2×2

ALBERTO S. CATTANEO

Inhaltsverzeichnis

Date: 20. Oktober 2012.

1. Einführung

In diesen Notizen werden wir uns mit 2×2-Matrizen befassen. Wir werden sie zur Auflösung von linearen Gleichungsystemen und von Systemen von linearen Differenzialgleichungen (je mit zwei Gleichungen und zwei Unbekannten) und zur Beschreibung von symmetrischen Bilinearformen auf der Ebene anwenden. Wir werden die natürlichen Äquivalenzrelationen in den drei Fällen einführen und die entsprechenden Normalformen herleiten.

All die hier eingeführten Begriffe und bewiesenen Sätze haben eine höherdimensionale Verallgemeinerung (s. z.B. [2]), zu der diese Notizen als „Aufwärmen" gelten.

Danksagung. Helen Riedtmann hat mich in der Bearbeitung der deutschen Version unterstützt; ich danke ihr vielmals.

2. Lineare Systeme

In diesem Abschnitt befassen wir uns mit linearen Systemen zweier Gleichungen mit zwei Unbekannten, x und y:

$$\begin{aligned} ax + by &= u \\ cx + dy &= v \end{aligned} \tag{2.1}$$

Wir werden sehen, dass „im Allgemeinen" solch ein System den Durchschnitt zweier Geraden in der Ebene darstellt und deshalb genau eine Lösung besitzt. Es gibt aber auch andere Möglichkeiten, die sog. entarteten Fälle.

Bemerkung 2.1. Die Variablen x und y heissen Unbekannte, die Konstanten a, b, c und d heissen Koeffizienten des Gleichungssystem, die Konstanten u and v heissen Werte. Insgesamt heissen a, b, c, d, u and v erweiterte Koeffizienten.

2.1. Geometrische Betrachtungen.

Wir beginnen eigentlich mit dem Studium einer einzigen Gleichung, z.B.,

$$ax + by = u. \tag{2.2}$$

Ist mindestens einer der zwei Koeffizienten a und b von Null verschieden, so stellt (2.2) eine Gerade dar. Ist z.B. $b \neq 0$, so können wir sie zur Normalform $y = -\frac{a}{b}x + \frac{u}{b}$ umformen. Ist $b = 0$ aber $a \neq 0$, so haben wir die senkrechte Gerade $x = \frac{u}{a}$.

Der Wert u spielt hier keine besondere Rolle. Es ist nur zu bemerken, dass für $u = 0$ die Gerade durch den Ursprung $x = y = 0$ geht. Sind a und b gegeben, so sind die durch (2.2) für verschiedene Werte von u beschriebenen Geraden zueinander parallel.

Bemerkung 2.2. Multiplizieren wir eine Gleichung mit einer Konstanten $\lambda \neq 0$, so bekommen wir eine äquivalente Gleichung (d.h. eine Gleichung mit genau der gleichen Lösungsmenge). Deshalb stellt die Gleichung

$$a'x + b'y = u'$$

die gleiche Gerade wie (2.2) dar, wenn $\exists \lambda \neq 0$, s.d. $a' = \lambda a$, $b' = \lambda b$ und $u' = \lambda u$. Gilt für ein $\lambda \neq 0$ $a' = \lambda a$ und $b' = \lambda b$ aber $u' \neq \lambda u$, so sind die zwei Geraden parallel, aber disjunkt.

Der entartete Fall entsteht, wenn $a = b = 0$ gilt. Da gibt es zwei Unterfälle: $u = 0$ und $u \neq 0$. Im ersteren Unterfall haben wir die Gleichung $0 = 0$, die von jedem Paar (x, y) aufgelöst wird und deshalb die ganze Ebene darstellt; im zweiteren Unterfall haben wir die Gleichung $0 = u \neq 0$, die keine Lösung besitzt und deshalb die leere Menge (oder bessere die leere Teilmenge der Ebene) darstellt.

Zusammenfassung: *Eine lineare Gleichung in zwei Unbekannten stellt die leere Menge oder eine Gerade oder die ganze Ebene dar.*

Bemerkung 2.3. Eine lineare Gleichung kann aber nie einen einzigen Punkt darstellen.

Wenn wir jetzt beide Gleichungen in (2.1) betrachten, sollen wir zwei Teilmengen der Ebene von der obigen Form miteinander schneiden:

(1) Stellt eine der Gleichungen die leere Menge dar, so ist der Durchschnitt auch leer und das System hat keine Lösung.
(2) Stellen beide Gleichungen die ganze Ebene dar, dann ist ihr Durchschnitt auch so und jedes Paar (x, y) ist eine Lösung.
(3) Stellt eine der Gleichungen eine Gerade und die andere die ganze Ebene dar, so ist ihr Durschnitt auch eine Gerade.
(4) Schliesslich, wenn beide Gleichungen eine Gerade darstellen, dann haben wir drei Unterfälle:
    (a) Die zwei Geraden sind gleich und deshalb ist ihr Durchschnitt auch eine Gerade (eigentlich die gleiche Gerade).
    (b) Die zwei Geraden sind parallel und ihr Durchschnitt ist deshalb leer.
    (c) Die zwei Geraden sind verschieden und nicht parallel und deshalb ist ihr Durchschnitt ein Punkt.

Zusammenfassung: *Das System* (2.1) *kann die leere Menge (1 oder 4b), einen Punkt (4c), eine Gerade (3 oder 4a) oder die ganze Ebene (2) darstellen.*

Algebraisch entsprechen die obigen Fälle zu folgenden Bedingungen:

(1) $(a = b = 0,\ u \neq 0)$ oder $(c = d = 0,\ v \neq 0)$.
(2) $a = b = u = 0$ und $c = d = v = 0$.
(3) $(a = b = u = 0$ und $cd \neq 0)$ oder $(ab \neq 0$ und $c = d = v = 0)$.
(4) $ab \neq 0$ $cd \neq 0$ und:
    (a) $\exists \lambda \neq 0$ s.d. $c = \lambda a$, $d = \lambda b$ und $v = \lambda u$.
    (b) $\exists \lambda \neq 0$ s.d. $c = \lambda a$, $d = \lambda b$, aber $v \neq \lambda u$.
    (c) Keiner der obigen Fälle.

Bemerkung 2.4. Es ist einfach zu bemerken, dass genau in den Fällen 1, 2, 3, 4a und 4b

$$ad - bc = 0$$

gilt. Deshalb kann der nichtentartete Fall, in dem die Lösung ein Punkt ist, durch die Bedingung $ad - bc \neq 0$ charakterisiert werden. Die Grösse

$$ad - bc \tag{2.3}$$

heisst Determinante des Gleichungssystems.

2.2. **Algebraische Betrachtungen.** Jetzt wollen wir das System (2.1) rein alegraisch auflösen.Wir beginnen mit dem nichtentarteten Fall 4c.

