, C_2(z), \sigma_{xz}, \sigma_{yz}\}.$$

bzgl. der Hintereinanderausführung (entspricht Matrizenmultiplikation) eine Gruppe bilden, da die Gruppeneigenschaften erfüllt sind.

0) $a, b \in \mathcal{D} \Rightarrow a \circ b \in \mathcal{D}$ und \circ ist assoziativ.

1) Es existiert ein neutrales Element $e \in \mathcal{D}$ mit

$$a \circ e = e \circ a = a \quad \text{für alle } a \in \mathcal{D}.$$

Hier $e = E$ (identische Abbildung).

2) Es existiert für alle $a \in \mathcal{D}$ ein inverses Element $a^{-1} \in \mathcal{D}$ mit

$$a \circ a^{-1} = a^{-1} \circ a = e.$$

In unserem Fall gilt sogar $a^{-1} = a$, d.h. jede Decktransformation ist zu sich selbst invers.

\mathcal{D} besteht also nur aus *Inversionen*.

Bemerkung 8 (a) Da $a \in \mathcal{D}$ orthogonal ist, folgt automatisch

$$a^{-1} = a^T = a$$

wegen der vorliegenden Matrix-Strukturen.

(b) Falls zusätzlich \circ kommutativ ist, so spricht man von einer *kommutativen* oder *abelschen* Gruppe. Insbesondere ist \mathcal{D} eine abelsche Gruppe.

(c) Die kleinste nicht-kommutative Gruppe von Decktransformationen wird durch das Ammoniakmolekül NH_3 realisiert.

5 Theorie der gewöhnlichen Differentialgleichungen

5.1 Beispiele

Es sei $-\infty < a < b < \infty$ ein endliches Intervall. Die stetige Funktion $y = f(x)$ sei auf $[a, b]$ definiert und in (a, b) n -mal stetig differenzierbar, d.h. die Ableitungen $y', \dots, y^{(n)}$ existieren in (a, b) und sind stetig.

Eine Gleichung zwischen der unabhängigen Variable x , der Funktion $y(x)$ und deren Ableitungen nennt man eine *gewöhnliche Differentialgleichung*. Die *Ordnung* der Differentialgleichung entspricht der Ordnung der höchsten Ableitung, die vorkommt. Damit ist eine gewöhnliche Differentialgleichung n -ter Ordnung in impliziter Form durch

$$F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0, \quad x \in (a, b)$$

und in expliziter Form durch

$$y^{(n)}(x) = G(x, y(x), y'(x), \dots, y^{(n-1)}(x)), \quad x \in (a, b)$$

gegeben. $y = f(x)$ heißt *Lösung* der Differentialgleichung, falls y n -mal stetig differenzierbar ist und die Differentialgleichung in (a, b) erfüllt.

Ist F eine ganze rationale Funktion in y, y', \dots , so kann man von einem *Grad* der Differentialgleichung sprechen. Dieser entspricht der höchsten Potenz, die von y, y', \dots gebildet wird. Insbesondere heißt eine Differentialgleichung der Form

$$\sum_{i=1}^n f_i(x)y^{(i)}(x) + f_0(x)y(x) + g(x) = 0$$

linear, da $y, y', \dots, y^{(n)}$ nur linear (in erster Potenz) vorkommen. Falls $g \equiv 0$, so sprechen wir von einer *homogenen* linearen Differentialgleichung n -ter Ordnung; falls $g \neq 0$, so ist sie *inhomogen*.

Beispiel 1 (a) $(y''(x))^3 + (y'(x))^2 + xy(x) = 0$ ist eine implizite Differentialgleichung zweiter Ordnung und dritten Grades.

(b) $y'''(x)y'(x) = y(x)$ ist eine implizite Differentialgleichung dritter Ordnung und zweiten Grades, und

$$\left(\frac{d^2y}{dx^2}\right)^2 - \frac{dy}{dx} + 2xy\frac{d^2y}{dx^2} + x^2y^2 = 0$$

ist eine implizite Differentialgleichung zweiter Ordnung und zweiten Grades.

(c)

$$y''(x) = x^3y'(x) + y(x)\sin x + e^x$$

ist eine inhomogene lineare Differentialgleichung zweiter Ordnung in expliziter Form.

Zur Bestimmung der *allgemeinen* Lösung einer Differentialgleichung n -ter Ordnung braucht man im allgemeinen n Integrationsschritte, die n beliebige Integrationskonstanten erzeugen. Wir betrachten als Beispiel $y''(x) = 0$. Dann erhalten wir nach zweimaliger Integration zuerst $y'(x) = c_1$ und schließlich $y(x) = c_1x + c_2$. Dabei sind c_1 und c_2 zwei voneinander unabhängige beliebige reelle Zahlen. Wenn durch zusätzliche Bedingungen eine oder mehrere Konstanten festgelegt werden, so erhält man eine spezielle oder *partikuläre* Lösung der Differentialgleichung. Haben wir im obigen Beispiel zusätzlich die beiden *Anfangsbedingungen*

$$y(0) = 1, \quad y'(0) = 2$$

vorgegeben, so erhalten wir aus $y(x) = c_1x + c_2$

$$y'(x) = c_1 \longrightarrow y'(0) = c_1 = 2, \quad y(0) = c_2 = 1,$$

d.h. die partikuläre Lösung lautet $y(x) = 2x + 1$. Wir geben einige Beispiele an, wo Differentialgleichungen auftreten.

Beispiel 2 (a) Bakterienwachstum

Die Wachstumsgeschwindigkeit $\frac{dn}{dt}$ einer Kolonie von Bakterien ist proportional zur Quadratwurzel der Zahl $n(t)$ der zu einem Zeitpunkt t vorhandenen Bakterien. Somit erhalten wir die Differentialgleichung erster Ordnung

$$\frac{dn}{dt} = k\sqrt{n}$$

mit einer Proportionalitätskonstante k . Falls $n \neq 0$, so erhalten wir

$$\frac{dn}{\sqrt{n}} = kdt,$$

und schließlich durch Integration $2\sqrt{n} = kt + c$, d.h. die allgemeine Lösung ist gegeben durch

$$n = n(t) = \frac{1}{4}(kt + c)^2. \quad (1)$$

Wir bemerken, daß auch $n \equiv 0$ eine Lösung der Ausgangsgleichung ist, die wir jedoch bei den äquivalenten Umformungen ausschließen mussten. Sie ist nicht in der allgemeinen Lösung (1) enthalten und wird deshalb als *singuläre* Lösung bezeichnet: Für $n = 0$ und $t = 0$ (keine Bakterien vorhanden) würden wir aus (1) folgern, dass $c = 0$ und damit wegen $n(t) = \frac{1}{4}k^2t^2$ ein quadratisches Wachstum vorliegt, obwohl gar keine Bakterien vorhanden sind! Man muss deshalb immer noch alle singuläre Lösungen einer Differentialgleichung bestimmen, falls solche existieren.

(b) Freier und verzögerter Fall

Die Funktion $x(t)$ beschreibe die Bewegung eines Massepunktes mit Masse $m \neq 0$ längs der x -Achse in Abhängigkeit von der Zeit $t \geq 0$ unter dem Einfluss einer Kraft $F = F(t, x(t), \dot{x}(t))$. Es sei wie üblich $\dot{x}(t) = \frac{dx}{dt}$, $\ddot{x}(t) = \frac{d^2x}{dt^2}$. Dann folgt nach dem Newtonschen Gesetz

$$F = m\ddot{x}(t) = f(t, x(t), \dot{x}(t)).$$

(b1) *freier Fall*

Es gilt $m\ddot{x} = mg$, wobei $g \approx 9,81 \text{ms}^{-2}$ die Erdbeschleunigung ist, oder äquivalent $\ddot{x} = g$. Daraus erhalten wir zunächst als allgemeine Lösung für die Geschwindigkeit v

$$v(t) = \dot{x}(t) = gt + c_1$$

und für die Ortskurve $x(t)$

$$x(t) = \frac{g}{2}t^2 + c_1t + c_0.$$

Geben wir uns zusätzlich als Anfangslage $x(0) = x_0$ und Anfangsgeschwindigkeit $\dot{x}(0) = v_0$ vor, so erhalten wir die partikuläre Lösung $x(t) = \frac{g}{2}t^2 + v_0t + x_0$.

(b2) *verzögerter Fall*

Wir nehmen jetzt an, dass zusätzlich ein Luftwiderstand proportional zur Geschwindigkeit vorliegt. Dann lautet die Differentialgleichung

$$m\ddot{x} = mg - \sigma\dot{x}, \tag{2}$$

wobei $\sigma > 0$ der Widerstandskoeffizient ist. Durch die Substitution $\dot{x} = v$ erhalten wir die inhomogene lineare Differentialgleichung

$$\dot{v} = -\frac{\sigma}{m}v + g. \tag{3}$$

Wir werden im Abschnitt 5.2.3 sehen, dass die Lösung von (3) durch

$$v(t) = c_1 e^{-\frac{\sigma}{m}t} + \frac{mg}{\sigma}$$

gegeben ist. Falls $v(0) = v_0$ ist, so erhalten wir

$$v_0 = c_1 + \frac{mg}{\sigma} \longrightarrow c_1 = v_0 - \frac{mg}{\sigma}$$

und damit die partikuläre Lösung von (3)

$$v(t) = \left(v_0 - \frac{mg}{\sigma}\right)e^{-\frac{\sigma}{m}t} + \frac{mg}{\sigma}.$$

Wegen $\dot{x}(t) = v(t)$ erhalten wir nach nochmaliger Integration

$$x(t) = \left(v_0 - \frac{mg}{\sigma}\right)\left(-\frac{m}{\sigma}\right)e^{-\frac{\sigma}{m}t} + \frac{mg}{\sigma}t + c_2.$$

Falls zusätzlich $x(0) = 0$ gilt, so ist

$$c_2 = \left(v_0 - \frac{mg}{\sigma}\right)\frac{m}{\sigma}.$$

Damit erhalten wir schließlich als partikuläre Lösung von (2)

$$x(t) = \frac{m}{\sigma}\left(v_0 - \frac{mg}{\sigma}\right)\left(1 - e^{-\frac{\sigma}{m}t}\right) + \frac{mg}{\sigma}t.$$

(c) radioaktiver Zerfall

$N(t) > 0$ bezeichne die Anzahl der Teilchen eines radioaktiven Materials zum Zeitpunkt $t \geq 0$.

Wir nehmen jetzt an, dass die Abnahme der Teilchenzahl beim radioaktiven Zerfall proportional zur vorhandenen Teilchenzahl erfolgt. Die den Prozess beschreibende Differentialgleichung lautet dann

$$-\frac{dN}{dt} = \alpha N,$$

wobei α eine Materialkonstante ist. Wir haben also eine homogene lineare Differentialgleichung erster Ordnung. Wegen $N > 0$ erhalten wir $\frac{dN}{N} = -\alpha dt$ und durch Integration $\ln N = -\alpha t + C$, und somit als allgemeine Lösung

$$N = N(t) = e^{-\alpha t + C} = e^C \cdot e^{-\alpha t} = A e^{-\alpha t}.$$

Unter der Anfangsbedingung $N(0) = N_0$ ergibt sich $A = e^C = N_0$. Die partikuläre Lösung lautet also

$$N(t) = N_0 e^{-\alpha t}.$$

Mit $t_{1/2}$ wird die Halbwertszeit bezeichnet, in der die Hälfte des anfangs vorhandenen Materials zerfallen ist, d.h. es gilt

$$\frac{1}{2}N_0 = N(t_{1/2}) = N_0 e^{-\alpha t_{1/2}}.$$

Daraus folgt

$$\frac{1}{2} = e^{-\alpha t_{1/2}}$$

und unter Verwendung der Logarithmus-Funktion

$$-\ln 2 = -\alpha t_{1/2},$$

d.h.

$$\alpha = \frac{\ln 2}{t_{1/2}}.$$

Damit ist man in der Lage, Rückschlüsse über das vorliegende radioaktive Material zu ziehen.

5.2 Differentialgleichungen 1. Ordnung

In diesem Abschnitt beschäftigen wir uns mit einigen Lösungsverfahren für gewöhnliche Differentialgleichungen 1. Ordnung.

5.2.1 Richtungsfelder – geometrische Lösungsmethode

Wir betrachten die explizite Differentialgleichung 1. Ordnung

$$y'(x) = f(x, y(x)). \quad (4)$$

In diesem Fall kann man (4) geometrisch deuten. Die Steigung der Lösungskurve $y = y(x)$ (Anstieg der Tangente) wird in jedem Punkt $(x, y(x))$ durch den Wert von $f(x, y(x))$ gegeben. Verläuft also die Lösungskurve $y(x)$ durch den Punkt (x, y) , so bestimmt der Anstieg $f(x, y)$ lokal den weiteren Verlauf. Ein Wertetripel (x, y, y') definiert somit ein *Linienelement*, d.h. ein Stück der Geraden mit dem vorgegebenen Anstieg $f(x, y)$ im Punkte (x, y) . Das *Richtungsfeld* in der (x, y) -Ebene bezeichnet dann die Gesamtheit der Linienelemente der Differentialgleichung (4). Die Kurven $f(x, y) \equiv c$, $c \in \mathbb{R}$, nennt man die *Isoklinen* der Differentialgleichung. Sie dienen dem systematischen Vorgehen zur graphischen Bestimmung des Richtungsfeldes, in das man dann mit guter Näherung Lösungskurven der Differentialgleichung einzeichnen kann.

Beispiel 3 (a) $y'(x) = x + y(x)$. In diesem Fall sind die Isoklinen die Geraden $x + y = c$.

Unter der Anfangsbedingung $y(0) = 1$ erhalten wir die partikuläre Lösung, die durch den Punkt $(0, 1)$ verläuft. Wir werden in 5.2.3 sehen, dass die allgemeine Lösung dieser inhomogenen linearen Differentialgleichung 1. Ordnung durch $y(x) = Ce^x - (1 + x)$, $C \in \mathbb{R}$, beschrieben wird.

(b) $y'(x) = \sqrt{y}$. In diesem Fall haben wir als Isoklinen $\sqrt{y} = c_1$, d.h. Geraden $y = c_1^2 = c$ mit $y \geq 0$.

Analog zu Beispiel 2 erhält man $y(x) = \frac{1}{4}(x + C)^2$, $C \in \mathbb{R}$. Wegen $y' \geq 0$ gilt $x + C \geq 0$. Weiterhin existiert die singuläre Lösung $y \equiv 0$.

(c) Die Isoklinen der Differentialgleichung $y'(x) = x$ sind gegeben durch die Geraden $x = c$.

Die allgemeine Lösung ist in diesem Fall $y(x) = \frac{x^2}{2} + C$, $C \in \mathbb{R}$.

Bemerkung 1 Das Isoklinenverfahren stellt ein gut anwendbares graphisches Lösungsverfahren für Differentialgleichungen dar (schneller Überblick über den Verlauf der Lösungskurve und

damit der Lösungen der Differentialgleichung, partikuläre Lösungen durch Vorgabe von Anfangsbedingungen können näherungsweise ermittelt werden).

Der Nachteil besteht darin, dass dieses Verfahren nur für Differentialgleichungen 1. Ordnung anwendbar ist und i.a. die Bestimmung der Lösungsfunktion $y = f(x)$ nur näherungsweise möglich ist.

In unserem weiteren Vorgehen sind wir deshalb an Algorithmen zur exakten Bestimmung der Lösung einer Differentialgleichung interessiert. Es zeigt sich, dass dieses Ziel für bestimmte Klassen von gewöhnlichen Differentialgleichungen erreichbar ist.

5.2.2 Trennbare Differentialgleichungen

Wir betrachten jetzt Differentialgleichungen der Form

$$y'(x) = \varphi(x, y(x)) = f(x)g(y). \quad (5)$$

In diesem Falle erhält man durch formales Umstellen

$$\frac{dy}{g(y)} = f(x)dx,$$

d.h. man kann die Variablen in der Gleichung separieren. Deshalb nennt man Differentialgleichungen vom Typ (5) separierbare oder *trennbare* Differentialgleichungen. Den Lösungsalgorithmus kann man formal folgendermaßen beschreiben.

- Man untersuche, ob $g(y) = 0$ eine singuläre Lösung von (5) ist.
- Falls die Bedingung $g(y) \neq 0$ erfüllt ist, führe man eine Trennung der Variablen durch und erhält

$$\frac{dy}{g(y)} = f(x)dx.$$

- Durch Integration erhält man die allgemeine Lösung

$$\int \frac{dy}{g(y)} = \int f(x)dx + \text{const.}$$

- Ist zusätzlich eine Anfangsbedingung durch $y(x_0) = y_0$ vorgegeben, so kann man die Integrationskonstante entsprechend bestimmen.

Satz 1 *Es gelte $-\infty < a < b < \infty$, $-\infty < c < d < \infty$ und $(x_0, y_0) \in (a, b) \times (c, d)$. Es sei $f(x)$ stetig in $[a, b]$ und $g(y)$ stetig differenzierbar in $[c, d]$ mit $g(y) \neq 0$ für alle $y \in [c, d]$. Dann lässt sich die die eindeutig bestimmte Lösung der Differentialgleichung*

$$y'(x) = f(x)g(y), \quad y(x_0) = y_0$$

aus

$$F(x, y) = \int_{y_0}^y \frac{du}{g(u)} - \int_{x_0}^x f(v)dv = 0, \quad x \in (a, b), y \in (c, d)$$

berechnen.

Beweis: Differenzieren wir die letzte Gleichung nach x , so erhalten wir mit der Kettenregel (sowie Ableitung nach der oberen Grenze)

$$\frac{1}{g(y)}y'(x) - f(x) = 0.$$

Da $F(x_0, y_0) = 0$, folgt (lokal) die explizite Bedingung $y_0 = y(x_0)$. ■

Beispiel 4 (a) Spezielle trennbare Differentialgleichungen sind $\frac{dy}{dx} = g(y)$, sowie die homogene lineare Differentialgleichung $\frac{dy}{dx} = f(x)y(x)$ und die inhomogene lineare Differentialgleichung $\frac{dy}{dx} = f(x)$.

(b) Analog kann man die trennbaren Differentialgleichungen

$$\frac{dx}{dy} = f_1(x)g_1(y)$$

bezüglich $x = x(y)$ betrachten.

(c) Wir untersuchen die trennbare Differentialgleichung

$$\frac{dy}{dx} = -\frac{x}{y}, \quad y(1) = 1,$$

d.h. es muss stets $y \neq 0$ gelten, und wir haben $f(x) = -x$ und $g(y) = \frac{1}{y}$. Damit sind die Voraussetzungen des Satzes 1 für $f(x)$, $x \in \mathbb{R}$, und $g(y)$, $y \in (-\infty, 0)$ bzw. $y \in (0, \infty)$ erfüllt. Nach dem obigen Algorithmus erhalten wir aus $ydy = -x dx$ die allgemeine Lösung $\frac{y^2}{2} = c - \frac{x^2}{2}$, d.h. die Kreise

$$x^2 + y^2 = 2c, \quad c \in \mathbb{R}.$$

Unter Berücksichtigung der Anfangsbedingung $y(1) = 1$ ergibt sich $c = 1$, d.h. $x^2 + y^2 = 2$. Da $y(1) = 1 > 0$, folgt hieraus wegen $y \neq 0$ die explizite Darstellung der Lösung (des Ausgangsproblems)

$$y(x) = +\sqrt{2 - x^2}, \quad 0 \leq |x| < \sqrt{2}.$$

(d) Für die trennbare Differentialgleichung

$$y'(x) = \frac{xe^{2x}}{y \cos y}, \quad y(0) = \frac{\pi}{4}$$

erhalten wir $f(x) = xe^{2x}$ und $g(y) = \frac{1}{y \cos y}$. Somit sind die Voraussetzungen von Satz 1 für $x \in \mathbb{R}$ und z.B. für $y \in (0, \pi/2)$ (bzgl. der Anfangsbedingung $y(0) = \pi/4$) erfüllt. Wegen

$$\int y \cos y dy = \int xe^{2x} dx + c$$

erhalten wir durch partielle Integration

$$y \sin y + \cos y = \frac{1}{4}(2x - 1)e^{2x} + c.$$

Die Anfangsbedingung $y(0) = \frac{\pi}{4}$ liefert

$$\frac{\pi}{4} \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} = -\frac{1}{4} + c$$

und somit $c = \frac{1}{4} + \frac{\sqrt{2}}{2}(\frac{\pi}{4} + 1)$. Wir erhalten die partikuläre Lösung (in impliziter Form)

$$y \sin y + \cos y = \frac{1}{4}(2x - 1)e^{2x} + \frac{\sqrt{2}}{2}(\frac{\pi}{4} + 1) + \frac{1}{4}.$$

Eine explizite Darstellung der Lösung (selbst lokal) hat hier wenig Sinn.

(e) Bevölkerungswachstum $P(t)$ in Abhängigkeit von der Zeit t in den Entwicklungsländern

Es wird beschrieben durch die Differentialgleichung

(i) $\frac{dP}{dt} = \alpha P^\beta(t), \alpha > 0, \beta > 1, P(0) = P_0$

bzw.

(ii) $\frac{dP}{dt} = \alpha P(t), \alpha > 0, P(0) = P_0.$

Man kann sich leicht durch Nachrechnen davon überzeugen, dass die Lösung für (i) durch

$$P(t) = \frac{1}{\left(\frac{1}{P_0^{\beta-1}} - \alpha(\beta-1)t\right)^{\frac{1}{\beta-1}}} \quad \text{für } 0 \leq t < T = \frac{1}{\alpha(\beta-1)P_0^{\beta-1}}$$

mit $P(t) \rightarrow \infty$ für $t \rightarrow T$ bzw. für (ii) durch

$$P(t) = P_0 e^{\alpha t}.$$

gegeben ist.

(f) Dosis – Wirkungsfunktion eines Medikamentes

Die Funktion $W(x)$ beschreibe die Wirkung des Medikamentes bei Einnahme von x Einheiten.

Damit gilt die Anfangsbedingung $W(0) = 0$. Für wachsendes x (x hinreichend groß) nähert sich $W(x)$ einer Wirkungssättigung $S > 0$. Somit gilt

$$W(x) = \frac{Sx}{x + A},$$

wobei $A > 0$ eine positive Konstante ist. S ist also die Asymptote für die Funktion $W(x)$. Es folgt, dass $W(x)$ der trennbaren Differentialgleichung

$$\frac{dW}{dx} = \frac{S(x + A) - Sx}{(x + A)^2} = \frac{SA}{(x + A)^2} = \frac{A}{S} \left(\frac{W}{x} \right)^2$$

genügt.

(g) Reaktionskinematik

Wir betrachten eine bimolekulare Reaktion, bei der **ein** Molekül vom Typ A mit **einem** vom Typ B **ein** Molekül vom Typ AB bildet. Zum Zeitpunkt $t = 0$ sind a Moleküle vom Typ A und b vom Typ B vorhanden, wobei $a \neq b$ gelte. $x(t)$ bezeichne die Anzahl der verbrauchten Moleküle in Abhängigkeit von der Zeit t . Für die Reaktionsgeschwindigkeit $\frac{dx}{dt}$ erhalten wir die trennbare Differentialgleichung

$$\frac{dx}{dt} = k(a - x)(b - x), \quad 0 \leq x \leq a, \quad 0 \leq x \leq b,$$

wobei $k > 0$ eine Konstante ist. (Die Reaktionsgeschwindigkeit nimmt also um so mehr ab, je mehr Moleküle verbraucht sind.) Falls $x \neq a$ und $x \neq b$ (sonst sind alle Moleküle eines Typs verbraucht), so gilt

$$\int \frac{dx}{(a - x)(b - x)} = kt + c.$$

Nach Partialbruchzerlegung, siehe Abschnitt 3.2, folgt aus

$$\frac{1}{(a - x)(b - x)} = \frac{A}{a - x} + \frac{B}{b - x}$$

$$1 = A(b - x) + B(a - x).$$

Für $x \rightarrow b$ erhalten wir $1 = B(a - b)$ sowie für $x \rightarrow a$ die Gleichung $1 = A(b - a)$, d.h.

$$B = \frac{1}{a - b}, \quad A = -\frac{1}{a - b}.$$

Somit gilt

$$\frac{1}{(a - x)(b - x)} = -\frac{1}{(a - b)(a - x)} + \frac{1}{(a - b)(b - x)}$$

und damit (man beachte $-x$ im Integranden)

$$\int \frac{dx}{(a - x)(b - x)} = \frac{1}{a - b} \ln(a - x) - \frac{1}{a - b} \ln(b - x)$$

$$= \frac{1}{a - b} \ln \frac{a - x}{b - x} = kt + c.$$

Berücksichtigen wir die Anfangsbedingung $x = 0$ für $t = 0$, so folgt

$$\frac{1}{a-b} \ln a - \frac{1}{a-b} \ln b = c,$$

d.h. es gilt

$$c = \frac{1}{a-b} \ln \frac{a}{b},$$

also

$$kt = \frac{1}{a-b} \ln \frac{a-x}{b-x} - \frac{1}{a-b} \ln \frac{a}{b} = \frac{1}{a-b} \ln \frac{b(a-x)}{a(b-x)}.$$

Stellen wir jetzt nach $x = x(t)$ um, so haben wir wegen

$$\begin{aligned} (a-b)kt &= \ln \frac{b(a-x)}{a(b-x)} \\ e^{(a-b)kt} &= \frac{b(a-x)}{a(b-x)} \\ a-x &= (b-x) \frac{a}{b} e^{(a-b)kt} \\ a &= x + ae^{(a-b)kt} - x \frac{a}{b} e^{(a-b)kt} \\ a(1 - e^{(a-b)kt}) &= x(1 - \frac{a}{b} e^{(a-b)kt}) \end{aligned}$$

schließlich

$$\begin{aligned} x &= \frac{a(1 - e^{(a-b)kt})}{1 - \frac{a}{b} e^{(a-b)kt}} \\ &= \frac{a(e^{(a-b)kt} - 1)}{\frac{1}{b}(ae^{(a-b)kt} - b)}, \end{aligned}$$

d.h.

$$x(t) = ab \frac{e^{(a-b)kt} - 1}{ae^{(a-b)kt} - b}. \quad (6)$$

Für $t \rightarrow \infty$ erhält man die Gesamtheit der verbrauchten Moleküle vom Typ A bzw. B:

Falls $a > b$, so ist $\lim_{t \rightarrow \infty} x(t) = ab \frac{1}{a} = b$

und für $a < b$ folgt $\lim_{t \rightarrow \infty} x(t) = ab \frac{-1}{-b} = a$. Somit wird die Reaktion beendet, wenn sämtliche Moleküle eines Stoffes verbraucht sind.

5.2.3 Lineare Differentialgleichungen

Es sei $-\infty < a < b < \infty$, f und g stetig auf $[a, b]$, und $x_0 \in [a, b]$. Nach Abschnitt 5.1 besitzt eine lineare Differentialgleichung 1. Ordnung die Form

$$y'(x) = f(x)y(x) + g(x), \quad x \in [a, b].$$

Falls $g \equiv 0$, so ist die Differentialgleichung homogen; ansonsten ist sie inhomogen.

Wir beschäftigen uns zunächst mit der Lösung der homogenen linearen Differentialgleichung mit Anfangsbedingung

$$y'(x) = f(x)y(x), \quad y(x_0) = y_0. \quad (7)$$

Satz 2

$$y(x) = y_0 \cdot e^{\int_{x_0}^x f(u)du}, \quad a \leq x \leq b, \quad (8)$$

ist die eindeutig bestimmte Lösung von (7).

Beweis: Nach Beispiel 4(a) ist (7) eine trennbare Differentialgleichung. Da alle Voraussetzungen von Satz 1 erfüllt sind, folgt damit die Existenz und die Unität der Lösung. Andererseits ergibt sich aus (8), dass $y(x_0) = y_0$ gilt. Differenzieren wir (8) nach x , so gilt (Ableitung nach der oberen Grenze)

$$y'(x) = f(x)y_0 \cdot e^{\int_{x_0}^x f(u)du} = f(x)y(x).$$

Wir bemerken, dass die triviale Lösung $y \equiv 0$ enthalten ist (in diesem Fall muss die Anfangsbedingung $y_0 = 0$ sein). ■

Beispiel 5 (a) Die allgemeine Lösung der homogenen linearen Differentialgleichung lautet

$$y_{hom}(x) = c \cdot e^{\int^x f(u)du}, \quad c \in \mathbb{R}.$$

Sie erhält man auch durch formales Vorgehen wie bei trennbaren Differentialgleichungen in 5.2.2.

$$\frac{dy}{dx} = f(x)y(x)$$

mit der singulären Lösung $y \equiv 0$. Falls $y \neq 0$ gilt, erhalten wir dann

$$\int \frac{dy}{y} = \int f(x)dx,$$

d.h.

$$\ln |y(x)| = \int^x f(u)du + \ln |C_1|, \quad C_1 \neq 0,$$

und damit

$$y(x) = \pm C_1 e^{\int^x f(u)du} = K e^{\int^x f(u)du}, \quad K \neq 0.$$

Beachten wir jetzt noch die singuläre Lösung, so folgt

$$y_{hom}(x) = c \cdot e^{\int^x f(u)du}, \quad c \in \mathbb{R}.$$

Weiterhin ist die Lösungsmenge der homogenen linearen Differentialgleichung linear, d.h. falls y_1 und y_2 Lösungen sind, so ist auch

$$y(x) := c_1 y_1(x) + c_2 y_2(x), \quad c_1, c_2 \in \mathbb{R},$$

eine Lösung.