Zunächst können wir annehmen, einer der Koeffizienten a, b, c und d ist von Null verschieden, z.B. $c \neq 0$. In diesem Falle können wir die zweite Gleichung mit $\frac{a}{c}$ multiplizieren und erhalten:

$$ax + \frac{ad}{c}y = \frac{a}{c}v.$$

Subtrahieren wir jetzt die 1. von der 2. Gleichung, bekommen wir

$$\left(\frac{ad}{c} - b\right)y = \frac{a}{c}v - u$$

oder, äquivalent,

$$(ad - bc)y = av - cu.$$

Da im Fall 4c $ad - bc \neq 0$ gilt, erhalten wir

$$y = \frac{av - cu}{ad - bc}. \tag{2.4}$$

Setzen wir das jetzt in die 2. Gleichung, so haben wir

$$cx = \frac{dcu - dav + adv - bcv}{ad - bc} = c\frac{du - bv}{ad - bc}$$

und deshalb

$$x = \frac{du - bv}{ad - bc}. \tag{2.5}$$

Im Falle $c = 0$ gilt $a \neq 0$ und $b \neq 0$ wegen $ad - bc \neq 0$. Die 2. Gleichung (jetzt $dy = v$) ergibt dann $y = \frac{v}{d}$, was mit (2.4) (für $c = 0$) übereinstimmt. Setzen wir das in die 1. Gleichung ein, dann bekommen wir $ax = -\frac{b}{d}v + u$ und deshalb $x = \frac{du-bv}{ad}$, was mit (2.5) (wieder für $c = 0$) übereinstimmt. Wir können diese Berechnungen im folgenden Satz zusammenfassen:

Satz 2.5. *Ist die Determinante $ad - bc$ von Null verschieden, so hat das Gleichungssystem* (2.1) *genau eine Lösung, nämlich*

$$\boxed{\begin{aligned} x &= \frac{du - bv}{ad - bc} \\ y &= \frac{av - cu}{ad - bc} \end{aligned}}. \tag{2.6}$$

2.3. **Das Gauss'sche Eliminationsvefahren.** Wir wollen jetzt das System (2.1) ohne Beschränkungen über die Determinante simplifizieren. Das folgende Verfahren kann man zu linearen Systemen einer beliebigen Anzahl Gleichungen und Unbekannten erweitern.

Fall 1. Alle Koeffizienten a, b, c, d sind Null. Dann haben wir das System

$$\begin{aligned} 0 &= u \\ 0 &= v \end{aligned} \tag{2.7}$$

das keine Lösung besitz, es sei denn es gilt $u = v = 0$. Im letzteren Falle ist jedes Paar (x, y) eine Lösung, d.h. die Lösungsmenge ist die ganze Ebene.

Fall 2. Die Koeffizienten a und c sind Null, aber mindestens einer der Koeffizienten b und d ist von Null verschieden. O.B.d.A.[1] können wir dann $b \neq 0$ annehmen (sonst vertauschen wir die zwei Gleichungen). Dividieren wie die 1. Gleichung durch b, so bekommen wir

$$\begin{aligned} y &= \frac{u}{b} \\ dy &= v \end{aligned}$$

Subtrahieren wir d mal die 1. Gleichung von der 2. Gleichung, dann erhalten wir schliesslich das System

$$\begin{aligned} y &= \frac{u}{b} \\ 0 &= v - \frac{du}{b} \end{aligned} \tag{2.8}$$

das keine Lösung besitzt, es sei denn es gilt $bv - du = 0$. Im letzteren Falle ist y durch die 1. Gleichung bestimmt, aber x ist ganz frei; deshalb ist die Lösungsmenge die waagerechte Gerade $y = \frac{u}{b}$.

Fall 3. Einer der zwei Koeffizienten a und c ist von Null verschieden. O.B.d.A. können wir jetzt $a \neq 0$ annehmen (sonst vertauschen wir die zwei Gleichungen). Dividieren wir die 1. Gleichung durch a, dann bekommen wir

$$\begin{aligned} x + \frac{b}{a}y &= \frac{u}{a} \\ cx + dy &= v \end{aligned}$$

Subtrahieren wir c mal die 1. Gleichung von der 2. Gleichung, so erhalten wir

$$\begin{aligned} x + \frac{b}{a}y &= \frac{u}{a} \\ \left(d - \frac{cb}{a}\right) y &= v - \frac{cu}{a} \end{aligned}$$

Es gibt jetzt zwei Unterfälle:

Fall 3.a. Ist die Determinante $ad - bc = 0$, so hat das System die Form

$$\begin{aligned} x + \frac{b}{a}y &= \frac{u}{a} \\ 0 &= v - \frac{cu}{a} \end{aligned} \tag{2.9}$$

Es hat keine Lösungen ausser, wenn $av - cu = 0$ gilt. Im letzteren Falle ist die Lösungsmenge die Gerade $x = -\frac{ba}{y} + \frac{u}{a}$.

[1] Ohne Beschränkung der Allgemeinheit

Fall 3.b. Gilt $ad-bc \neq 0$, dann können wir die 2. Gleichung durch $d-\frac{cb}{a}$ dividieren und

$$\begin{aligned} x+\frac{b}{a}y &= \frac{u}{a} \\ y &= \frac{av-cu}{ad-cb} \end{aligned}$$

erhalten. Schliesslich können wir $\frac{b}{a}$ mal die 2. Gleichung von der 1. Gleichung subtrahieren und

$$\begin{aligned} x &= \frac{du-bv}{ad-bc} \\ y &= \frac{av-cu}{ad-bc} \end{aligned} \tag{2.10}$$

bekommen, was wieder (2.6) ist.

2.4. **Bemerkungen.**

Bemerkung 2.6. Man soll eigentlich bemerken, dass (2.9) folgt aus der Annahme, dass $a \neq 0$ gilt, aber die Determinante verschwindet. Im Fall 3 mit $a = 0$ gilt aber $c \neq 0$. Wir können die zwei Gleichungen Vertauschen und wie oben verfahren. Verschwindet die Determinante, so ist $b = 0$ und wir haben

$$\begin{aligned} x+\frac{d}{c}y &= \frac{v}{c} \\ 0 &= u \end{aligned} \tag{2.11}$$

Dieses System hat keine Lösungen ausser, wenn $u = 0$ gilt. Im letzteren Falle ist die Lösungsmenge die Gerade $x = -\frac{dc}{y} + \frac{v}{c}$.

Bemerkung 2.7. Was wir im letzten Unterabschnitt bewiesen haben, ist, dass das System (2.1) äquivalent zu einem der fünf Systemen (2.7), (2.8), (2.9), (2.11) oder (2.10) ist. Das enstsprechende System hat die sog. Stufengestalt.

Bemerkung 2.8. In den Gleichungssystemen (2.7), (2.8) und (2.10) sind die Koeffizienten entweder 1 oder 0. In (2.9) und (2.11) können die Koeffizienten $\frac{b}{a}$ und $\frac{d}{c}$ auf der anderen Seite jeden Wert annehmen. Mit folgendem Trick können wir aber dieses Problem umgehen: Wir definieren nämlich im Falle von (2.9) (der andere Fall wird auf einer sehr ähnlichen Weise studiert) neue Unbekannte x' und y' durch

$$\begin{aligned} x' &:= x+\frac{b}{a}y \\ y' &:= y \end{aligned}$$

Man bemerke, dass dieses System bez. x und y lösbar ist. Deshalb haben wir die inverse Transformation

$$\begin{aligned} x &= x'-\frac{b}{a}y' \\ y &= y' \end{aligned}$$

Bez. der neuen Unbekannten hat jetzt das System (2.9) die Form

$$\begin{aligned} x' &= \frac{u}{a} \\ 0 &= v-\frac{cu}{a} \end{aligned} \tag{2.12}$$

Die Koeffizienten sind jetzt, auch in diesem Falle, entweder 1 oder 0.

Zusammenfassung: *Jedes lineare Gleichungssystem ist, bis auf eine Neudefinition der Unbekannten, zu einem System, dessen Koeffizienten entweder* 1 *oder* 0 *sind, äquivalent.*

Bemerkung 2.9. Man kann auch das System (2.8) in eine zu (2.12) ähnliche Form überführen, indem man neue Unbekannte durch

$$\begin{aligned} x' &:= y \\ y' &:= x \end{aligned}$$

definiert. Hier auch kann man die inverse Transformation

$$\begin{aligned} x &= y' \\ y &= x' \end{aligned}$$

berechnen und aus (2.8) das System

$$\begin{aligned} x' &= \frac{u}{b} \\ 0 &= v - \frac{du}{b} \end{aligned} \tag{2.13}$$

gewinnen.