(b) Die allgemeine Lösung von $y'(x) = y(x) \sin x$ lautet

$$y_{hom}(x) = c \cdot e^{\int \sin u \, du} = c \cdot e^{-\cos x}.$$

(c) Wir suchen die Lösung von

$$y'(x) = \frac{1}{x} y(x), \quad y(1) = 2.$$

In diesem Fall ist

$$f(x) = \frac{1}{x}$$

auf $\mathbb{R} \setminus \{0\}$ stetig. Wir suchen also die allgemeine Lösung $y_{hom}(x)$ für $x > 0$ entsprechend der Anfangsbedingung. Es gilt für die allgemeine Lösung wegen $\ln|x| = \ln x$ für $x > 0$

$$y_{hom}(x) = c \cdot e^{\ln x} = cx.$$

Wegen $y(1) = 2$ erhalten wir schließlich $c = 2$ und somit die partikuläre Lösung $y(x) = 2x$.

Wir beschäftigen uns jetzt mit der Lösbarkeit der inhomogenen linearen Differentialgleichung

$$y'(x) = f(x)y(x) + g(x), \tag{9}$$

wobei g die Inhomogenität bezeichnet. Dazu verwenden wir als Methode die *Variation der Konstanten*. Nach Beispiel 5(a) ist die allgemeine Lösung der zugehörigen homogenen linearen Differentialgleichung durch

$$y_{hom}(x) = c \cdot e^{\int f(u) \, du}$$

gegeben. Unser Ansatz für die Lösung von (9) lautet jetzt

$$y(x) = c(x) \cdot e^{\int f(u) \, du}, \tag{10}$$

d.h. wir substituieren die reelle Konstante c durch eine reelle Funktion $c(x)$. Setzen wir (10) in die Differentialgleichung (9) ein, so erhalten wir (Ableitung nach der oberen Grenze)

$$y'(x) = c'(x) \cdot e^{\int f(u) \, du} + \underbrace{c(x) \cdot e^{\int f(u) \, du}}_{y(x)} \cdot f(x).$$

Durch Vergleich mit (9) erhalten wir für $c'(x)$ die Bestimmungsgleichung

$$g(x) = c'(x) \cdot e^{\int^x f(u)du},$$

d.h.

$$c'(x) = g(x) \cdot e^{-\int^x f(u)du}$$

und nach Integration

$$c(x) = c_0 + \int^x g(v) \cdot e^{-\int^v f(u)du} dv. \quad (11)$$

Setzen wir jetzt $c(x)$ in (10) ein, so erhalten wir die *allgemeine* Lösung von (9)

$$y_{inh}(x) = c_0 \cdot e^{\int^x f(u)du} + e^{\int^x f(u)du} \left(\int^x e^{-\int^v f(u)du} \cdot g(v)dv \right) \quad (12)$$

Die Anfangsbedingung $y(x_0) = y_0$ liefert schließlich die partikuläre Lösung

$$y_{part}(x) = y_0 \cdot e^{\int_{x_0}^x f(u)du} + e^{\int_{x_0}^x f(u)du} \left(\int_{x_0}^x e^{-\int^v f(u)du} \cdot g(v)dv \right) \quad (13)$$

Satz 3 *Es seien $f(x)$ und $g(x)$ stetig auf $[a, b]$ und $x_0 \in [a, b]$.*

(a) *Die allgemeine Lösung y_{inh} der inhomogenen linearen Differentialgleichung (9) ist durch (12) gegeben.*

(b) *Die eindeutig bestimmte Lösung der inhomogenen linearen Differentialgleichung (9) mit der Anfangsbedingung $y(x_0) = y_0$ ist (13).*

Bemerkung 2 (a) Die Lösungsstruktur ist analog zu der von linearen Gleichungssystemen, siehe Abschnitt lineare Algebra: Die *allgemeine* Lösung der inhomogenen linearen Differentialgleichung $y_{inh}(x)$ ergibt sich aus der *allgemeinen* Lösung der korrespondierenden homogenen linearen Differentialgleichung $y_{hom}(x)$ und einer *speziellen* Lösung der inhomogenen linearen Differentialgleichung $y_s(x)$:

Es gelte

$$\begin{aligned} y_1'(x) &= f(x)y_1(x) + g(x) \\ y_2'(x) &= f(x)y_2(x) + g(x). \end{aligned}$$

Dann folgt

$$(y_1 - y_2)'(x) = f(x)(y_1 - y_2)(x)$$

(b) Es sei

$$y_{hom}(x) = c \cdot e^{\int^x f(u)du} =: c \cdot h(x), \quad c \in \mathbb{R},$$

die allgemeine Lösung des homogenen Problems. Dann lautet der Ansatz für *eine spezielle* Lösung des zugehörigen inhomogenen Problems

$$y'(x) = f(x)y(x) + g(x)$$

mittels der Methode der Variation der Konstanten

$$y_s(x) = c(x)h(x).$$

Nach der letzten Bemerkung ist es ausreichend, wenn man nur eine spezielle Lösung bestimmt.

Dann folgt wegen $h(x) \neq 0$ und $h'(x) = f(x)h(x)$

$$\begin{aligned} y'_s(x) &= c'(x)h(x) + c(x)h'(x) = c'(x)h(x) + c(x)(f(x)h(x)) \\ &= c'(x)h(x) + f(x)y_s(x) \stackrel{!}{=} f(x)y_s(x) + g(x) \end{aligned}$$

aus

$$c'(x)h(x) = g(x)$$

die Bestimmungsgleichung für $c'(x)$, die durch

$$c'(x) = g(x) \frac{1}{h(x)}$$

gegeben ist, und durch Integration schließlich $c(x)$.

(c) Falls die inhomogene lineare Differentialgleichung durch

$$y'(x) = ay(x) + g(x)$$

gegeben ist, wobei a eine **Konstante** ist und $g(x)$ die Inhomogenität bezeichnet, so existieren spezielle Ansätze, siehe die folgende Tabelle. Dabei hat der Lösungsansatz die gleiche Struktur wie die Inhomogenität. Das Lösungsverfahren besteht darin, mit dem Ansatz in die Differentialgleichung einzugehen und die Koeffizienten durch Koeffizientenvergleich zu bestimmen.

Ansätze für Lösungen der inhomogenen linearen Differentialgleichung erster Ordnung mit konstantem Koeffizienten a und Inhomogenität $g(x)$

Inhomogenität $g(x)$	Lösungsansatz
$P_m(x) = \sum_{i=0}^m b_i x^i$	$Q_m(x) = \sum_{i=0}^m a_i x^i$, falls $a \neq 0$
$[P_m(x)]e^{bx}$	$[Q_m(x)]e^{bx}$, falls $b \neq a$
	$x[Q_m(x)]e^{bx}$, falls $b = a$
$[P_m(x)] \cos bx$	$[Q_m(x)] \cos bx + [\tilde{Q}_m(x)] \sin bx$
$[P_m(x)] \sin bx$	$[Q_m(x)] \cos bx + [\tilde{Q}_m(x)] \sin bx$
$[P_m(x)]e^{cx} \cos bx$	$[Q_m(x)]e^{cx} \cos bx + [\tilde{Q}_m(x)]e^{cx} \sin bx$
$[P_m(x)]e^{cx} \sin bx$	$[Q_m(x)]e^{cx} \cos bx + [\tilde{Q}_m(x)]e^{cx} \sin bx$

Beispiel 6 (a) Wir bestimmen die allgemeine Lösung der inhomogenen linearen Differentialgleichung

$$y'(x) = y(x) \sin x + \sin x,$$

d.h. wir haben

$$f(x) = g(x) = \sin x.$$

1. *Schritt* Nach Beispiel 5(b) ist die allgemeine Lösung der homogenen linearen Differentialgleichung $y'(x) = y(x) \sin x$ gegeben durch

$$y_{hom}(x) = c \cdot e^{-\cos x}.$$

2. *Schritt* Die Variation der Konstanten

$$y_s(x) = c(x)e^{-\cos x}$$

liefert nach Bemerkung 2(b)

$$c'(x) = \sin x \cdot e^{\cos x}$$

und somit

$$c(x) = -e^{\cos x} + c_1$$

gilt. Folglich lautet die allgemeine Lösung der inhomogenen linearen Differentialgleichung

$$y_{inh}(x) = \left(-e^{\cos x} + c_1 \right) e^{-\cos x} = c_1 \cdot e^{-\cos x} - 1.$$

3. *Schritt* Falls zusätzlich die Anfangsbedingung $y(0) = 0$ vorgegeben ist, erhalten wir wegen $0 = c_1 e^{-1} - 1$, dass $c_1 = e$ gilt, und die partikuläre Lösung durch

$$y_{part}(x) = e^{1-\cos x} - 1$$

gegeben ist.

(b) Im obigen Beispiel ist die allgemeine Lösung der homogenen linearen Differentialgleichung durch $y_{hom}(x) = c \cdot e^{-\cos x}$ gegeben. Durch Probieren findet man, dass die inhomogene lineare Differentialgleichung die spezielle Lösung $y_s(x) = -1$ hat. Somit erhalten wir auch diesem Fall die Struktur

$$y_{inh}(x) = y_{hom}(x) + y_s(x).$$

(c) Wir behandeln die Differentialgleichung des Gleichstromkreises $U = RI(t) + L \frac{dI}{dt}$ mit Anfangsbedingung $I(0) = 0$. Damit ist die Differentialgleichung des Stromes I in Abhängigkeit von der Zeit t eine inhomogene lineare Differentialgleichung. Die zugehörige homogene Differentialgleichung ist gegeben durch

$$\frac{dI}{dt} = -\frac{R}{L}I(t),$$

wobei wie üblich R den Ohmschen Widerstand, U die Spannung und L die Induktivität bezeichnet.

Die allgemeine Lösung der homogenen linearen Differentialgleichung ist

$$I(t) = c \cdot e^{-\frac{R}{L}t}.$$

Nach der Tabelle erhalten wir als Ansatz für eine spezielle Lösung der inhomogenen linearen Differentialgleichung $I_s(t) = A$, A Konstante. Einsetzen ergibt wegen

$$U = R \cdot A$$

die Beziehung $A = \frac{U}{R}$. Damit lautet die allgemeine Lösung der inhomogenen linearen Differentialgleichung

$$I(t) = \frac{U}{R} + c \cdot e^{-\frac{R}{L}t}.$$

Beachten wir noch die Anfangsbedingung $I(0) = 0$, so erhalten wir mit $c = -\frac{U}{R}$ schließlich die partikuläre Lösung

$$I(t) = \frac{U}{R} \left(1 - e^{-\frac{R}{L}t} \right).$$

(d) Wir bestimmen jetzt Lösungen mittels der in der Tabelle angegebenen speziellen Ansätze.

(d1) $y'(x) = y(x) + \beta$, $\beta \in \mathbb{R}$. In diesem Fall ist

$$y_{hom}(x) = c \cdot e^x$$

und der Ansatz für die spezielle Lösung lautet

$$y_s(x) = A,$$

wobei A eine Konstante ist. Durch Einsetzen in die Differentialgleichung erhalten wir die Gleichung $0 = A + \beta$, d.h. $A = -\beta$. Damit ergibt sich für die allgemeine Lösung der inhomogenen linearen Differentialgleichung

$$y_{inh}(x) = y_{hom}(x) + y_s(x) = c \cdot e^x - \beta.$$

(d2) Es sei $y'(x) = 3y(x) + e^{2x}$. In diesem Fall ist $3 = a \neq b = 2$. Der Ansatz lautet deshalb

$$y_s(x) = A \cdot e^{2x}.$$

Durch Einsetzen erhalten wir

$$2A \cdot e^{2x} = 3A \cdot e^{2x} + e^{2x},$$

und somit nach Koeffizientenvergleich wegen $2A = 3A + 1$, dass $A = -1$ und folglich

$$y_s(x) = -e^{2x}$$

gilt. Da $y_{hom}(x) = c \cdot e^{3x}$, ist die allgemeine Lösung durch

$$y_{inh}(x) = y_{hom}(x) + y_s(x) = c \cdot e^{3x} - e^{2x}$$

gegeben.

(d3) Wir untersuchen $y'(x) = 3y(x) + e^{3x}$. In diesem Falle gilt $3 = a = b$. Der Ansatz ist somit

$$y_s(x) = Ax \cdot e^{3x}.$$

Durch Einsetzen und Koeffizientenvergleich erhalten wir

$$3Ax \cdot e^{3x} + A \cdot e^{3x} = 3Ax \cdot e^{3x} + e^{3x},$$

d.h. $A = 1$. Somit ist

$$y_s(x) = x \cdot e^{3x}$$

und analog zu (d2)

$$y_{inh}(x) = c \cdot e^{3x} + x \cdot e^{3x}.$$

(e) Man weise nach, dass die im Beispiel 4 angegebene Lösung des verzögerten Falls durch

$$v(t) = c_1 \cdot e^{-\frac{\sigma}{m}t} + \frac{mg}{\sigma}$$

gegeben ist.

(f) Wir zeigen, dass die allgemeine Lösung von Beispiel 3(a) im Abschnitt der Differentialgleichung

$$y'(x) = y(x) + x$$

durch

$$y_{inh}(x) = c \cdot e^x - (1 + x)$$

gegeben ist.

Zunächst gilt

$$y_{hom}(x) = ce^x, \quad c \in \mathbb{R}.$$

Wegen $g(x) = x$ lautet der spezielle Ansatz

$$y_s(x) = Ax + b.$$

Damit folgt

$$y'_s(x) = A \stackrel{!}{=} Ax + b + x = (A + 1)x + b.$$

Durch Koeffizientenvergleich folgt $A = b$, $A + 1 = 0$, d.h.

$$A = b = -1.$$

Somit lautet die spezielle Lösung $y_s(x) = -x - 1$, und die allgemeine Lösung ist

$$y_{inh}(x) = ce^x - (x + 1).$$

Unter der Anfangsbedingung $y(0) = 1$ ergibt sich tatsächlich

$$y_{part}(x) = e^x - (1 + x).$$

5.3 Lineare Differentialgleichungen 2. Ordnung

In diesem Abschnitt beschäftigen wir uns mit einigen Lösungsverfahren für lineare Differentialgleichungen 2. Ordnung, die sich auch auf höhere Ordnungen übertragen lassen.

5.3.1 Lineare Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten

Es sei $I \subset \mathbb{R}$ stets ein **endliches Intervall**. Wir betrachten zunächst lineare Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten, die in der Physik und Chemie eine zentrale Bedeutung besitzen. Ihre allgemeine Form ist

$$y''(x) + ay'(x) + by(x) = f(x), \quad x \in I,$$

wobei a und b reelle Konstanten sind, und die auf dem Intervall I stetige Funktion f die Inhomogenität bezeichnet. Durch derartige Differentialgleichungen werden z.B. in der Mechanik oder Elektrodynamik ungedämpfte, gedämpfte und erzwungene Schwingungen eines harmonischen Oszillators bzw. in der Chemie die Normalschwingungen eines diatomaren Moleküls beschrieben. Wir beschäftigen uns zunächst mit der **homogenen Gleichung**

$$y''(x) + ay'(x) + by(x) = 0, \tag{14}$$

die auf dem endlichen Intervall I definiert sei. Die Gleichung ist linear. Damit ergibt sich, dass

$$c_1 y_1(x) + c_2 y_2(x), \quad c_1, c_2 \in \mathbb{C},$$

eine Lösung von (14) ist, falls y_1, y_2 bereits Lösungen sind. Das folgt aus

$$\begin{aligned} & (c_1 y_1 + c_2 y_2)''(x) + a(c_1 y_1 + c_2 y_2)'(x) + b(c_1 y_1 + c_2 y_2)(x) \\ &= c_1 (y_1''(x) + ay_1'(x) + by_1(x)) + c_2 (y_2''(x) + ay_2'(x) + by_2(x)) \\ &= 0 \end{aligned}$$

für $x \in I$. Somit ist für zwei Lösungen y_1 und y_2 von (14) auch jede (komplexe) Linearkombination der Form $y = c_1 y_1 + c_2 y_2$ mit beliebigen reellen oder komplexen Zahlen c_1 und c_2 eine Lösung ist.

Ohne Beweis geben wir folgenden wichtigen Satz an.

Satz 4 (a) *Aus der Lösungsmenge einer homogenen linearen Differentialgleichung zweiter Ordnung (14) lassen sich stets zwei linear unabhängige Lösungsfunktionen y_1 und y_2 auswählen. Mehr als zwei Lösungen sind jedoch stets linear abhängig.*

(b) Es seien y_1 und y_2 zwei linear unabhängige Lösungen von (14). Dann lässt sich die allgemeine Lösung der linearen Differentialgleichung (14) stets als Linearkombination $y_{hom} = c_1 y_1 + c_2 y_2$ darstellen.

Bemerkung 3 Zwei Lösungen y_1 und y_2 heißen *linear unabhängig*, falls aus

$$0 = c_1 y_1(x) + c_2 y_2(x)$$

für alle $x \in I$ stets folgt, dass $c_1 = c_2 = 0$ gilt. Ansonsten heißen die Lösungen *linear abhängig*.

In Analogie zur Lösungsstruktur bei linearen Gleichungssystemen suchen wir ein ausgezeichnetes System von Lösungen von (14), aus denen man durch Linearkombination jede Lösung darstellen kann. Damit ist folgende Definition naheliegend.

Definition 1 Jedes System aus zwei linear unabhängigen Lösungen der linearen homogenen linearen Differentialgleichung (14) heißt *Fundamentalsystem* oder *Fundamentalebasis* der Differentialgleichung.

Als Kriterium für den Nachweis eines Fundamentalsystems verwendet man folgende Eigenschaft.

Satz 5 Die Funktionen y_1 und y_2 seien Lösungen von (14). Sie bilden ein Fundamentalsystem der Differentialgleichung, falls die sogenannte *Wronski-Determinante* $W(x) \neq 0$ für alle $x \in \bar{I}$ ist. Dabei ist $W(x)$ für $x \in \bar{I}$ definiert durch

$$W(x) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix}.$$

Diese Definition verallgemeinert somit einen bekannten Sachverhalt für linear unabhängige Vektoren. Wir beschäftigen uns jetzt mit dem Lösungsverfahren von (14). Dazu verwenden wir den Exponentialfunktions-Ansatz

$$y(x) = e^{\lambda x}.$$

Hierbei kann der Parameter λ auch eine *komplexe* Zahl sein. Durch Einsetzen in (14) erhalten wir wegen $e^{\lambda x} \neq 0$ folgende Bestimmungsgleichung für λ :

$$\lambda^2 + a\lambda + b = 0. \tag{15}$$

Die Gleichung (15) nennt man die *charakteristische Gleichung* für die Differentialgleichung (14).

Nach dem Fundamentalsatz der Algebra, siehe Abschnitt 2.3, gibt es folgende Möglichkeiten

- es existieren zwei verschiedene reelle Lösungen $\lambda_1 \neq \lambda_2$,

- es existiert eine reelle Doppellösung $\lambda_1 = \lambda_2$,
- es existieren zwei konjugiert komplexe Lösungen $\lambda_1 = \alpha + i\beta$ und $\lambda_2 = \alpha - i\beta$ mit $\beta \neq 0$.

Wir konstruieren jetzt daraus für jeden der Fälle ein Paar reeller Lösungen y_1 und y_2 .

Im *ersten* Fall erhalten wir für die Wronski-Determinante von $y_1 = e^{\lambda_1 x}$ und $y_2 = e^{\lambda_2 x}$ wegen

$$\lambda_1 = -\frac{a}{2} + \sqrt{\frac{a^2}{4} - b} \neq -\frac{a}{2} - \sqrt{\frac{a^2}{4} - b} = \lambda_2$$

$$W(x) = \begin{vmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 \cdot e^{\lambda_1 x} & \lambda_2 \cdot e^{\lambda_2 x} \end{vmatrix} = (\lambda_2 - \lambda_1)e^{(\lambda_1 + \lambda_2)x} \neq 0.$$

Somit hat die allgemeine Lösung in diesem Falle die Gestalt

$$\boxed{y_{\text{hom}}(x) = C_1 \cdot e^{\lambda_1 x} + C_2 \cdot e^{\lambda_2 x}}, \quad C_1, C_2 \in \mathbb{R}.$$

Für den *zweiten* Fall erhalten wir zunächst wegen $\lambda = \lambda_1 = \lambda_2$, d.h.

$$\lambda = -\frac{a}{2}, \quad \lambda^2 = b,$$

nur eine unabhängige Lösung $y_1(x) = e^{\lambda x}$. In Analogie zum Abschnitt 5.2.3 (oder mittels Variation der Konstanten) wählen wir als zweiten Kandidaten für das Fundamentalsystem die Funktion $y_2(x) = x \cdot e^{\lambda x}$. Mit diesem Ansatz erhalten wir wegen

$$y_2'(x) = \lambda x \cdot e^{\lambda x} + e^{\lambda x} = (\lambda x + 1)e^{\lambda x},$$

$$y_2''(x) = \lambda(\lambda x + 1)e^{\lambda x} + \lambda e^{\lambda x} = (\lambda^2 x + 2\lambda)e^{\lambda x}$$

und

$$\lambda = -\frac{a}{2}, \quad \lambda^2 = b$$

durch Einsetzen

$$\begin{aligned} y_2''(x) + ay_2'(x) + by_2(x) &= e^{\lambda x}(\lambda^2 x + 2\lambda + a(\lambda x + 1) + bx) \\ &= e^{\lambda x}(\lambda^2 x + 2\lambda + (-2\lambda)(\lambda x + 1) + \lambda^2 x) \\ &= e^{\lambda x} \cdot 0 = 0, \end{aligned}$$

dass $y_2(x) = x \cdot e^{\lambda x}$ ebenfalls eine Lösung ist. Weiterhin gilt wegen

$$W(x) = \begin{vmatrix} x \cdot e^{\lambda x} & e^{\lambda x} \\ (1 + \lambda x)e^{\lambda x} & \lambda \cdot e^{\lambda x} \end{vmatrix} = -e^{2\lambda x} \neq 0,$$

dass die so gewählten Funktionen y_1 und y_2 tatsächlich ein Fundamentalsystem bilden. Die allgemeine Lösung ist also in diesem Falle durch

$$\boxed{y_{hom}(x) = C_1 \cdot x e^{\lambda x} + C_2 \cdot e^{\lambda x} = (C_1 x + C_2) e^{\lambda x}}, \quad C_1, C_2 \in \mathbb{R},$$

gegeben.

Im *dritten* Fall haben wir die beiden komplexen Lösungen $\lambda_1 = \alpha + i\beta$ und $\lambda_2 = \alpha - i\beta$ mit

$$\alpha = -\frac{a}{2}, \quad \beta = \sqrt{b - \frac{a^2}{4}} \neq 0.$$

Wir machen zunächst den komplexen Ansatz

$$z_1(x) = e^{\lambda_1 x} = e^{\alpha x} (\cos \beta x + i \sin \beta x)$$

und

$$z_2(x) = e^{\lambda_2 x} = e^{\alpha x} (\cos \beta x - i \sin \beta x).$$

Dabei verwendeten wir die *Eulersche Formel* (siehe Abschnitt 1: Komplexe Zahlen)

$$e^{i\beta x} = \cos \beta x + i \sin \beta x.$$

Somit erhalten wir zwei konjugiert komplexe Funktionen $z_1(x)$, $z_2(x) = \overline{z_1(x)}$. Wegen

$$\frac{z + \bar{z}}{2} = \operatorname{Re} z, \quad \frac{z - \bar{z}}{2i} = \operatorname{Im} z$$

(siehe Abschnitt 1: Komplexe Zahlen) folgt

$$\begin{aligned} \frac{1}{2}(e^{\lambda_1 x} + e^{\lambda_2 x}) &= \frac{1}{2}e^{\alpha x} (\cos \beta x + i \sin \beta x + \cos \beta x - i \sin \beta x) = e^{\alpha x} \cos \beta x, \\ \frac{1}{2i}(e^{\lambda_1 x} - e^{\lambda_2 x}) &= \frac{1}{2i}e^{\alpha x} (\cos \beta x + i \sin \beta x - \cos \beta x + i \sin \beta x) = e^{\alpha x} \sin \beta x, \end{aligned}$$

und daraus können wir die zwei reelle Lösungen

$$y_1(x) = e^{\alpha x} \cos \beta x, \quad y_2(x) = e^{\alpha x} \sin \beta x$$

(selbst überprüfen) konstruieren. Diese bilden ebenfalls ein Fundamentalsystem. Es gilt

$$W(x) = \begin{vmatrix} e^{\alpha x} \cos \beta x & e^{\alpha x} \sin \beta x \\ e^{\alpha x} (\alpha \cos \beta x - \beta \sin \beta x) & e^{\alpha x} (\alpha \sin \beta x + \beta \cos \beta x) \end{vmatrix}.$$

Daraus folgt aber

$$W(x) = e^{2\alpha x} \beta (\sin^2 \beta x + \cos^2 \beta x) = \beta \cdot e^{2\alpha x} \neq 0.$$

Die allgemeine Lösung lautet also

$$\boxed{y_{hom}(x) = e^{\alpha x} (C_1 \cos \beta x + C_2 \sin \beta x)}, \quad C_1, C_2 \in \mathbb{R}.$$

Beispiel 7 (a) Man bestimme die allgemeine Lösung von

$$y''(x) + 4y'(x) + a_0y(x) = 0 \quad (16)$$

für die Werte $a_0 = 3, 4, 5$.

Die charakteristische Gleichung lautet

$$\lambda^2 + 4\lambda + a_0 = 0.$$

(a1) Für $a_0 = 3$ folgt

$$\lambda^2 + 4\lambda + 3 = 0.$$

In diesem Fall erhalten wir $\lambda_1 = -1$ und $\lambda_2 = -3$. Die allgemeine Lösung von (16) ist somit

$$y_{hom}(x) = C_1e^{-x} + C_2e^{-3x}, \quad C_1, C_2 \in \mathbb{R}.$$

(a2) Für $a_0 = 4$ erhalten wir wegen

$$\lambda^2 + 4\lambda + 4 = 0$$

die doppelte reelle Nullstelle $\lambda_1 = \lambda_2 = -2$. Die allgemeine Lösung von (16) ist folglich

$$y_{hom}(x) = e^{-2x}(C_1x + C_2), \quad C_1, C_2 \in \mathbb{R}.$$

(a3) Schließlich gilt für $a_0 = 5$

$$\lambda^2 + 4\lambda + 5 = 0.$$

Wir erhalten die zwei komplexen Nullstellen $\lambda_1 = -2 + i$ und $\lambda_2 = -2 - i$. Damit hat die allgemeine Lösung von (16) die Form

$$y_{hom}(x) = e^{-2x}(C_1 \sin x + C_2 \cos x), \quad C_1, C_2 \in \mathbb{R}.$$

Eine spezielle Lösung erhält man, wenn die Konstanten C_1 und C_2 fixiert sind. Das kann z.B. durch die Vorgabe zweier Anfangsbedingungen geschehen.

(b) (Der elektrische Schwingkreis) Wie üblich bezeichnen wir mit C die Kapazität, mit L die Induktivität und mit R den Ohmschen Widerstand eines Stromkreises. Für den Strom I in Abhängigkeit von der Zeit t im Schwingkreis gilt dann folgende Differentialgleichung

$$\frac{d^2I}{dt^2} + \frac{R}{L} \frac{dI}{dt} + \frac{1}{LC} I = 0. \quad (17)$$

(b1) Der *ungedämpfte Schwingkreis*, d.h. es gelte $R = 0$. In diesem Falle haben wir die homogene lineare Differentialgleichung

$$\frac{d^2I}{dt^2} + \frac{1}{LC} I = 0.$$

Mit

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

bezeichnet man wie üblich die Kreisfrequenz. Damit erhalten wir

$$\frac{d^2 I}{dt^2} + \omega_0^2 I = 0. \quad (18)$$

Die allgemeine Lösung lautet somit

$$I(t) = A \cos \omega_0 t + B \sin \omega_0 t.$$

Ersetzen wir die Konstanten durch $A = K \sin \varphi$ und $B = K \cos \varphi$ und verwenden wir das Additionstheorem für die Sinus-Funktion, so erhalten wir mit

$$K^2 = A^2 + B^2, \quad \varphi = \arctan \frac{A}{B}, \quad \varphi \in [0, 2\pi),$$

die Darstellung

$$I(t) = K \sin(\omega_0 t + \varphi).$$

Außerdem gilt

$$\frac{dI}{dt} = \omega_0 K \cos(\omega_0 t + \varphi)$$

Die Anfangsbedingungen seien durch $I(0) = 0$ und $\frac{dI}{dt}(0) = \frac{U}{L}$ gegeben.

Somit folgt aus diesen Anfangsbedingungen $\varphi = 0$ (für $\varphi = \pi$ ergibt sich $\sin(x + \pi) = -\sin x$) und $K = \frac{U}{\omega_0 L}$. Unsere partikuläre Lösung für (17) lautet somit

$$\boxed{I(t) = \frac{U}{\omega_0 L} \sin \omega_0 t}.$$

(b2) Der *gedämpfte Schwingkreis* mit Ohmschen Widerstand $R \neq 0$.

Wir setzen

$$\omega_0 = \frac{1}{\sqrt{LC}} \quad \text{und} \quad \frac{R}{L} = 2\delta.$$

Damit geht (18) über in

$$\frac{d^2 I}{dt^2} + 2\delta \frac{dI}{dt} + \omega_0^2 I = 0. \quad (19)$$

In diesem Fall erhalten wir als Nullstellen für die charakteristische Gleichung

$$\lambda_{1,2} = -\delta \pm \sqrt{\delta^2 - \omega_0^2}.$$

In Abhängigkeit von δ und ω müssen wir jetzt folgende drei Fälle betrachten. Dabei bezeichnet die Zahl

$$D = \frac{\delta}{\omega_0}$$

den Dämpfungsgrad.