3. Matrizen

Die Resultate von Abschnitt 2 lassen sich mit der Matrixbezeichnung schön zusammenfassen und stark erweitern.

3.1. **Koeffizientenmatrizen.**

Das System (2.1) ist durch die Koeffizienten (a, b, c, d) und die Werte (u, v) völlig bestimmt. Es ist üblich die Koeffizienten in einer Tabelle

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{3.1}$$

zusammenzufassen. Solch eine Tabelle heisst Koeffizientenmatrix des linearen Systems und ist ein Beispiel einer 2×2-Matrix (d.h. einer Matrix mit 2 Zeilen und 2 Spalten).

Die Determinante der Koeffizientenmatrix wird als die Determinante des entsprechenden Gleichungssystems, d.h. durch Formel (2.3), definiert:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} := \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc \tag{3.2}$$

Bemerkung 3.1. Beide Notationen $\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right|$ und $\det \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ sind sehr üblich.

Ist $\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| \neq 0$, kann man die Lösung (2.6) zum System (2.1) wie folgt umschreiben:

$$x = \frac{\begin{vmatrix} u & b \\ v & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$$

$$y = \frac{\begin{vmatrix} a & u \\ c & v \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$$

Diese Formeln sind die sog. Cramersche Regel.

Zur Erinnerung: Der Nenner in jeder Formel ist die Determinante der Koeffizientenmatrix. Der Zähler der ersten (bzw. zweiten) Formel ist die Determinante der Matrix, die aus der Koeffizientenmatrix durch das Vertausch der ersten (bzw. zweiten) Spalte durch die Wertenspalte $\begin{smallmatrix} u \\ v \end{smallmatrix}$ entsteht.

Bemerkung 3.2 (Zeilenstufenform)*.* Wir können Bemerkung 2.7 wie folgt ausdrücken: *Das lineare Gleichungssystem* (2.1) *ist zu einem System äquivalent, dessen Koeffizientenmatrix eine der folgenden 4 Zeilenstufenformen hat:*

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & \star \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

wobei $\star$ eine beliebige Zahl sein kann.[2]

Bemerkung 3.3 (Normalform)*.* Wenn wir Neudefinitionen der Unbekannten, wie in Bemerkungen 2.8 und 2.9, erlauben, können wir die Koeffizientenmatrix zu einer der folgenden 3 Normalformen überführen:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

3.2. **Matrixmultiplikation.** Die Koeffizienten des linearen Systems (2.1) haben wir in der Koeffizientenmatrix (3.1) zusammengefasst. Wir können auch die Unbekannten x, y und die Werte u, v in dem Unbekanntenvektor $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ und in dem Wertenvektor $\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)$ zusammenfassen. Das System selbst können wir durch die „Rechenregel"

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

kompakt schreiben als

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}. \tag{3.3}$$

[2]Für die Matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ mit $ad - bc = 0$ haben wir früher $\star = \frac{b}{a}$, falls $a \neq 0$, und $\star = \frac{d}{c}$, falls $a = 0$ und $c \neq 0$, erhalten.

Noch besser können wir Abkürzungen für die Koeffizientenmatrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, für den Unbekanntenvektor $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ und für den Wertenvektor $\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)$ einführen:

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

$$\mathbf{x} := \begin{pmatrix} x \\ y \end{pmatrix},$$

$$\mathbf{u} := \begin{pmatrix} u \\ v \end{pmatrix}.$$

Damit können wir schliesslich ds System als

$$\mathsf{A} \cdot \mathbf{x} = \mathbf{u}$$

schreiben.

Bemerkung 3.4. Wir werden normalerweise Matrizen durch leichtfette Grossbuchstaben (wie z.B. A) und Vektoren durch fette Kleinbuchstaben (wie z.B. $\mathbf{x}$ oder $\mathbf{u}$) bezeichnen. Eine andere übliche Notation für Vektoren ist Kleinbuchstaben mit einem Pfeil (wie z.B. $\vec{x}$). Oft werden Matrizen und Vektoren einfach durch Buchstaben (wie z.B. A oder v) bezeichnet.

Bemerkung 3.5. Wir werden oft den Multiplikationspunkt weglassen. Unser System würden wir einfach als

$$\mathsf{A}\mathbf{x} = \mathbf{u}$$

schreiben.

Es ist jetzt sehr nützlich die Multiplikation einer Matrix mit einem Vektor zur Multiplikation zweier Matrizen zu erweitern:

$$\begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} := \begin{pmatrix} ea + fc & eb + fd \\ ga + hc & gb + hd \end{pmatrix}. \tag{3.4}$$

Bemerkung 3.6. Man kann diese Formel aus (3.3) gewinnen, indem man die Matrix $\mathsf{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ als Nebeneinanderstellung der Vektoren $\mathbf{a} := \left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$ und $\mathbf{b} := \left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)$ betrachtet:

$$\mathsf{A} = \begin{pmatrix} \mathbf{a} & \mathbf{b} \end{pmatrix}.$$

Führen wir $\mathsf{B} := \begin{pmatrix} e & f \\ g & h \end{pmatrix}$ ein, so haben wir nämlich

$$\mathsf{BA} = \begin{pmatrix} \mathsf{B}\mathbf{a} & \mathsf{B}\mathbf{b} \end{pmatrix}.$$

Satz 3.7. *Die Operation von Matrizen auf Vektoren ist assoziativ, d.h.*

$$\boxed{\mathsf{B}(\mathsf{A}\mathbf{x}) = (\mathsf{BA})\mathbf{x}}$$

für alle Matrizen A, B *und alle Vektoren* $\mathbf{x}$.

Beweis. Wir schreiben $\mathsf{B} := \begin{pmatrix} e & f \\ g & h \end{pmatrix}$, $\mathsf{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ und $\mathbf{x} := \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$. Dann gelten

$$\mathsf{B}(\mathsf{A}\mathbf{x}) = \begin{pmatrix} e & f \\ g & h \end{pmatrix} \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} =$$
$$= \begin{pmatrix} eax + eby + fcx + fdy \\ gax + gby + hcx + hdy \end{pmatrix}$$

und

$$(\mathsf{BA})\mathbf{x} = \left(\begin{pmatrix} e & f \\ g & h \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ea+fc & eb+fd \\ ga+hc & gb+hd \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = $$
$$= \begin{pmatrix} eax+eby+fcx+fdy \\ gax+gby+hcx+hdy \end{pmatrix}$$

und die zwei Ergebnisse sind gleich. □

Aus Bemerkung 3.6 folgt jetzt, dass das Matrizenprodukt assoziativ ist. Nämlich:

Folgesatz 3.8. *Für alle Matrizen* $\mathsf{A}, \mathsf{B}, \mathsf{C}$ *gilt*

$$\boxed{(\mathsf{AB})\mathsf{C} = \mathsf{A}(\mathsf{BC})}\,.$$

Bemerkung 3.9. Man schreibt normalerweise Ausdrücke wie $\mathsf{AB}\mathbf{x}$ und ABC ohne Klammern, da es wegen der Assoziativität keine Zweideutigkeit gibt.

Bemerkung 3.10. Die Matrixmultiplikation ist aber i.A. **nicht kommutativ**; d.h. i.A.

$$\boxed{\mathsf{BA} \neq \mathsf{AB}}\,.$$

Deshalb ist es wichtig, auf die Reihenordnung in der Matrixmultiplikation Acht zu geben.

Z.B.