α): $\delta < \omega_0$ (kleine Dämpfung $D < 1$), d.h. der Ohmsche Widerstand R ist relativ klein gegenüber der Induktivität L und der Kapazität C

Wir erhalten

$$\lambda_{1,2} = -\delta \pm i\sqrt{\omega_0^2 - \delta^2} = -\delta \pm i\omega_e$$

mit der Eigenfrequenz $\omega_e = \sqrt{\omega_0^2 - \delta^2}$. Somit lautet die allgemeine Lösung von (19)

$$I(t) = e^{-\delta t}(A \cos \omega_e t + B \sin \omega_e t).$$

Man erhält schließlich die gedämpfte Schwingung

$$I(t) = e^{-\delta t} K \sin(\omega_e t + \varphi) = e^{-\delta t} K (\sin \omega_e t \cos \varphi + \cos \omega_e t \sin \varphi),$$

falls $K \cos \varphi = B$ und $K \sin \varphi = A$, d.h. wir setzen $K = \sqrt{A^2 + B^2}$ und $\varphi = \arctan \frac{A}{B}$. Dabei bezeichnet $K e^{-\delta t}$ die „Amplitude“ und

$$\omega_e = \sqrt{\omega_0^2 - \delta^2} = \omega_0 \sqrt{1 - D^2}$$

die Eigenfrequenz. Durch Zurücksubstituieren gilt

$$I(t) = e^{-\frac{R}{2L}t} K \sin \left(\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}} t + \varphi \right).$$

β): $\delta = \omega_0$ (aperiodischer Grenzfall $D = 1$)

Wegen $\lambda_1 = \lambda_2 = -\delta$ erhalten wir die allgemeine Lösung

$$I(t) = (C_1 t + C_2) e^{-\frac{R}{2L}t},$$

die höchstens eine (reelle) Nullstelle t_0 haben kann, falls $C_1 t_0 + C_2 = 0$ erfüllbar ist.

γ): $\delta > \omega_0$ (aperiodische Schwingung: starke Dämpfung $D > 1$)

Wegen $\lambda_1 = -\delta + \sqrt{\delta^2 - \omega_0^2}$ und $\lambda_2 = -\delta - \sqrt{\delta^2 - \omega_0^2}$ erhalten wir mit $\gamma = \sqrt{\delta^2 - \omega_0^2}$ als allgemeine Lösung

$$I(t) = e^{-\delta t}(C_1 e^{\gamma t} + C_2 e^{-\gamma t}).$$

Aus

$$\begin{aligned} C_1 e^{\gamma t} + C_2 e^{-\gamma t} &= A \cosh \gamma t + B \sinh \gamma t \\ &= \frac{A}{2}(e^{\gamma t} + e^{-\gamma t}) + \frac{B}{2}(e^{\gamma t} - e^{-\gamma t}) \end{aligned}$$

folgt

$$C_1 = \frac{A+B}{2}, \quad C_2 = \frac{A-B}{2}.$$

Setzen wir also $C_1 + C_2 = A$ und $C_1 - C_2 = B$, so gilt

$$I(t) = e^{-\delta t} (A \cosh \gamma t + B \sinh \gamma t)$$

und nach Zurücksstitution erhalten wir die allgemeine Lösung von (19) bei starker Dämpfung

$$I(t) = e^{-\frac{R}{2L}t} \left(A \cosh \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}} t + B \sinh \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}} t \right).$$

Auch in diesem Fall existiert höchstens eine (reelle) Nullstelle. Unter Berücksichtigung der Anfangsbedingungen können auch im Fall (b2) die noch vorhandenen Konstanten bestimmt werden.

Weiterhin gilt allgemein folgende Aussage.

Satz 6 *Das Anfangswertproblem der linearen homogene Differentialgleichung 2. Ordnung mit konstanten Koeffizienten (14) mit den Anfangsbedingungen*

$$y(x_0) = y_0, \quad y'(x_0) = y_1, \quad x_0 \in I,$$

besitzt genau eine Lösung $y_{part}(x)$.

Es seien $-\infty < \alpha < \beta < \infty$ zwei reelle Zahlen. Wir untersuchen jetzt das inhomogene Problem mit stetiger Funktion $f(x)$, $x \in [\alpha, \beta]$, d.h. wir haben

$$y''(x) + ay'(x) + by(x) = f(x), \quad x \in [\alpha, \beta]. \quad (20)$$

Auch hier verwenden wir die *Methode der Variation der Konstanten*, um eine Lösung von (20) zu erhalten. Wir nehmen an, dass die allgemeine Lösung des zugehörigen homogenen Problems

$$y''(x) + ay'(x) + by(x) = 0, \quad x \in [\alpha, \beta], \quad (21)$$

durch $y(x) = c_1 y_1(x) + c_2 y_2(x)$ gegeben ist. Dabei sind c_1 und c_2 beliebige reelle Zahlen. Unser Ansatz für die Lösung von (20) lautet somit

$$y(x) = c_1(x) y_1(x) + c_2(x) y_2(x). \quad (22)$$

Zur Bestimmung der unbekanntenen Funktionen $c_1(x)$ und $c_2(x)$ müssen wir zunächst die Ableitungen $y'(x)$ und $y''(x)$ berechnen. Es gilt

$$y'(x) = c_1(x)y_1'(x) + c_2(x)y_2'(x) + c_1'(x)y_1(x) + c_2'(x)y_2(x).$$

Die Idee besteht jetzt darin, die Funktionen $c_1(x)$ und $c_2(x)$ so zu bestimmen, dass

$$c_1'(x)y_1(x) + c_2'(x)y_2(x) = 0 \quad (23)$$

gilt. Damit erhalten wir

$$y'(x) = c_1(x)y_1'(x) + c_2(x)y_2'(x). \quad (24)$$

Analog berechnet man daraus

$$y''(x) = c_1(x)y_1''(x) + c_2(x)y_2''(x) + c_1'(x)y_1'(x) + c_2'(x)y_2'(x). \quad (25)$$

Wir erhalten damit einerseits aus der Differentialgleichung (20)

$$y''(x) + ay'(x) + by(x) = f(x)$$

und andererseits durch Einsetzen von (22), (24) und (25)

$$\begin{aligned} & c_1(x)y_1''(x) + c_2(x)y_2''(x) + c_1'(x)y_1'(x) + c_2'(x)y_2'(x) \\ & + ac_1(x)y_1'(x) + c_2(x)y_2'(x) + bc_1(x)y_1(x) + c_2(x)y_2(x) \\ & = c_1(x)\left(y_1''(x) + ay_1'(x) + by_1(x)\right) + c_2(x)\left(y_2''(x) + ay_2'(x) + by_2(x)\right) \\ & + c_1'(x)y_1'(x) + c_2'(x)y_2'(x) \\ & = f(x). \end{aligned}$$

Da y_1 und y_2 ein Fundamentalsystem bilden und somit Lösungen des korrespondierenden homogenen Problems sind, ergeben die beiden ersten Terme den Wert Null. Somit erhalten wir

$$c_1'(x)y_1'(x) + c_2'(x)y_2'(x) = f(x), \quad x \in [\alpha, \beta], \quad (26)$$

d.h. wir haben insgesamt die zwei Bestimmungsgleichungen (23) und (26) zur Berechnung der unbekanntenen Funktionen $c_1'(x)$ und $c_2'(x)$. Wir erhalten also das Gleichungssystem

$$\begin{pmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{pmatrix} \begin{pmatrix} c_1'(x) \\ c_2'(x) \end{pmatrix} = W(x) \begin{pmatrix} c_1'(x) \\ c_2'(x) \end{pmatrix} = \begin{pmatrix} 0 \\ f(x) \end{pmatrix} \quad (27)$$

Da y_1 und y_2 ein Fundamentalsystem bilden, gilt für die Wronski-Determinante, dass $W(x) \neq 0$ für $x \in [\alpha, \beta]$ ist. Somit hat das GLS (27) eine **eindeutig bestimmte Lösung**

$$c_1'(x), c_2'(x).$$

Eine anschließende Integration liefert uns die beiden gesuchten Funktionen

$$c_1(x), c_2(x).$$

Es gilt auch hier folgende Aussage.

Satz 7 (a) Die allgemeine Lösung $y_{inh}(x)$ der linearen inhomogenen Differentialgleichung 2. Ordnung mit konstanten Koeffizienten (21) hat die Struktur

$$y_{inh}(x) = y_{hom}(x) + y_s(x).$$

Dabei ist $y_{hom}(x) = c_1 y_1(x) + c_2 y_2(x)$ die allgemeine Lösung des zugehörigen homogenen Problems (14), d.h. $\{y_1, y_2\}$ bilden ein Fundamentalsystem für (14), und c_1, c_2 sind beliebig wählbare reelle Zahlen. $y_s(x)$ ist eine spezielle Lösung des inhomogenen Problems (21).

(b) Das Anfangswertproblem der linearen inhomogenen Differentialgleichung 2. Ordnung mit konstanten Koeffizienten (21) mit den beiden Anfangsbedingungen

$$y(x_0) = y_0, \quad y'(x_0) = z_0, \quad x_0 \in [\alpha, \beta],$$

besitzt genau eine Lösung $y_{part}(x)$.

Beispiel 8 Man bestimme die Lösung des Anfangswertproblems

$$y''(x) + y(x) = \cos x, \quad x \in [0, 1], \quad y(0) = \pi, \quad y'(0) = 0.$$

Die allgemeine Lösung des homogenen Problems

$$y''(x) + y(x) = 0$$

ist wegen $\lambda^2 + 1 = 0$ durch

$$y_{hom}(x) = c_1 \cos x + c_2 \sin x$$

gegeben, d.h. wir haben das Fundamentalsystem $\{\cos x, \sin x\}$. Die Methode der Variation der Konstanten ergibt als Ansatz

$$y_s(x) = c_1(x) \cos x + c_2(x) \sin x.$$

Nach Cramerscher Regel erhalten wir unter Verwendung von (27) die beiden Gleichungen

$$c_1'(x) = \frac{-\cos x \sin x}{\cos^2 x + \sin^2 x} = -\cos x \sin x$$

und

$$c_2'(x) = \cos^2 x.$$

Die Integration der Gleichung von $c_1'(x)$ ergibt

$$c_1(x) = - \int \cos x \sin x dx = \frac{\cos^2 x}{2} + c = -\frac{\sin^2 x}{2} + \tilde{c}.$$

Durch partielle Integration folgt aus der Gleichung für $c_2'(x)$, dass

$$c_2(x) = \int \cos^2 x dx = \frac{\sin x \cos x + x}{2} + C$$

gilt. Damit ist eine *spezielle* Lösung des inhomogenen Problems durch

$$y_s(x) = \frac{-\sin^2 x \cos x + \sin^2 x \cos x + x \sin x}{2} = \frac{x \sin x}{2}$$

gegeben. Somit lautet die *allgemeine* Lösung des inhomogenen Problems

$$y_{inh}(x) = \frac{x \sin x}{2} + c_1 \cos x + c_2 \sin x, \quad x \in [0, 1].$$

Aus den Anfangsbedingungen folgt schließlich

$$y(0) = \pi = c_1$$

und

$$y'(0) = \frac{\sin x + x \cos x}{2} - \pi \sin x + c_2 \cos x|_{x=0} = c_2 = 0.$$

Wir erhalten also die *partikuläre* Lösung

$$y_{part}(x) = \frac{x \sin x}{2} + \pi \cos x, \quad x \in [0, 1].$$

In Analogie zu linearen Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten kann man auch hier mit **speziellen Ansätzen** arbeiten, die in ihrer Struktur der jeweils vorliegenden Inhomogenität angepasst sind (siehe Tabelle). Damit kann man sich ebenfalls eine spezielle Lösung $y_s(x)$ der inhomogenen Differentialgleichung beschaffen.

Beispiel 9 Wir suchen eine spezielle Lösung von Beispiel 8, d.h. von

$$y''(x) + y(x) = \cos x, \quad x \in [0, 1].$$

Nach Tabelle gilt $\beta = 1$, $P_0(x) = 1$ und $p(\lambda) = \lambda^2 + 1$. Somit ist $\beta i = i$ eine einfache Nullstelle von p . Der Ansatz lautet wegen $T_0(x) = A \cos x + B \sin x$

$$y_s(x) = x(A \cos x + B \sin x) = Ax \cos x + Bx \sin x.$$

Es gilt weiter

$$y_s'(x) = A \cos x + B \sin x + x(B \cos x - A \sin x) = \cos x(A + Bx) + \sin x(B - Ax)$$

und

$$\begin{aligned}y_s''(x) &= -\sin x(A + Bx) + \cos x(B - Ax) + \cos x(B) + \sin x(-A) \\ &= \cos x(2B - Ax) + \sin x(-2A - Bx).\end{aligned}$$

Somit erhalten wir durch Einsetzen in die Differentialgleichung

$$y_s''(x) + y_s(x) = \cos x(Ax + 2B - Ax) + \sin x(Bx - 2A - Bx) = \cos x.$$

Ein Koeffizientenvergleich ergibt also

$$A = 0, B = \frac{1}{2}.$$

Wir erhalten also wie im Beispiel 8 die Lösung

$$y_s(x) = \frac{x \sin x}{2}.$$

Spezielle Ansätze für Lösungen der inhomogenen linearen Differentialgleichung mit konstanten Koeffizienten a_j und Störfunktion $f(x)$:

$$y^{(n)}(x) + \sum_{j=0}^{n-1} a_j y^{(j)}(x) = f(x)$$

Inhomogenität $f(x)$	Lösungsansatz (p := charakteristisches Polynom)
$P_m(x) = \sum_{i=0}^m b_i x^i$	$Q_m(x) = \sum_{i=0}^m A_i x^i$, falls $p(0) \neq 0$ $x^\nu [Q_m(x)]$, falls 0 ν -fache Nullstelle von p
$[P_m(x)]e^{\alpha x}$	$[Q_m(x)]e^{\alpha x}$, falls $p(\alpha) \neq 0$ $x^\nu [Q_m(x)]e^{\alpha x}$, falls α ν -fache Nullstelle von p
$[P_m(x)] \cos \beta x$, $\beta \neq 0$	$T_m(x) := [Q_m(x)] \cos \beta x + [\tilde{Q}_m(x)] \sin \beta x$, falls $p(i\beta) \neq 0$ $x^\nu [T_m(x)]$, falls $i\beta$ ν -fache Nullstelle von p
$[P_m(x)] \sin \beta x$, $\beta \neq 0$	$T_m(x)$, falls $p(i\beta) \neq 0$ $x^\nu [T_m(x)]$, falls $i\beta$ ν -fache Nullstelle von p
$[P_m(x)]e^{\alpha x} \cos \beta x$, $\beta \neq 0$	$\tilde{T}_m(x) := [T_m(x)]e^{\alpha x}$, falls $p(\alpha + i\beta) \neq 0$ $x^\nu [\tilde{T}_m(x)]$, falls $\alpha + i\beta$ ν -fache Nullstelle von p
$[P_m(x)]e^{\alpha x} \sin \beta x$, $\beta \neq 0$	$\tilde{T}_m(x)$, falls $p(\alpha + i\beta) \neq 0$ $x^\nu [\tilde{T}_m(x)]$, falls $\alpha + i\beta$ ν -fache Nullstelle von p

5.3.2 Lineare Differentialgleichungen 2. Ordnung mit variablen Koeffizienten

In diesem Abschnitt wenden wir uns jetzt der Lösung von Differentialgleichungen der Form

$$y''(x) + a(x)y'(x) + b(x)y(x) = q(x) \quad (28)$$

zu. Dabei nehmen wir an, dass die Funktionen $a(x)$, $b(x)$ und $q(x)$ beliebig oft differenzierbar sind und in einer Umgebung $U(x_0)$ in Potenzreihen entwickelbar sind. Bei der Konstruktion der Lösungen verwenden wir bekannte Eigenschaften von Potenzreihen, siehe Abschnitt 3: Taylorreihen und Potenzreihen.

Differentialgleichungen der Form (28) haben eine große Bedeutung in der Physik und Chemie. Zur Lösung verwendet man als spezielle Methode den sogenannten *Potenzreihenansatz*, d.h. man schreibt die Lösung y in der Form

$$y(x) = \sum_{j=0}^{\infty} a_j (x - x_0)^j.$$

Anschließend versucht man, alle auftretenden reellen Koeffizienten a_j durch Einsetzen von $y(x)$ in die Differentialgleichung (28) zu bestimmen. Man hat also im Gegensatz zu den speziellen Ansätzen im letzten Abschnitt nicht mehr nur endlich viele Koeffizienten zu finden. Unter Verwendung der Eigenschaften von Potenzreihen (siehe Abschnitt 3.2.) kann man folgende Aussage beweisen.

Satz 8 *Die Funktionen $a(x)$, $b(x)$ und $q(x)$ seien auf dem Intervall $(x_0 - R, x_0 + R)$, $R > 0$, als Potenzreihen im Punkt $x = x_0$ entwickelbar. Dann gilt dasselbe für die Lösungen $y(x)$ der Differentialgleichung (28).*

Wir suchen also eine Funktion $y(x) = \sum_{j=0}^{\infty} a_j (x - x_0)^j$, die Lösung der obigen Differentialgleichung (28) sein soll. Wenn wir $y(x)$ in (28) einsetzen, müssen wir dann die Koeffizienten a_j , $j \in \mathbb{N}_0$, derart bestimmen, dass die Funktion $y(x)$ auch tatsächlich der Differentialgleichung (28) genügt. Wir konkretisieren diese Methode durch einige einfache und trotzdem sehr wichtige Beispiele, wobei wir nur das erste Beispiel ausführlich behandeln werden.

Die Airysche Differentialgleichung

$$y''(x) - xy(x) = 0. \quad (29)$$

Satz 8 sichert zunächst, dass alle Lösungen als Potenzreihen mit Konvergenzradius $R = \infty$ in $x = x_0$ entwickelbar sind. Wir wählen $x_0 = 0$ und gehen mit dem Ansatz

$$y(x) = \sum_{j=0}^{\infty} a_j x^j$$

in die Gleichung (29) ein. Wegen

$$y'(x) = \sum_{j=0}^{\infty} (j+1)a_{j+1}x^j, \quad y''(x) = \sum_{j=0}^{\infty} (j+2)(j+1)a_{j+2}x^j \quad (30)$$

folgt

$$\underbrace{\sum_{j=0}^{\infty} (j+1)(j+2)a_{j+2}x^j}_{2a_2 + \sum_{j=1}^{\infty} \dots} - \underbrace{\sum_{j=0}^{\infty} a_j x^{j+1}}_{\sum_{j=1}^{\infty} a_{j-1} x^j} = 0,$$

also

$$2a_2 + \sum_{j=1}^{\infty} [(j+1)(j+2)a_{j+2} - a_{j-1}]x^j = 0. \quad (31)$$

Durch Koeffizientenvergleich erhalten wir das unendliche Gleichungssystem

$$2a_2 = 0, \quad (j+1)(j+2)a_{j+2} - a_{j-1} = 0, \quad j = 1, 2, \dots \quad (32)$$

Dabei bleiben die Koeffizienten a_0 und a_1 durch die Differentialgleichung (30) unbestimmt (freie Parameter: werden durch Anfangsbedingungen fixiert). Wegen $a_2 = 0$ erhalten wir aus (32)

$$a_5 = a_8 = a_{11} = \dots = 0, \quad \text{d.h. } a_{3j-1} = 0 \quad \text{für } j = 1, 2, \dots \quad (33)$$

Weiterhin ergibt sich

$$a_3 = \frac{a_0}{2 \cdot 3}, \quad a_6 = \frac{a_3}{5 \cdot 6} = \frac{a_0}{2 \cdot 3 \cdot 5 \cdot 6} = \frac{a_0}{2 \cdot 5 \cdot 3^2 \cdot 2!} \quad (34)$$

und

$$a_9 = \frac{a_6}{8 \cdot 9} = \frac{a_0}{2 \cdot 5 \cdot 3^2 \cdot 2! \cdot 8 \cdot 9} = \frac{a_0}{2 \cdot 5 \cdot 8 \cdot 3^3 \cdot 3!}, \quad (35)$$

oder allgemein

$$a_{3j} = \frac{a_0}{2 \cdot 5 \cdot 8 \cdot \dots \cdot (3j-1) \cdot 3^j \cdot j!} \quad \text{für } j = 1, 2, \dots \quad (36)$$

Analog erhält man

$$a_{3j+1} = \frac{a_1}{4 \cdot 7 \cdot 10 \cdot \dots \cdot (3j+1) \cdot 3^j \cdot j!} \quad \text{für } j = 1, 2, \dots \quad (37)$$

Damit gilt folgende Aussage.

Satz 9 Die allgemeine Lösung der Airyschen Differentialgleichung (29) (mit freien Konstanten a_0 und a_1) ist für alle $x \in \mathbb{R}$ gegeben durch

$$y(x) = a_0 \left(1 + \sum_{j=1}^{\infty} \frac{x^{3j}}{2 \cdot 5 \cdot 8 \cdot \dots \cdot (3j-1) \cdot 3^j \cdot j!} \right) + a_1 \left(x + \sum_{j=1}^{\infty} \frac{x^{3j+1}}{4 \cdot 7 \cdot 10 \cdot \dots \cdot (3j+1) \cdot 3^j \cdot j!} \right) \quad (38)$$

Bemerkung 4 (a) Die Differentialgleichung ist nach dem englischen Mathematiker Sir George Airy (1801-1892) benannt. Diese Gleichung tritt bei Problemen der Lichtbeugung und in der Theorie der Beugung von Radiowellen an der Erdoberfläche auf.

(b) Durch Vorgabe von zwei Anfangsbedingungen $y(0) = c_0$ und $y'(0) = c_1$ kann man die freien Konstante $a_0 = c_0$ und $a_1 = c_1$ eindeutig bestimmen.

Die Hermitesche Differentialgleichung

$$y''(x) - 2xy'(x) + \lambda y(x) = 0. \quad (39)$$

Hierbei ist $\lambda \in \mathbb{R}$ ein fixierter Parameter, d.h. wir behandeln eine Familie von Differentialgleichungen.

Wir gehen mit dem Ansatz

$$y(x) = \sum_{k=0}^{\infty} c_k x^k$$

in (39) ein und erhalten unter Verwendung von (30)

$$\underbrace{\sum_{k=0}^{\infty} (k+1)(k+2)c_{k+2}x^k}_{2c_2 + \sum_{k=1}^{\infty} \dots} - 2 \underbrace{\sum_{k=0}^{\infty} (k+1)c_{k+1}x^{k+1}}_{\sum_{k=1}^{\infty} kc_k x^k} + \lambda \underbrace{\sum_{k=0}^{\infty} c_k x^k}_{c_0 + \sum_{k=1}^{\infty} \dots} = 0.$$

Damit gilt

$$(2c_2 + \lambda c_0) + \sum_{k=1}^{\infty} [(k+1)(k+2)c_{k+2} - 2kc_k + \lambda c_k] x^k = 0.$$

Somit sind c_0 und c_1 frei wählbar, und die anderen Koeffizienten c_k werden, ausgehend von c_0 und c_1 , rekursiv durch

$$c_{k+2} = \frac{2k - \lambda}{(k+1)(k+2)} c_k, \quad k = 0, 1, \dots, \quad (40)$$

bestimmt. Somit ergibt sich z.B.

$$\begin{aligned} c_2 &= \frac{-\lambda}{1 \cdot 2} c_0 = \frac{-\lambda}{2!} c_0 \\ c_4 &= \frac{4 - \lambda}{4 \cdot 3} c_2 = \frac{-[4 - \lambda]\lambda}{4!} c_0 \\ c_6 &= \frac{8 - \lambda}{6 \cdot 5} c_4 = \frac{-[8 - \lambda][4 - \lambda]\lambda}{6!} c_0 \end{aligned}$$

sowie

$$\begin{aligned} c_3 &= \frac{2 - \lambda}{2 \cdot 3} c_1 = \frac{2 - \lambda}{3!} c_1 \\ c_5 &= \frac{6 - \lambda}{5 \cdot 4} c_3 = \frac{[6 - \lambda][2 - \lambda]}{5!} c_1 \\ c_7 &= \frac{10 - \lambda}{7 \cdot 6} c_5 = \frac{[10 - \lambda][6 - \lambda][2 - \lambda]}{7!} c_1. \end{aligned}$$

Unter Verwendung von Satz 8 erhalten wir damit folgendes Ergebnis.

Satz 10 Die allgemeine Lösung der Hermiteschen Differentialgleichung (39) (mit freiem reellen Parameter λ) ist für alle $x \in \mathbb{R}$ gegeben durch

$$y(x) = c_0 \left(1 - \frac{\lambda}{2!} x^2 - \sum_{k=2}^{\infty} \frac{[4(k-1) - \lambda][4(k-2) - \lambda] \cdots [4 - \lambda] \lambda}{(2k)!} x^{2k} \right) + c_1 \left(x + \sum_{k=1}^{\infty} \frac{[4k - 2 - \lambda][4(k-1) - 2 - \lambda] \cdots [2 - \lambda]}{(2k+1)!} x^{2k+1} \right). \quad (41)$$

Bemerkung 5 (a) Für $c_0 = 0$ erhält man eine ungerade, und für $c_1 = 0$ eine gerade Lösung der Hermiteschen Differentialgleichung. Somit setzt sich die allgemeine Lösung additiv aus einer geraden und einer ungeraden zusammen.

(b) Es sei $\lambda \geq 0$ eine gerade Zahl, d.h. es gelte $\lambda = 2n$, $n \in \mathbb{N}_0$. Dann folgt aus (40), dass $c_{n+2} = c_{n+4} = c_{n+6} = \cdots = 0$. Somit bricht eine der beiden eingeklammerten Reihen in (41) ab: Falls n gerade ist, so ist die erste dieser Reihen ein gerades Polynom vom Grade n .

Falls n ungerade ist, so ist die zweite Reihe ein ungerades Polynom vom Grade n .

Multipliziert man diese Polynome mit

$$(-1)^{\frac{n}{2}} \cdot 2^{\frac{n}{2}} \cdot 3 \cdot 5 \cdots (n-1)$$

bei geradem n und mit

$$(-1)^{\frac{n-1}{2}} \cdot 2^{\frac{n+1}{2}} \cdot 3 \cdot 5 \cdots n,$$

so erhält man die *Hermiteschen Polynome* $H_n(x)$, $n \in \mathbb{N}_0$. Sie beginnen mit

$$H_0(x) = 1, H_1(x) = 2x, H_2(x) = -2 + 4x^2, H_3(x) = -12x + 8x^3, \dots$$

Es lässt sich weiterhin zeigen, dass für die Hermiteschen Polynome folgende Eigenschaften erfüllt sind:

- Für das n -te Hermitesche Polynom $H_n(x)$ gilt

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}.$$

- Für $n \in \mathbb{N}$ erhalten wir

$$H'_n(x) = 2nH_{n-1}(x), H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x).$$

- Die Hermiteschen Polynome sind die eindeutigen Lösungen der Differentialgleichung

$$y''(x) - 2xy'(x) + 2ny(x) = 0, n \in \mathbb{N}_0, \quad (42)$$

mit den Anfangsbedingungen

$$H_n(0) = (-1)^{\frac{n}{2}} \cdot 2^{\frac{n}{2}} \cdot 3 \cdot 5 \cdots (n-1), \quad H'_n(0) = 0, \quad \text{falls } n \text{ gerade} \quad (43)$$

bzw.

$$H_n(0) = 0, \quad H'_n(0) = (-1)^{\frac{n-1}{2}} \cdot 2^{\frac{n+1}{2}} \cdot 3 \cdot 5 \cdots n, \quad \text{falls } n \text{ ungerade.} \quad (44)$$

(c) Die Differentialgleichung (39) ist nach dem französischen Mathematiker Charles Hermite (1822-1901) benannt. Sie tritt z.B. bei der Untersuchung von Molekülschwingungen auf.

Weitere wichtige Differentialgleichungen

Abschließend geben wir noch einige weitere Differentialgleichungen 2. Ordnung mit variablen Koeffizienten an, die man mittels Potenzreihenansatz lösen kann.

- Die Legendresche Differentialgleichung mit reellem Parameter λ

$$(1-x^2)y''(x) - 2xy'(x) + \lambda(\lambda+1)y(x) = 0, \quad |x| < 1, \quad (45)$$

tritt z.B. bei geodätischen Problemen auf. Falls $\lambda = n \in \mathbb{N}_0$ gilt, so sind die allgemeine Lösungen durch die sogenannten *Legendreschen Polynome* $P_n(x)$, $n \in \mathbb{N}_0$, gegeben. Sie beginnen mit

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = -\frac{1}{2} + \frac{3}{2}x^2, \quad P_3(x) = -\frac{3}{2}x + \frac{5}{2}x^3, \dots,$$

Sie sind nach dem französischen Mathematiker Adrien-Marie Legendre (1752-1833) benannt.

- Die Tschebyscheffsche Differentialgleichung mit reellem Parameter λ

$$(1-x^2)y''(x) - xy'(x) + \lambda^2 y(x) = 0, \quad |x| < 1, \quad (46)$$

ist nach dem russischen Mathematiker Pafnutij L. Tschebyscheff (1821-1894) benannt. Für $\lambda = n \in \mathbb{N}_0$ erhält man die sogenannten *Tschebyscheffschen Polynome* $T_n(x)$. Sie beginnen mit

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = -\frac{1}{2} + x^2, \quad T_3(x) = -\frac{3}{4}x + x^3, \dots,$$

und haben eine große Bedeutung bei Approximationsaufgaben.

5.4 Systeme von linearen Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten

In diesem Abschnitt beschäftigen wir uns mit Lösungsverfahren für Systeme von linearen Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten. Sie treten bei der Untersuchung reaktionskinetischer Probleme und im Zusammenhang mit Molekülschwingungen auf und sind deshalb für Chemiker von großem Interesse.