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ a & 0 \end{pmatrix}$$

aber

$$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix}$$

und die Ergebnisse sind verschieden, wenn der Parameter $a \neq 1$ ist.

Bemerkung 3.11. Es gibt zwei sehr wichtige Matrizen: die Einheitsmatrix

$$\mathbf{1} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

und die Nullmatrix

$$\mathbf{0} := \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Sie haben folgende Eigenschaften: Für jede Matrix A gilt

$$\begin{aligned} \mathbf{1}\mathsf{A} = \mathsf{A}\mathbf{1} &= \mathsf{A}, \\ \mathbf{0}\mathsf{A} = \mathsf{A}\mathbf{0} &= \mathbf{0}. \end{aligned}$$

Bemerkung 3.12. Die Einheitsmatrix spielt die Rolle der identischen Transformation; d.h. es gilt

$$\mathbf{1x} = \mathbf{x}$$

für jeden Vektor $\mathbf{x}$.

Bemerkung 3.13. Eine andere übliche Bezeichnung für die Einheitsmatrix ist Id, was für „identische Abbildung“ steht.

Eine wichtige Eigenschaft, die uns später sehr nützlich sein wird, ist die Verträglichkeit der Matrixmultiplikation mit der Berechnung der Determinante. Nämlich

Satz 3.14. *Für je zwei Matrizen* A, B *gilt*

$$\boxed{\det(\mathsf{BA}) = \det \mathsf{B} \det \mathsf{A}}\,.$$

Beweis. Wir haben, für $\mathsf{B} := \left(\begin{smallmatrix} e & f \\ g & h \end{smallmatrix}\right)$ und $\mathsf{A} := \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$,

$$\begin{aligned}\det(\mathsf{BA}) = \begin{vmatrix} ea+fc & eb+fd \\ ga+hc & gb+hd \end{vmatrix} = \\ &= eagb + eahd + fcgb + fchd - ebga - ebhc - fdga - fdhc = \\ &= eahd + fcgb - ebhc - fdga\end{aligned}$$

und

$$\det \mathsf{B} \det \mathsf{A} = (eh - fg)(ad - bc) = ehad + fgbc - ehbc - fgad.$$

Die zwei Ergebnisse sind deutlich gleich. □

Bemerkung 3.15. Da die Multiplikation von Zahlen kommutativ ist, folgt aus dem Satz ohne weitere Berechnungen, dass

$$\boxed{\det \mathsf{BA} = \det \mathsf{AB}}$$

für je zwei Matrizen A, B. D.h. die Determinante „vergisst“ die Nichtkommutativität der Matrixmultiplikation.

3.3. **Systemumformungen.** Durch die Matrixmultiplikation können wir die üblichen Umformungen von linearen Gleichungssystemen ausdrücken. Nämlich entspricht zu jeder Umformung des Systems $\mathsf{A}\mathbf{x} = \mathbf{u}$ eine Matrix B so, dass das umgeformte System als

$$\mathsf{BA}\mathbf{x} = \mathsf{B}\mathbf{u}$$

geschrieben wird.

In diesem Abschnitt werden wir durchweg

$$\mathsf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} u \\ v \end{pmatrix}$$

setzen.

3.3.1. *Vertauschen der Gleichungen.* Das wird durch $\mathsf{B} =$

$$\boxed{\mathsf{P} := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}$$

erhalten. Nämlich gilt $\mathsf{P}\mathbf{u} = \left(\begin{smallmatrix} v \\ u \end{smallmatrix}\right)$ und $\mathsf{PA}\mathbf{x} = \left(\begin{smallmatrix} cx+dy \\ ax+by \end{smallmatrix}\right)$.

3.3.2. *Skalierung einer Gleichung.* Die Skalierung der ersten Gleichung durch eine Konstante $\lambda \neq 0$ wird durch $\mathsf{B} =$

$$\boxed{\mathsf{S}_1(\lambda) := \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}}$$

erhalten.

Bemerkung 3.16. Die Skalierung der zweiten Gleichung durch eine Konstante $\lambda \neq 0$ wird durch $\mathsf{B} = \mathsf{S}_2(\lambda) := \left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right)$ erhalten. Diese Matrizen sind nicht unabhängig von den bereits eingeführten Matrizen P und $\mathsf{S}_1(\lambda)$, denn es gilt $\mathsf{S}_2(\lambda) = \mathsf{P}\mathsf{S}_1(\lambda)\mathsf{P}$.

Bemerkung 3.17. Für $\lambda = 1$ verändert sich das System nicht. Die Matrix $\mathbf{1} := \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) = \mathsf{S}_1(1) = \mathsf{S}_2(1)$ ist die Einheitsmatrix.

3.3.3. *Addierung von Gleichungen.* Die Addierung der ersten zur zweiten Gleichung erhält man durch $\mathsf{B} =$

$$\boxed{\mathsf{Q} := \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}}.$$

Bemerkung 3.18. Die Addierung der zweiten zur ersten Gleichung erhält man durch $\mathsf{B} = \tilde{\mathsf{Q}} := \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, die von den bereits eingeführten Matrizen nicht unabhängig ist: $\tilde{\mathsf{Q}} = \mathsf{PQP}$. Die Addierung der ersten Gleichung mal α zur zweiten Gleichung erhält man durch $\mathsf{B} = \mathsf{T}(\alpha) := \left(\begin{smallmatrix} 1 & 0 \\ \alpha & 1 \end{smallmatrix}\right)$. Das kann man auch durch die bereits eingeführten Matrizen ausdrücken. Nämlich, $\mathsf{T}(0) = \mathbf{1} = \mathsf{S}_1(1)$ und für $\alpha \neq 0$ hat man $\mathsf{T}(\alpha) = \mathsf{S}_1(\alpha^{-1})\mathsf{Q}\mathsf{S}_1(\alpha)$.

3.3.4. *Weitere Bemerkungen.*

Bemerkung 3.19. Die Matrizen P, Q und $\mathsf{S}_1(\lambda)$, $\lambda \neq 0$, heissen Elementarmatrizen.

Bemerkung 3.20 (Zeilenstufenform). Wir haben gesehen (Bemerkung 2.7), dass man ein lineares Gleichungssystem durch Umformungen in eine Stufengestalt überführen kann. Äquivalent (Bemerkung 3.2) kann man durch Umformungen eine Koeffizientenmatrix in Zeilestufenform überführen. D.h.: Durch endlich viele Multiplikationen von links mit Elementarmatrizen kann man jede Matrix in eine der vier Zeilenstufenformen in Bemerkung 3.2 überführen. Anders ausgedrückt:

Hilfssatz 3.21. *Sei* A *eine beliebige Matrix. Dann gibt es endlich viele Elementarmatrizen* B_1, B_2, …, B_n *so, dass*

$$\tilde{\mathsf{A}} := \mathsf{B}_1\mathsf{B}_2\cdots\mathsf{B}_n\mathsf{A}$$

eine der 4 Zeilenstufenormen

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & \star \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

besitzt, wobei $\star$ *eine Zahl ist.*

I.A. können wir ein System $\mathsf{A}\mathbf{x} = \mathbf{u}$ mit einer beliebigen Matrix B multiplizieren.

Frage: Wann sind die Systeme $\mathsf{A}\mathbf{x} = \mathbf{u}$ und $\mathsf{BA}\mathbf{x} = \mathsf{B}\mathbf{u}$ äquivalent?

Um diese Frage zu beantworten, müssen wir verstehen, wann man durch eine Matrix B „dividieren“ kann. Das ist das Inversionsproblem und wir werden sehen, es ist mit dem Begriff der Determinante eng verknüpft (s. Satz 3.23).