5.4.1 Homogene Systeme

Wir beginnen mit folgendem Einführungsbeispiel. Beim radioaktiven Zerfall eines Materials M_1 entsteht eine Substanz M_2 , die wiederum zerfällt. Im folgendem bezeichnen wir mit t die Zeit und mit m_i die Masse von M_i , $i = 1, 2$. Dann gelte

$$\Delta m_1(t) = \underbrace{m_1(t+h) - m_1(t)}_{<0} = -\alpha h m_1(t), \quad \alpha > 0,$$

sowie

$$\Delta m_2(t) = \underbrace{m_2(t+h) - m_2(t)}_{?} = -\underbrace{\beta h m_2(t)}_{\text{Zerfall}} + \underbrace{\alpha h m_1(t)}_{\text{Zuwachs}}, \quad \beta > 0.$$

Damit erhalten wir für $h \rightarrow 0$ folgendes System von Differentialgleichungen 1. Ordnung

$$\begin{aligned} \dot{m}_1(t) &= -\alpha m_1(t) \\ \dot{m}_2(t) &= \alpha m_1(t) - \beta m_2(t). \end{aligned}$$

Wir beschäftigen uns jetzt mit der Lösbarkeit eines allgemeinen Systems von *zwei* linearen Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten a_{ij} , $i, j = 1, 2$, der Form

$$\begin{aligned} y_1'(x) &= a_{11}y_1(x) + a_{12}y_2(x) \\ y_2'(x) &= a_{21}y_1(x) + a_{22}y_2(x). \end{aligned}$$

Damit erhalten wir als Matrixschreibweise

$$\begin{pmatrix} y_1'(x) \\ y_2'(x) \end{pmatrix} = \underbrace{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}_A \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix} \quad (47)$$

bzw.

$$\vec{y}'(x) = A\vec{y}(x), \quad x \in I,$$

für die gesuchte Funktion $\vec{y}(x) = (y_1(x), y_2(x))$. Hierbei ist A eine reelle $(2, 2)$ -Matrix und I ein endliches reelles Intervall. Wir bemerken, dass die folgenden Resultate auch auf Systeme mit

$n \geq 2$ Gleichungen übertragen werden können.

In Analogie zu den Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten arbeiten wir auch hier mit einem Exponentialfunktions-Ansatz der Form

$$\vec{y}(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} e^{\lambda x}.$$

Durch Einsetzen in (47) erhalten wir, dass $\vec{y}(x)$ genau dann eine nichttriviale Lösung des Systems (47) ist, wenn $\vec{y}(x)$ eine nichttriviale Lösung des Eigenwertproblems

$$\vec{y}'(x) = A\vec{y}(x) = \lambda\vec{y}(x),$$

d.h. von

$$(A - \lambda E)\vec{y}(x) = \vec{0} \tag{48}$$

ist. Nach den Resultaten aus der linearen Algebra (siehe Abschnitt 4.5: Eigenwertprobleme) folgt somit die Bedingung

$$\det(A - \lambda E) = 0.$$

Wir müssen also die Nullstellen der *charakteristischen Gleichung* 2. Grades

$$P_2(\lambda) = \det(A - \lambda E) = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0$$

bestimmen. Nach dem Fundamentalsatz der Algebra (siehe Abschnitt 2.3: Integration gebrochen-rationaler Funktionen) können für $n = 2$ folgende **drei** Fälle bzgl. der Nullstellen $\lambda_1, \lambda_2 \in \mathbb{C}$ vorliegen:

1. Fall: $\lambda_1 \neq \lambda_2$ sind *verschiedene reelle* Nullstellen. In diesem Falle setzen wir

$$\vec{y}(x) = \begin{pmatrix} c_{11} \\ c_{12} \end{pmatrix} e^{\lambda_1 x} + \begin{pmatrix} c_{21} \\ c_{22} \end{pmatrix} e^{\lambda_2 x} = \vec{y}_1(x) + \vec{y}_2(x),$$

wobei man die Koeffizienten

$$\vec{c}_i = \begin{pmatrix} c_{i1} \\ c_{i2} \end{pmatrix}, \quad i = 1, 2,$$

durch Einsetzen in die Differentialgleichung (47) bestimmen kann. Da $e^{\lambda x} \neq 0$ gilt, sind sie nichttriviale Lösungen des homogenen linearen Gleichungssystems

$$(A - \lambda_i E)\vec{c}_i = \vec{0}, \quad i = 1, 2.$$

2. Fall: $\lambda := \lambda_1 = \lambda_2$ ist doppelte reelle Nullstelle. In diesem Falle gehen wir analog wie im Abschnitt 5.3: Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten vor und setzen

$$\vec{y}(x) = \vec{c}(x)e^{\lambda x},$$

wobei

$$\vec{c}(x) = \begin{pmatrix} c_{11} \\ c_{12} \end{pmatrix} x + \begin{pmatrix} c_{21} \\ c_{22} \end{pmatrix}$$

ein Vektor ist, dessen Koordinaten Polynome höchstens ersten Grades in x sind. Wir erhalten also

$$\vec{y}(x) = \begin{pmatrix} c_{11} \\ c_{12} \end{pmatrix} x e^{\lambda x} + \begin{pmatrix} c_{21} \\ c_{22} \end{pmatrix} e^{\lambda x} = \vec{y}_1(x) + \vec{y}_2(x).$$

Auch in diesem Falle erhält man die Koeffizienten c_{ij} , $i, j = 1, 2$, durch Einsetzen in die Differentialgleichung (48).

3. Fall: Wir haben zwei konjugiert komplexe Nullstelle $\lambda_1 = \alpha + i\beta$ und $\lambda_2 = \alpha - i\beta$ mit

$$\beta \neq 0.$$

In diesem Falle verfahren wir analog zum Abschnitt 5.3: Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten. Wir setzen

$$\begin{aligned} \vec{y}(x) &= \begin{pmatrix} e^{\alpha x} (c_{11} \cos \beta x + c_{21} \sin \beta x) \\ e^{\alpha x} (c_{12} \cos \beta x + c_{22} \sin \beta x) \end{pmatrix} = \begin{pmatrix} c_{11} \\ c_{12} \end{pmatrix} e^{\alpha x} \cdot \cos \beta x + \begin{pmatrix} c_{21} \\ c_{22} \end{pmatrix} e^{\alpha x} \cdot \sin \beta x \\ &= \vec{y}_1(x) + \vec{y}_2(x), \end{aligned}$$

wobei man die Koeffizienten

$$\vec{c}_i = \begin{pmatrix} c_{i1} \\ c_{i2} \end{pmatrix}, \quad i = 1, 2,$$

durch Einsetzen in die Differentialgleichung (47) bestimmt.

Allgemein erhält man unter Verwendung der Linearität des Differentialgleichungssystems folgendes Resultat. Man vergleiche insbesondere mit Satz 4.

Satz 11 (a) *Aus der Lösungsmenge des homogenen linearen Systems*

$$\vec{y}'(x) = A\vec{y}(x), \quad x \in I, \tag{49}$$

von Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten lassen sich stets zwei linear unabhängige Lösungen

$$\vec{y}_1(x) = (y_{11}(x), y_{12}(x)), \quad \vec{y}_2(x) = (y_{21}(x), y_{22}(x))$$

auswählen. Mehr als zwei Lösungen sind jedoch stets linear abhängig.

(b) *Es seien $\vec{y}_1(x)$ und $\vec{y}_2(x)$ zwei linear unabhängige Lösungen von (49). Dann lässt sich die allgemeine Lösung der linearen Differentialgleichung (49) stets als Linearkombination*

$$\vec{y}(x) = c_1 \vec{y}_1(x) + c_2 \vec{y}_2(x)$$

darstellen.

Bemerkung 6 Zwei Lösungen $\vec{y}_1(x)$ und $\vec{y}_2(x)$ heißen *linear unabhängig*, falls das homogene lineare Gleichungssystem

$$\vec{0} = c_1 \vec{y}_1(x) + c_2 \vec{y}_2(x)$$

für alle $x \in I$ nur die triviale Lösung $c_1 = c_2 = 0$ hat. Ansonsten heißen die Lösungen *linear abhängig*.

In Analogie zur Lösungsstruktur bei linearen Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten suchen wir auch hier nach einem ausgezeichneten System von Lösungen für (49), aus denen man durch Linearkombination jede Lösung darstellen kann. Damit ist folgende Definition naheliegend.

Definition 2 Jedes System aus zwei linear unabhängigen Lösungen von (49) heißt *Fundamentalsystem* oder *Fundamentalebasis* des Differentialgleichungssystems (49).

Auch hier kann man als Kriterium für den Nachweis eines Fundamentalsystems folgende Eigenschaft verwenden.

Satz 12 Die Funktionen $\vec{y}_1(x) = (y_{11}(x), y_{12}(x))$ und $\vec{y}_2(x) = (y_{21}(x), y_{22}(x))$ seien Lösungen von (49). Sie bilden ein Fundamentalsystem des Differentialgleichungssystems, falls die Wronski-Determinante $W(x) \neq 0$ für alle $x \in \bar{I}$ ist. Dabei ist $W(x)$ für $x \in \bar{I}$ definiert durch

$$W(x) = \begin{vmatrix} y_{11}(x) & y_{21}(x) \\ y_{12}(x) & y_{22}(x) \end{vmatrix}.$$

Wir wenden die Lösungstheorie jetzt auf einige Beispiele an.

Beispiel 10 (a) Wir lösen für $t \in [0, 1]$ das Anfangswertproblem

$$\begin{aligned} \dot{x}(t) &= 4x(t) + y(t) \\ \dot{y}(t) &= -2x(t) + y(t) \end{aligned} \tag{50}$$

mit den Anfangsbedingungen

$$x(0) = 1, y(0) = 0. \tag{51}$$

Erster Schritt Bestimmung der Nullstellen

Wir erhalten

$$P_2(\lambda) = \begin{vmatrix} 4 - \lambda & 1 \\ -2 & 1 - \lambda \end{vmatrix} = \lambda^2 - 5\lambda + 6 = 0.$$

Die Nullstellen sind $\lambda_1 = 3$ und $\lambda_2 = 2$.

Zweiter Schritt Lösungsansatz

Nach den obigen Untersuchungen (1. Fall) erhalten wir

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} A \\ C \end{pmatrix} e^{3t} + \begin{pmatrix} B \\ D \end{pmatrix} e^{2t}. \quad (52)$$

Dritter Schritt Berechnung der reellen Koeffizienten A, B, C und D

Durch Einsetzen in (50) erhalten wir

$$3Ae^{3t} + 2Be^{2t} = (4A + C)e^{3t} + (4B + D)e^{2t}.$$

Ein Koeffizientenvergleich ergibt

$$C = -A, \quad D = -2B.$$

(Ein analoges Resultat liefert auch

$$3Ce^{3t} + 2De^{2t} = (-2A + C)e^{3t} + (-2B + D)e^{2t}.)$$

Die allgemeine Lösung des Systems (50) lautet

$$x_{hom}(t) = Ae^{3t} + Be^{2t}, \quad y_{hom}(t) = -Ae^{3t} - 2Be^{2t}. \quad (53)$$

Wir setzen $A = 1, B = 0$ bzw. $A = 0$ und $B = 1$. Die Lösungen

$$\vec{y}_1(t) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{3t}, \quad \vec{y}_2(t) = \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^{2t}$$

sind wegen

$$W(t) = \begin{vmatrix} 1e^{3t} & 1e^{2t} \\ -1e^{3t} & -2e^{2t} \end{vmatrix} = -2e^{5t} + 1e^{5t} = -e^{5t} \neq 0$$

für alle $t \in \mathbb{R}$ linear unabhängig und bilden ein Fundamentalsystem.

Vierter Schritt Einarbeiten der Anfangsbedingungen

Aus (51) und (53) erhalten wir $A + B = 1$ und $-A - 2B = 0$, d.h. es gilt

$$A = 2, \quad B = -1.$$

Die Lösung des Problems (50), (51) ist

$$x_{part}(t) = 2e^{3t} - e^{2t}, \quad y_{part}(t) = -2e^{3t} + 2e^{2t}.$$

(b) Wir betrachten für $t \in [0, 1]$

$$\begin{aligned} \dot{u}(t) &= 3u(t) - 4v(t) \\ \dot{v}(t) &= u(t) - v(t) \end{aligned} \quad (54)$$

mit den Anfangsbedingungen

$$u(0) = 3, v(0) = 1. \quad (55)$$

Erster Schritt Bestimmung der Nullstellen

Wir erhalten

$$P_2(\lambda) = \begin{vmatrix} 3 - \lambda & -4 \\ 1 & -1 - \lambda \end{vmatrix} = \lambda^2 - 2\lambda + 1 = 0.$$

Die Nullstellen sind $\lambda_1 = \lambda_2 = 1$.

Zweiter Schritt Lösungsansatz

Nach den obigen Untersuchungen (2. Fall) erhalten wir

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} A \\ C \end{pmatrix} e^t + \begin{pmatrix} B \\ D \end{pmatrix} t e^t. \quad (56)$$

Dritter Schritt Berechnung der reellen Koeffizienten A, B, C und D

Durch Einsetzen in (54) erhalten wir

$$(A + Bt)e^t + Be^t = 3(A + Bt)e^t - 4(C + Dt)e^t.$$

Ein Koeffizientenvergleich ergibt

$$C = \frac{A}{2} - \frac{B}{4}, D = \frac{B}{2}.$$

Die allgemeine Lösung des Systems (54) lautet

$$u_{hom}(t) = (A + Bt)e^t, \quad v_{hom}(t) = \left(\frac{A}{2} - \frac{B}{4} + \frac{B}{2}t \right) e^t. \quad (57)$$

Wir setzen $A = 1, B = 0$ bzw. $A = 0$ und $B = 1$. Die Lösungen sind dann

$$\vec{y}_1(t) = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix} e^t, \quad \vec{y}_2(t) = \begin{pmatrix} t \\ -\frac{1}{4} + \frac{1}{2}t \end{pmatrix} e^t$$

und wegen

$$W(t) = \begin{vmatrix} e^t & te^t \\ \frac{1}{2}e^t & (-\frac{1}{4} + \frac{t}{2})e^t \end{vmatrix} = (-\frac{1}{4} + \frac{t}{2})e^{2t} - \frac{t}{2}e^{2t} = -\frac{1}{4}e^{2t} \neq 0$$

für alle $t \in \mathbb{R}$ linear unabhängig und bilden ein Fundamentalsystem.

Vierter Schritt Einarbeiten der Anfangsbedingungen

Aus (55) und (57) erhalten wir $A = 3$ und $A/2 - B/4 = 1$, d.h. es gilt

$$A = 3, B = 2.$$

Die Lösung des Problems (54), (55) ist

$$u_{part}(t) = (3 + 2t)e^t, \quad v_{part}(t) = (1 + t)e^t.$$

(c) Wir untersuchen für $t \in [0, 1]$

$$\begin{aligned} \dot{u}(t) &= 3u(t) + 2v(t) \\ \dot{v}(t) &= -5u(t) + v(t) \end{aligned} \quad (58)$$

mit den Anfangsbedingungen

$$u(0) = v(0) = 2. \quad (59)$$

Erster Schritt Bestimmung der Nullstellen

Wir erhalten

$$P_2(\lambda) = \begin{vmatrix} 3 - \lambda & 2 \\ -5 & 1 - \lambda \end{vmatrix} = \lambda^2 - 4\lambda + 13 = 0.$$

Die Nullstellen $\lambda_1 = 2 + 3i$ und $\lambda_2 = 2 - 3i$ sind konjugiert komplex.