3.4. **Die Inversion von Matrizen.**

Definition 3.22. Eine Matrix A heisst invertierbar, wenn es eine Matrix B gibt, sodass

$$\mathsf{AB} = \mathbf{1}$$

gilt, wobei

$$\mathbf{1} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

die Einheitsmatrix ist.

Satz 3.23. *Eine Matrix ist genau dann invertierbar, wenn ihre Determinante nicht verschwindet.*

Beweis. **a)** Sei A invertierbar. Aus Satz 3.14 und aus der Gleichung $\det \mathsf{AB} = \mathbf{1}$ folgt, dass $\det \mathsf{A} \det \mathsf{B} = 1$ gilt. Deshalb ist $\det \mathsf{A} \neq 0$ (sonst hätten wir $0 = 1$).

b) Sei $\det \mathsf{A} \neq 0$. Dank Satz 2.5 (oder auch dank Bemerkung 2.4) sind in diesem Falle die Systeme

$$\mathsf{A} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{und} \qquad \mathsf{A} \begin{pmatrix} w \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

eindeutig lösbar. Diese Systeme können wir auch als

$$\mathsf{A} \begin{pmatrix} x & w \\ y & z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

zusammensetzen, was impliziert, dass die Matrix $\mathsf{B} = \left(\begin{smallmatrix} x & w \\ y & z \end{smallmatrix}\right)$ existiert und eigentlich eindeutig bestimmt ist. □

Bemerkung 3.24. Ist A invertierbar, so heisst die eindeutig bestimmte (s. Ende des obigen Beweises) Matrix B, die $\mathsf{AB} = \mathbf{1}$ erfüllt, die zu A inverse Matrix. Sie wird üblicherweise mit A^{-1} bezeichnet. D.h.:

$$\mathsf{AA}^{-1} = \mathbf{1}.$$

Satz 3.25. *Sei* $\mathsf{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ *invertierbar. Dann gilt:*

$$\det(\mathsf{A}^{-1}) = \frac{1}{\det \mathsf{A}}, \tag{3.5a}$$

$$\mathsf{A}^{-1} = \frac{1}{\det \mathsf{A}} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \tag{3.5b}$$

$$\mathsf{A}^{-1}\mathsf{A} = \mathbf{1}. \tag{3.5c}$$

Bemerkung 3.26. Die einfach zu erinnernden Formeln im obigen Satz sind sehr wichtig. Insbesondere erinnert man sich an (3.5b), indem man ausser der Division durch die Determinante bemerkt, dass die Diagonaleinträge der Matrix auf der rechten Seite einfach die miteinander vertauschten Diagonaleinträge von A sind, während die Einträge ausserhalb der Diagonale die von A mit entgegengesetzen Vorzeichen sind. Dank Gleichung (3.5a) ist es nicht zweideutig $\det \mathsf{A}^{-1}$ zu schreiben: Es kann nämlich $\det(\mathsf{A}^{-1})$ oder $(\det \mathsf{A})^{-1}$ bedeuten, aber die zwei Ausdrücke sind gleich.

Beweis. Gleichung (3.5a) folgt aus $\det \mathsf{A} \det(\mathsf{A}^{-1}) = \det(\mathsf{AA}^{-1}) = \det \mathbf{1} = 1$. Gleichungen (3.5b) und (3.5c) folgen aus den expliziten Rechnungen

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} \det \mathsf{A} & 0 \\ 0 & \det \mathsf{A} \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

□

Bemerkung 3.27. Als einfache Beispiele von invertierbaren Matrizen stehen die Elementarmatrizen: nämlich gelten $\det \mathsf{P} = -1$, $\det \mathsf{S}_1(\lambda) = \lambda$ und $\det \mathsf{Q} = 1$. Man kann auch die Inversen einfach berechnen:[3]

$$\mathsf{P}^{-1} = \mathsf{P}, \qquad \mathsf{S}_1(\lambda)^{-1} = \mathsf{S}_1(\lambda^{-1}), \qquad \mathsf{Q}^{-1} = \mathsf{S}_1(-1)\mathsf{Q}\mathsf{S}_1(-1). \tag{3.6}$$

Das zeigt, *die Inversen von Elementarmatrizen lassen sich als Produkte von Elementarmatrizen schreiben.*

Bemerkung 3.28. Die eindeutig bestimmte Lösung zum System $\mathsf{A}\mathbf{x} = \mathbf{u}$ mit $\det \mathsf{A} \neq 0$ kann man gewinnen, indem man das System durch A^{-1} umformt. Das so erhaltene äquivalente System ist einfach

$$\mathbf{x} = \mathsf{A}^{-1}\mathbf{u},$$

das explizit die Lösung ergibt. Wenn $\mathsf{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, $\mathbf{x} = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$, und $\mathbf{u} = \left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)$, dann haben wir dank (3.5b)

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\det \mathsf{A}} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{\det \mathsf{A}} \begin{pmatrix} du - bv \\ -cu + av \end{pmatrix},$$

was (2.6) entspricht.

Bemerkung 3.29. Aus den Sätzen 3.14 und 3.23 folgt, dass das Produkt zweier invertierbaren Matrizen invertierbar ist. Sind nämlich A und B invertierbar, so gelten $\det \mathsf{A} \neq 0$ und $\det \mathsf{B} \neq 0$. Daraus folgt $\det \mathsf{AB} \neq 0$, was impliziert, dass AB invertierbar ist. Die zu AB inverse Matrix ist einfach zu berechnen: Wegen der Assoziativität gilt nämlich

$$\mathsf{B}^{-1}\mathsf{A}^{-1}\mathsf{AB} = \mathbf{1}.$$

Deshalb gilt

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

An diese sehr wichtige Formel (die Formel für das Inverse im nichtkommutativen Fall) kann man sich so erinnern: Die inverse Operation zum Anziehen eines Hemdes und danach einer Jacke ist das Ausziehen der Jacke und danach des Hemdes. Diese Formel kann man einfach zu einem beliebigen Produkt von invertierbaren Matrizen erweitern:

$$(\mathsf{A}_1\mathsf{A}_2 \cdots \mathsf{A}_n)^{-1} = \mathsf{A}_n^{-1} \cdots \mathsf{A}_2^{-1}\mathsf{A}_1^{-1}.$$

Definition 3.30. Sei

$$\mathrm{GL}_2 := \{\text{invertierbare } 2 \times 2\text{-Matrizen}\} = \{2 \times 2\text{-Matrizen } \mathsf{A} : \det \mathsf{A} \neq 0\}.$$

Es folgt, dass GL_2 eine Gruppe ist (GL steht für Englisch „general linear"). Mehr über Gruppen wird in Anhang A erklärt.

Folgender Satz ergibt eine Charakterisierung der invertierbaren Matrizen:

Satz 3.31. *Jede invertierbare Matrix lässt sich als endliches Produkt von Elementarmatrizen schreiben.*

Man sagt dann, die 2×2-Elementarmatrizen seien die Erzeuger der Gruppe GL_2.

[3] Man denke an die Bedeutung von den inversen Transformationen auf einem Gleichungssystem.

Beweis. Wir haben im Hilfssatz 3.21 gesehen, dass es für jede Matrix A endlich viele Elementarmatrizen B_1, B_2, $\ldots$, B_n gibt, sodass

$$\tilde{\mathsf{A}} := \mathsf{B}_1\mathsf{B}_2 \cdots \mathsf{B}_n\mathsf{A}$$

eine der 4 Zeilenstufenormen besitzt. Da die Elementarmatrizen invertierbar sind (s. Bemerkung 3.27), ist $\tilde{\mathsf{A}}$ invertierbar, wenn A so ist. Aber die einzige invertierbare Matrix in Zeilestufenform ist die Einheitsmatrix $\mathbf{1} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Es folgt

$$\mathbf{1} = \mathsf{B}_1\mathsf{B}_2 \cdots \mathsf{B}_n\mathsf{A}$$

und deshalb (s. Bemerkung 3.29)

$$\mathsf{A} = \mathsf{B}_n^{-1} \cdots \mathsf{B}_2^{-1}\mathsf{B}_1^{-1}.$$

Dank (3.6) können wir schliesslich die rechte Seite als Produkt Elementarmatrizen schreiben. □

3.5. **Auflösung von linearen Gleichungssystemen, Äquivalenz von Matrizen und Normalform.** Wir kommen jetzt zurück zum Problem der Auflösung von linearen Glecihungssystemen. Wir beginnen mit folgender

Definition 3.32. Zwei Systeme heissen äquivalent, wenn man das erste ins zweite durch endliche viele folgender elementarer Umformungen überführen kann:

(1) Vertausch der Gleichungen.
(2) Skalierung der ersten Gleichung.
(3) Addierung der ersten zur zweiten Gleichung.

Dann haben wir den

Satz 3.33. *Zwei Systeme* $\mathsf{A}\mathbf{x} = \mathbf{u}$ *und* $\tilde{\mathsf{A}}\mathbf{x} = \tilde{\mathbf{u}}$ *sind genau dann äquivalent, wenn es eine* invertierbare *Matrix* B *gibt mit* $\tilde{\mathsf{A}} = \mathsf{B}\mathsf{A}$ *und* $\tilde{\mathbf{u}} = \mathsf{B}\mathbf{u}$ *(d.h. man erhält das zweite System durch Multiplikation des ersten mit* B*).*

Beweis. Sind die zwei Systeme äquivalent, dann definieren wir B als Produkt der Elementarmatrizen, die zu den elementaren Umformungen entsprechen (s. Abschnitt 3.3).

Erhält man das zweite System durch Multiplikation des ersten durch B, so schreiben wir B dank Satz 3.31 als Produkt Elementarmatrizen. Das zweite System erhält man dann durch die entsprechenden Umformungen. □

Durch Äquivalenz können wir immer ein System in die Stufengestalt umformen (Bemerkung 2.7). In Bemerkung 2.8 haben wir weiter gesehen, dass eine System durch eine Neudefinition der Unbekannten weiter simplifiziert werden kann: Nämlich kann man die Koeffizientenmatrix zu einer der 3 Normalformen

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

überführen (Bemerkung 3.3).

Als Neudefinition der Unbekannten können wir jede invertierbare lineare Transformation erlauben: Definieren wir nähmlich

$$\mathbf{x}' = \mathsf{C}\mathbf{x},$$

so wird das lineare System $\mathsf{A}\mathbf{x} = \mathbf{u}$ zum System $\mathsf{A}\mathsf{C}^{-1}\mathbf{x}' = \mathbf{u}$ überführt, das auch linear ist. Wir können $\mathbf{x}$ aus $\mathbf{x}'$ durch

$$\mathbf{x} = \mathsf{C}^{-1}\mathbf{x}'$$

zurück gewinnen. Die zwei Systeme sind dann, in diesem erweiterten Sinne, äquivalent.

Zusammenfassung: *Zwei lineare Gleichungssysteme* $\mathsf{A}\mathbf{x} = \mathbf{u}$ *und* $\tilde{\mathsf{A}}\mathbf{x}' = \tilde{\mathbf{u}}$ *sind bis auf eine Neudefinition der Unbekannten äquivalent, wenn es zwei* invertierbare *Matrizen* B *und* C *gibt mit* $\tilde{\mathsf{A}} = \mathsf{B}\mathsf{A}\mathsf{C}^{-1}$ *und* $\tilde{\mathbf{u}} = \mathsf{B}\mathbf{u}$.

Das motiviert folgende

Definition 3.34. Zwei Matrizen A und $\tilde{\mathsf{A}}$ heissen äquivalent, wenn es invertierbare Matrizen B und C gibt, sodass

$$\boxed{\tilde{\mathsf{A}} = \mathsf{B}\mathsf{A}\mathsf{C}^{-1}}$$

gilt. In diesem Falle schreibt man,

$$\tilde{\mathsf{A}} \sim \mathsf{A}.$$

Satz 3.35. *Die Äquivalenz von Matrizen ist eine Äquivalenzrelation, d.h.*

Reflexivität: $\mathsf{A} \sim \mathsf{A}\ \forall \mathsf{A}$*;*
Symmetrie: $\tilde{\mathsf{A}} \sim \mathsf{A} \Rightarrow \mathsf{A} \sim \tilde{\mathsf{A}}$*;*
Transitivität: $\tilde{\mathsf{A}} \sim \mathsf{A},\ \hat{\mathsf{A}} \sim \tilde{\mathsf{A}} \Rightarrow \hat{\mathsf{A}} \sim \mathsf{A}$.

Beweis.

Reflexivität: Man nimmt $\mathsf{B} = \mathsf{C} = \mathbf{1}$.
Symmetrie: Nach der Definition haben wir $\tilde{\mathsf{A}} = \mathsf{B}\mathsf{A}\mathsf{C}^{-1}$ und deshalb $\mathsf{A} = \mathsf{B}^{-1}\tilde{\mathsf{A}}\mathsf{C}$. Da B^{-1} und C^{-1} auch invertiebar sind, haben wir $\mathsf{A} \sim \tilde{\mathsf{A}}$.
Transitivität: Wir haben $\tilde{\mathsf{A}} = \mathsf{B}\mathsf{A}\mathsf{C}^{-1}$ und $\hat{\mathsf{A}} = \tilde{\mathsf{B}}\tilde{\mathsf{A}}\tilde{\mathsf{C}}^{-1}$, wobei $\mathsf{B}, \mathsf{C}, \tilde{\mathsf{B}}, \tilde{\mathsf{C}}$ invertierbare Matrizen sind. Deshalb $\hat{\mathsf{A}} = \hat{\mathsf{B}}\mathsf{A}\hat{\mathsf{C}}^{-1}$ mit $\hat{\mathsf{B}} = \tilde{\mathsf{B}}\mathsf{B}$ und $\hat{\mathsf{C}} = \tilde{\mathsf{C}}\mathsf{C}$, die auch invertierbar sind.

□

Wir können jetzt Bemerkung 3.3 wie folgt ausdrücken und verbessern:

Satz 3.36. *Jede 2×2-Matrix ist zu einer der folgenden Matrizen äquivalent (Normalform)*

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Diese drei Matrizen sind zueinander nicht *äquivalent.*

Beweis. Die erste Aussage ist einfach eine Umdeutung von Bemerkung 3.3.

Um die zweite Aussage zu zeigen, bemerken wir zunächst, dass

$$\mathsf{B}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\mathsf{C}^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

für alle B und C; deshalb ist die Nullmatrix $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ nur zu sich selbst äquivalent (das zeigt auch, dass notwendigerweise die ursprüngliche Matrix A die Nullmatrix war).

Um zu zeigen, dass die anderen zwei Matrizen auf der Liste zueinander nicht äquivalent sind, bemerken wir, dass aus $\tilde{\mathsf{A}} = \mathsf{BAC}^{-1}$ die Gleichung

$$\det \tilde{\mathsf{A}} = \frac{\det \mathsf{B}}{\det \mathsf{C}} \det \mathsf{A}$$

folgt. Deshalb sind entweder beide Determinanten von zwei äquivalenten Matrizen gleich Null oder keine der zwei verschwindet. Da $\left|\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right| = 0$ und $\left|\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right| = 1$, sind die zwei Matrizen zueinander nicht äquivalent. □

Bemerkung 3.37. Aus dem letzten Teil des Beweises folgt, dass man die Äquivalenzrelation von Matrizen auf die Teilmenge GL_2 der invertierbaren Matrizen einschränken kann und dass jede invertierbare Matrix zur Einheitsmatrix äquivalent ist (in anderen Worten: GL_2 besteht aus einer einzigen Äquivalenzklasse).