Zweiter Schritt Lösungsansatz

Nach den obigen Untersuchungen (3. Fall) erhalten wir

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} A \\ C \end{pmatrix} e^{2t} \cos 3t + \begin{pmatrix} B \\ D \end{pmatrix} e^{2t} \sin 3t \quad (60)$$

Dritter Schritt Berechnung der Koeffizienten

Durch Einsetzen in (58) erhalten wir

$$e^{2t}[(2A + 3B) \cos 3t + (-3A + 2B) \sin 3t] = e^{2t}[(3A + 2C) \cos 3t + (3B + 2D) \sin 3t].$$

Ein Koeffizientenvergleich ergibt bei $\cos 3t$ bzw. $\sin 3t$

$$C = -\frac{A}{2} + \frac{3}{2}B, \quad D = -\frac{3}{2}A - \frac{B}{2}.$$

Die allgemeine Lösung des Systems (58) lautet

$$u_{hom}(t) = e^{2t}(A \cos 3t + B \sin 3t), \quad v_{hom}(t) = \frac{e^{2t}}{2}[(-A + 3B) \cos 3t - (3A + B) \sin 3t]. \quad (61)$$

Wir setzen $A = 1, B = 0$ bzw. $A = 0$ und $B = 1$. Damit erhalten wir die Lösungen

$$\vec{y}_1(t) = \begin{pmatrix} \cos 3t \\ -\frac{1}{2} \cos 3t - \frac{3}{2} \sin 3t \end{pmatrix} e^{2t}, \quad \vec{y}_2(t) = \begin{pmatrix} \sin 3t \\ \frac{3}{2} \cos 3t - \frac{1}{2} \sin 3t \end{pmatrix} e^{2t}.$$

Sie sind wegen

$$W(t) = \begin{vmatrix} e^{2t} \cos 3t & e^{2t} \sin 3t \\ -\frac{1}{2}e^{2t}(\cos 3t + 3 \sin 3t) & \frac{1}{2}e^{2t}(3 \cos 3t - \sin 3t) \end{vmatrix} = \frac{3}{2}e^{4t} \neq 0$$

für alle $t \in \mathbb{R}$ linear unabhängig und bilden ein Fundamentalsystem.

Vierter Schritt Einarbeiten der Anfangsbedingungen

Aus (59) und (61) erhalten wir $A = 2$ und $-A + 3B = 4$, d.h. es gilt

$$A = 2, B = 2.$$

Die Lösung des Problems (58), (59) ist

$$u_{part}(t) = 2e^{2t}(\cos 3t + \sin 3t), \quad v_{part}(t) = 2e^{2t}(\cos 3t - 2 \sin 3t).$$

(d) Beispiel aus der Reaktionskinetik (Papula, Mathematik für Chemiker, S. 362)

Wir geben ein Beispiel, wo $n = 3$ Gleichungen auftreten. Das Vorgehen ist analog zum früheren Fall $n = 2$, mit entsprechenden Modifikationen für mehrfache Nullstellen.

Bei einer chemischen Reaktion wandeln sich Moleküle vom Typ X zunächst in Moleküle vom Typ Y und anschließend in Moleküle vom Typ Z . Die Molekularkonzentrationen zum Zeitpunkt t seien durch $c_X(t)$, $c_Y(t)$ und $c_Z(t)$ gegeben. Die Reaktion kann dann durch folgendes System von *drei* linearen Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten beschrieben werden.

$$\begin{aligned} c'_X &= -k_1 c_X \\ c'_Y &= k_1 c_X - k_2 c_Y \\ c'_Z &= k_2 c_Y, \end{aligned}$$

wobei für die Konstanten $k_1 \neq k_2$ gilt. In diesem Fall haben wir die drei verschiedenen Eigenwerte

$$\lambda_1 = 0, \lambda_2 = -k_1, \lambda_3 = -k_2.$$

Man zeige, dass die spezielle Lösung des Systems mit den Anfangsbedingungen

$$c_X(0) = a, c_Y(0) = b, c_Z(0) = c$$

durch

$$\begin{aligned} c_X(t) &= ae^{-k_1 t} \\ c_Y(t) &= a \frac{k_1}{k_2 - k_1} e^{-k_1 t} + \frac{b(k_2 - k_1) - ak_1}{k_2 - k_1} e^{-k_2 t} \\ c_Z(t) &= a + b + c - a \frac{k_2}{k_2 - k_1} e^{-k_1 t} - \frac{b(k_2 - k_1) - ak_1}{k_2 - k_1} e^{-k_2 t} \end{aligned}$$

gegeben ist.

Bemerkung 7 Wir schränken uns jetzt wieder auf den Fall $n = 2$ ein.

(a) Wie in den obigen Beispiel 10 (a)-(c) kann auch allgemein, entsprechend des Lösungsansatzes für die drei Fälle, stets ein Fundamentalsystem $\{\vec{y}_1(x), \vec{y}_2(x)\}$ konstruiert werden. Somit gilt für die allgemeine Lösung von (49)

$$\vec{y}_{hom}(x) = c_1 \vec{y}_1(x) + c_2 \vec{y}_2(x).$$

(b) Falls A eine reelle symmetrische Matrix ist, so sind alle Eigenwerte reell (siehe Abschnitt: Lineare Algebra). Folglich kann dann nur Fall 1 oder Fall 2 auftreten.

(c) Falls für (49) zwei Anfangsbedingungen in der Form

$$\vec{y}(x_0) = \vec{y}_0, \quad x_0 \in I,$$

vorgegeben sind, so existiert wie im Satz 7 genau eine Lösung $\vec{y}_{part}(x)$.

5.4.2 Inhomogene Systeme

Wir beschäftigen uns jetzt mit dem *inhomogenen System* von *zwei* linearen Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten der Form

$$\vec{y}'(x) = A\vec{y}(x) + \vec{s}(x), \quad x \in I, \quad (62)$$

wobei $\vec{s}(x) = (s_1(x), s_2(x))$ eine stetige Störfunktion ist, die die Inhomogenität erzeugt.

Auch hier verwenden wir die Methode der *Variation der Konstanten*. Die allgemeine Lösung des zugehörigen homogenen Problems

$$\vec{y}'(x) = A\vec{y}(x), \quad x \in I, \quad (63)$$

sei

$$\vec{y}_{hom}(x) = c_1 \vec{y}_1(x) + c_2 \vec{y}_2(x).$$

Hierbei sind c_1 und c_2 beliebige reelle Zahlen, und die beiden Lösungen $\{\vec{y}_1(x), \vec{y}_2(x)\}$ bilden ein Fundamentalsystem. Unser Ansatz für eine spezielle Lösung von (62) lautet somit

$$\vec{y}_s(x) = c_1(x) \vec{y}_1(x) + c_2(x) \vec{y}_2(x). \quad (64)$$

Zur Bestimmung der unbekannt Funktionen $c_1(x)$ und $c_2(x)$ müssen wir zunächst die Ableitung von $\vec{y}_s(x)$ berechnen. Es gilt

$$\begin{aligned} \vec{y}'_s(x) &= c'_1(x) \vec{y}_1(x) + c'_2(x) \vec{y}_2(x) + \underbrace{c_1(x) \vec{y}'_1(x) + c_2(x) \vec{y}'_2(x)}_{A(c_1(x) \vec{y}_1(x) + c_2(x) \vec{y}_2(x))} \\ &= c'_1(x) \vec{y}_1(x) + c'_2(x) \vec{y}_2(x) + A\vec{y}_s(x) \\ &= \vec{s}(x) + A\vec{y}_s(x). \end{aligned}$$

Damit erhalten wir folgendes inhomogenes lineares Gleichungssystem zur Berechnung der Funktionen $c'_1(x)$ und $c'_2(x)$.

$$\underbrace{\begin{pmatrix} y_{11}(x) & y_{21}(x) \\ y_{12}(x) & y_{22}(x) \end{pmatrix}}_{Y(x)} \begin{pmatrix} c'_1(x) \\ c'_2(x) \end{pmatrix} = \begin{pmatrix} s_1(x) \\ s_2(x) \end{pmatrix}. \quad (65)$$

Da $\det(Y(x)) = W(x) \neq 0$ gilt, hat dieses inhomogene lineare Gleichungssystem genau eine Lösung $(c'_1(x), c'_2(x))$, die man z.B. mittels Cramerscher Regel bestimmen kann. Durch Integration erhält man dann die Funktionen $(c_1(x), c_2(x))$. Somit gelten analog zu Satz 7 folgende Aussagen.

Satz 13 *Wir betrachten das inhomogene lineare Differentialgleichungssystem 1. Ordnung mit konstanten Koeffizienten (62).*

(a) *Falls $\vec{y}_{inh}(x)$ die allgemeine Lösung von (62) ist, so gilt*

$$\vec{y}_{inh}(x) = \vec{y}_{hom}(x) + \vec{y}_s(x).$$

Dabei ist $\vec{y}_{hom}(x) = c_1\vec{y}_1(x) + c_2\vec{y}_2(x)$ die allgemeine Lösung des zugehörigen homogenen Problems (63), d.h. $\{\vec{y}_1(x), \vec{y}_2(x)\}$ bilden ein Fundamentalsystem für (63), und c_1, c_2 sind beliebig wählbare reelle Zahlen. $\vec{y}_s(x)$ ist eine spezielle Lösung des inhomogenen Problems (62).

(b) *Das Problem (62) mit den Anfangsbedingungen*

$$\vec{y}(x_0) = \vec{y}_0, \quad x_0 \in I,$$

besitzt genau eine Lösung.

Beispiel 11 Wir bestimmen die allgemeine Lösung des inhomogenen Systems

$$\begin{aligned} \dot{x}(t) &= 4x(t) + y(t) - 36t \\ \dot{y}(t) &= -2x(t) + y(t) - 2e^t. \end{aligned} \quad (66)$$

Das zugehörige homogene Problem lautet

$$\begin{aligned} \dot{x}(t) &= 4x(t) + y(t) \\ \dot{y}(t) &= -2x(t) + y(t). \end{aligned} \quad (67)$$

Die Matrix A ist dann durch

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

und das charakteristische Polynom durch

$$P_2(\lambda) = \begin{vmatrix} 4 - \lambda & 1 \\ -2 & 1 - \lambda \end{vmatrix} = (4 - \lambda)(1 - \lambda) + 2 = \lambda^2 - 5\lambda + 6$$

gegeben. Damit haben wir wegen

$$\lambda_{1/2} = \frac{5}{2} \pm \sqrt{\frac{25}{4} - \frac{24}{4}}$$

die beiden reellen Nullstellen $\lambda_1 = 3$ und $\lambda_2 = 2$. Das korrespondierende Fundamentalsystem für (67) ist damit durch

$$\left\{ \begin{pmatrix} c_{11} \\ c_{12} \end{pmatrix} e^{3t}, \begin{pmatrix} c_{21} \\ c_{22} \end{pmatrix} e^{2t} \right\}$$

gegeben, wobei entsprechend des Systems (66) bestimmte Beziehungen zwischen den Koeffizienten c_{11}, \dots, c_{22} bestehen. Unser Ansatz für die allgemeine Lösung lautet somit

$$x_{hom}(t) = c_{11}e^{3t} + c_{21}e^{2t}, \quad y_{hom}(t) = c_{12}e^{3t} + c_{22}e^{2t}.$$

Durch Einsetzen in die Differentialgleichung (67) und Koeffizientenvergleich erhalten wir aus

$$\dot{x}_{hom}(t) = 3c_{11}e^{3t} + 2c_{21}e^{2t} = (4c_{11} + c_{12})e^{3t} + (4c_{21} + c_{22})e^{2t},$$

dass

$$3c_{11} = 4c_{11} + c_{12} \implies c_{12} = -c_{11} =: -C_1$$

$$2c_{21} = 4c_{21} + c_{22} \implies c_{22} = -2c_{21} =: -2C_2$$

gelten muss. Somit ist die allgemeine Lösung des zugehörigen homogenen Problems durch

$$x_{hom}(t) = C_1 e^{3t} + C_2 e^{2t}, \quad y_{hom}(t) = -C_1 e^{3t} - 2C_2 e^{2t} \quad (68)$$

gegeben. Wählen wir $C_1 = 1$ und $C_2 = 1$, so erhalten wir mit

$$\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{3t}, \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^{2t} \right\}$$

ein Fundamentalsystem des homogenen Problems. Entsprechend der früheren Überlegungen verwenden wir folgenden Ansatz für die Bestimmung einer speziellen Lösung von (66)

$$x_s(t) = C_1(t)e^{3t} + C_2(t)e^{2t}, \quad y_s(t) = -C_1(t)e^{3t} - 2C_2(t)e^{2t}. \quad (69)$$

Damit erhalten wir das GLS

$$\begin{pmatrix} e^{3t} & e^{2t} \\ -e^{3t} & -2e^{2t} \end{pmatrix} \begin{pmatrix} \dot{C}_1(t) \\ \dot{C}_2(t) \end{pmatrix} = \begin{pmatrix} -36t \\ -2e^t \end{pmatrix}.$$

Mittels Cramerscher Regel erhält man wegen

$$W(t) = \begin{vmatrix} e^{3t} & e^{2t} \\ -e^{3t} & -2e^{2t} \end{vmatrix} = -e^{5t} \neq 0,$$

dass

$$\dot{C}_1(t) = \frac{\begin{vmatrix} -36t & e^{2t} \\ -2e^t & -2e^{2t} \end{vmatrix}}{-e^{5t}} = \frac{72te^{2t} + 2e^{3t}}{-e^{5t}} = -72te^{-3t} - 2e^{-2t}$$

und

$$\dot{C}_2(t) = \frac{\begin{vmatrix} e^{3t} & -36t \\ -e^{3t} & -2e^t \end{vmatrix}}{-e^{5t}} = \frac{-2e^{4t} - 36te^{3t}}{-e^{5t}} = 2e^{-t} + 36te^{-2t}$$

gelten muss. Durch Integration (partielle Integration) erhalten wir

$$C_1(t) = 24te^{-3t} + 8e^{-3t} + e^{-2t} + c_1, \quad C_2(t) = -18te^{-2t} - 9e^{-2t} - 2e^{-t} + c_2.$$

Damit erhalten wir nach (69) mit $c_1 = c_2 = 0$ die spezielle Lösung des inhomogenen Problems

$$x_s(t) = 6t - 1 - e^t, \quad y_s(t) = 12t + 10 + 3e^t.$$

Die gesuchte allgemeine Lösung von (66) lautet damit wegen (68)

$$\begin{aligned} x_{inh}(t) &= x_s(t) + x_{hom}(t) = 6t - 1 - e^t + C_1e^{3t} + C_2e^{2t}, \\ y_{inh}(t) &= y_s(t) + y_{hom}(t) = 12t + 10 + 3e^t - C_1e^{3t} - 2C_2e^{2t}. \end{aligned}$$

Bemerkung 8 (a) Wir bemerken, dass man auch hier mit *speziellen Ansätzen* arbeiten kann (wie früher bei inhomogenen linearen Differentialgleichungen), die der Störfunktion $\vec{s}(t)$ angepasst sind, um eine spezielle Lösung des inhomogenen Problems zu finden.

(b) Abschließend betonen wir noch einmal, dass alle Ergebnisse aus dem Abschnitt 5.4 sich auf Systeme mit $n \geq 2$ Gleichungen analog ausdehnen lassen.