Wir können schliesslich die Klassifizierung und Auflösung des linearen Systems

$$\mathsf{A}\mathbf{x} = \mathbf{u}$$

wiederholen und ein bisschen verbessern.

3.5.1. *Fall 1:* A *ist die Nullmatrix.* In diesem Falle haben wir das System $0 = \mathbf{u}$, dass keine Lösung besitzt, es sei denn es gilt $\mathbf{u} = 0$. Im letzteren Falle ist jedes $\mathbf{x}$ eine Lösung.

3.5.2. *Fall 2:* $\det \mathsf{A} = 0$ *aber* $\mathsf{A} \neq \mathbf{0}$. In diesem Falle ist das System zu

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{x}' = \tilde{\mathbf{u}}$$

äquivalent. Dieses System hat keine Lösung ausser, wenn $\tilde{\mathbf{u}}$ von der Form $\left(\begin{smallmatrix} \tilde{u} \\ 0 \end{smallmatrix}\right)$ ist. Ist die Äquivalenz $\tilde{\mathsf{A}} = \mathsf{BAC}^{-1}$ durch eine Matrix $\mathsf{B} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$ erhalten, so ist

$$gu + hv = 0$$

die Bedingung auf den Werten für die Lösbarkeit des Systems. Ist die Bedingung erfüllt, dann ist das System durch

$$\mathbf{x} = \mathsf{C}^{-1} \begin{pmatrix} eu + fv \\ y' \end{pmatrix}$$

aufgelöst, wobei y' ein freier Parameter ist.

3.5.3. *Fall 3:* A *ist invertierbar.* In diesem Falle hat das System genau die Lösung

$$\mathbf{x} = \mathsf{A}^{-1}\mathbf{u}.$$

3.6. **Lineare Abbildungen.** Es ist für weitere Bemerkungen wichtig, die Linearität besser zu erklären. Zunächst müssen wir die Skalarmultiplikation definieren: Für eine Zahl λ (ein „Skalar") und einen Vektor $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ definiert man

$$\lambda \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}.$$

Sind zwei Vektoren $\left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ gegeben, so definiert man die Addition

$$\begin{pmatrix} v \\ w \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} v + x \\ w + y \end{pmatrix}.$$

Bemerkung 3.38. Zeichnet man einen Vektor $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ als einen Pfeil vom Ursprung $(0,0)$ zum Punkt (x,y) auf der Ebene, so entspricht die Skalarmultiplikation einer Skalierung des Vektors und die Addition der Parallelogrammregel.

Es folgt dann aus einer einfachen Berechnung, dass

$$\mathsf{A}(\lambda\mathbf{x}+\mu\mathbf{y}) = \lambda\mathsf{A}\mathbf{x}+\mu\mathsf{A}\mathbf{y} \tag{3.7}$$

für jede Matrix A and alle Vektoren $\mathbf{x}$ und $\mathbf{y}$.

Bemerkung 3.39. Wir können $\mathbf{x} \mapsto \mathsf{A}\mathbf{x}$ als eine Abbildung $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ betrachten. Man sagt wegen (3.7), diese Abbildung sei linear. Im Allgemeinen sagt man, eine Abbildung $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ sei linear, wenn

$$\phi(\lambda\mathbf{x}+\mu\mathbf{y}) = \lambda\phi(\mathbf{x})+\mu\phi(\mathbf{y}), \qquad \forall \mathbf{x},\mathbf{y}$$

gilt. Man kann zeigen, dass dies genau dann passiert, wenn $\phi(\mathbf{x})$ von der Form $\mathsf{A}\mathbf{x}$ für irgendeine Matrix A ist.

Man kann Skalarmultiplikation und Addition auf Matrizen erweitern:

$$\lambda\begin{pmatrix} a & b \\ c & d \end{pmatrix} := \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} := \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix}. \tag{3.8}$$

Man kann sehr einfach zeigen, dass

$$\mathsf{A}(\mathsf{B}+\mathsf{C}) = \mathsf{AB}+\mathsf{AC} \qquad \text{und} \qquad (\mathsf{A}+\mathsf{B})\mathsf{C} = \mathsf{AC}+\mathsf{BC}$$

gelten. D.h. die Multiplikation von Matrizen ist distributiv.

4. Lineare Differenzialgleichungen

Eine Differenzialgleichung ist eine Gleichung, deren Unbekannte eine Funktion ist,[4] die zusammen mit ihren Ableitungen vorkommt. Die Ordnung einer Differenzialgleichung ist die höchste vorkommende Ableitungsordnung der Unbekannten.

Ein bekanntes Beispiel ist die Gleichung

$$\dot{x} = ax, \tag{4.1}$$

wobei a eine gegebene Konstante und $x(t)$ die Unbekannte ist. Diese Differenzialgleichung erster Ordnung wird zur Beschreibung von vielen Phänomenen in der Physik (z.B. dem radioaktiven Zerfall) und in der Wirtschaft (Zinseszins) angewandt.

Ein Beispiel einer Gleichung zweiter Ordnung ist das Newtonsche Gesetz (in einer Dimension) $F = ma$, wo F die Kraft, m das Mass und a die Beschleunigung eines Teilchens sind. In mathematischer Formulation haben wir

$$\ddot{x} = \frac{1}{m}F(t,x,\dot{x}), \tag{4.2}$$

wobei $x(t)$ die unbekannte Bahn, $m > 0$ eine gegebene Konstante und F eine gegebene Funktion sind.

In diesen Notizen werden wir uns mit linearen Differenzialgleichungen oder linearen Systemen von Differenzialgleichungen (mit zwei Unbekannten Funktionen) mit konstanten Koeffizienten befassen. Eine Differenzialgleichung heisst linear, wenn die

[4] Wir werden hier nur gewöhnliche Differenzialgleichungen betrachten, wo die Unbekannte eine Funktion einer einzigen Variable ist.

Unbekannte zusammen mit ihren Ableitungen nur linear vorkommt. Sie hat konstante Koeffizienten, wenn die vorkommenden Koeffizienten (d.h. alle Grösse ausser der Unbekannten und ihren Ableitungen) konstant sind. Das allgemeinste Beispiel einer linearen Differenzialgleichung erster Ordnung ist[5]

$$\dot{x} = a(t)x + b(t),$$

von der (4.1) ein Spezialfall ist. Sie hat konstante Koeffizienten, wenn a und b nicht von t abhängen, wie z.B. in (4.1). Das Newtonsche Gesetz (4.2) definiert eine lineare Differenzialgleichung mit konstanten Koeffizienten, wenn die Kraft F die Gestalt

$$F = -kx - \gamma\dot{x}$$

hat, wobei k und γ konstant sind. Der geschwindigkeitsabhängige Term beschreibt die Reibung und aus physikalischen Gründen ist $\gamma \geq 0$. Der erste Term beschreibt für $k > 0$ die sog. harmonische Kraft (und das ist eine Annäherung von jedem System in der Nähe von einem Punkt von stabilem Gleichgewicht). Die entsprechende Gleichung

$$m\ddot{x} = -kx - \gamma\dot{x} \tag{4.3}$$

beschreibt den sog. gedämpften harmonischen Oszillator.

4.1. **Lineare Differenzialgleichung mit konstanten Koeffizienten.** Wir beginnen mit Gleichung (4.1). Es ist bekannt, dass eine Lösung die Exponenzialfunktion e^{at} ist. Wir zeigen jetzt, dass sie bis auf eine multiplikative Konstante die einzige Lösung ist. Dazu definieren wir

$$X(t) := \mathrm{e}^{-at}x(t). \tag{4.4}$$

Dank der Ableitungsregel für Produkte erhalten wir dann

$$\dot{X}(t) = -a\mathrm{e}^{-at}x(t) + \mathrm{e}^{-at}\dot{x}(t) = \mathrm{e}^{-at}(-ax(t) + \dot{x}).$$

Deshalb ist x genau dann eine Lösung zu (4.1), wenn X eine Lösung zu

$$\dot{X} = 0$$

ist, d.h., wenn X konstant ist. Setzen wir $X(t) = K$, so erhalten wir, aus (4.4), die allgemeine Lösung zu (4.1):

$$x(t) = K\mathrm{e}^{at}.$$

Bemerkung 4.1. Die Konstante K kann festgelegt werden, wenn man statt nur Gleichung (4.1) auch die Anfangsbedingung

$$x(t_0) = x_0$$

betrachtet, wobei t_0 und x_0 gegeben sind. Eine Differenzialgleichung zusammen mit einer Anfangsbedingung heisst Anfangswertproblem. In unserem Fall hat das Anfangswertproblem die Lösung

$$x(t) = x_0\mathrm{e}^{a(t-t_0)}.$$

[5]Nach der üblichen Notation schreibt man explizit die unabhängige Variable t nur als Argument der Funktionen a und b, aber nicht der unbekannten Funktion x. Sind die Funktionen a oder b konstant, dann schreibt man das Argument nicht und es ist klar von der Formel, dass es sich um Konstanten händelt. Eine Lösung zur Differenzialgleichung wird doch als Funktion $x(t)$ geschrieben.

Das allgemeinste Beispiel einer linearen Differenzialgleichung erster Ordnung mit Konstanten Koeffizienten ist

$$\dot{x} = ax + b, \tag{4.5}$$

wobei a und b gegebene Konstanten sind. Ist $b \neq 0$, so heisst die Gleichung nichthomogen. Der Fall $b = 0$, d.h. (4.1), heisst die assoziierte homogene Differenzialgleichung.

Die nichthomogene Gleichung kann auch durch (4.4) aufgelöst werden; nämlich

$$\dot{X} = \mathrm{e}^{-at} b,$$

wenn x eine Lösung ist. Deshalb

$$X(t) = -\frac{1}{a}\mathrm{e}^{-at} b + K,$$

wobei K eine Konstante ist, und

$$x(t) = -\frac{b}{a} + K\mathrm{e}^{at},$$

wenn $a \neq 0$; sonst haben wir

$$x(t) = X(t) = bt + K.$$

Auch in diesem Fall legt die Anfangsbedingung $x(t_0) = x_0$ die Konstante K fest. Die Lösung ist dann

$$x(t) = -\frac{b}{a} + \left(x_0 + \frac{b}{a}\right)\mathrm{e}^{a(t-t_0)},$$

falls $a \neq 0$; sonst

$$x(t) = b(t - t_0) + x_0.$$

Bemerkung 4.2. Dieses Verfahren wirkt eigentlich auch im Falle, wenn b nicht konstant ist (was wir später brauchen werden), d.h. für die inhomogene Gleichung

$$\dot{x} = ax + b(t), \tag{4.6}$$

wobei $b(t)$ eine gegebene Funktion ist.[6] Nämlich folgt jetzt aus (4.4), dass

$$\dot{X} = \mathrm{e}^{-at} b(t),$$

wenn x eine Lösung zu (4.6) ist. Deshalb ist

$$X(t) = \int_0^t \mathrm{e}^{-as} b(s)\,\mathrm{d}s + K,$$

wobei K eine Integrationskonstante ist. Es folgt, dass

$$x(t) = \mathrm{e}^{at}\int_0^t \mathrm{e}^{-as} b(s)\,\mathrm{d}s + K\mathrm{e}^{at}.$$

Die Anfangsbedingung $x(t_0) = x_0$ legt die Konstante K fest:

$$K = \mathrm{e}^{-at_0} x_0 - \int_0^{t_0} \mathrm{e}^{-as} b(s)\,\mathrm{d}s$$

und wir können die entsprechende Lösung als

$$x(t) = \mathrm{e}^{a(t-t_0)} x_0 + \int_{t_0}^t \mathrm{e}^{a(t-s)} b(s)\,\mathrm{d}s$$

[6] Aus technischen Gründen nehmen wir an, dass b eine stetige Funktion ist.

schreiben.

Bemerkung 4.3. Ein sehr einfaches Speziallfall — noch einfacher als der Fall, in dem b konstant ist — kommt vor, wenn $b(t) = c\,\mathrm{e}^t$, wobei c eine Konstante ist. Wir überlassen dem Leser diese Übung.

4.2. **Systeme von linearen Differenzialgleichungen erster Ordnung mit konstanten Koeffizienten.** Wir wollen das System

$$\begin{aligned} \dot{z}_1 &= az_1 + bz_2 \\ \dot{z}_2 &= cz_1 + dz_2 \end{aligned} \tag{4.7}$$

auflösen, wobei z_1 und z_2 die unbekannten Funktionen und a, b, c, d gegebene Konstante sind. In Matrixnotation führen wir die unbekannte Verktorfunktion

$$\mathbf{z} := \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

und die Matrix der Koeffizienten

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

ein. Das System kann man dann als

$$\dot{\mathbf{z}} = \mathsf{A}\mathbf{z} \tag{4.8}$$

schreiben.

Bemerkung 4.4. Die Differenzialgleichung (4.3) für den gedämpften harmonischen Oszillator kann man als System erster Ordnung schreiben: Man führt die Geschwindigkeit v als neue unbekannte Funktion ein. Die Beziehung $\dot{x} = v$ sieht man als Differenzialgleichung. Die zweite Ableitung $\ddot{x}$ von x sieht man als erste Ableitung von v. Deshalb hat man das System

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= -\frac{k}{m}x - \frac{\gamma}{m}v \end{aligned} \tag{4.9}$$

Wir können das System in Matrixform wie in (4.8) schreiben durch

$$\mathbf{z} := \begin{pmatrix} x \\ v \end{pmatrix}$$

und

$$\mathsf{A} := \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\gamma}{m} \end{pmatrix}.$$

4.2.1. *Spezialfall 1: Diagonalmatrix.* Das System (4.8) kann man sehr einfach auflösen, wenn die Koeffizientenmatrix A eine $\mathsf{Diagonalmatrix}$ ist; d.h. wenn alle Einträge von A ausserhalb der Hauptdiagonale gleich Null sind. Wir nehmen nämlich an, dass A die Form

$$\boxed{\mathsf{A} := \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}} \tag{4.10}$$

hat, wobei λ_1 und λ_2 gegebene Konstanten sind. Das System (4.7) besteht dann aus zwei entkoppelten Differenzialgleichungen

$$\dot{z}_1 = \lambda_1 z_1$$
$$\dot{z}_2 = \lambda_2 z_2$$

die wir wie in 4.1 auflösen können:

$$z_1(t) = \mathrm{e}^{\lambda_1 t} K_1$$
$$z_2(t) = \mathrm{e}^{\lambda_2 t} K_2$$

wobei K_1 und K_2 Konstanten sind. Schreiben wir

$$\mathbf{K} := \begin{pmatrix} K_1 \\ K_2 \end{pmatrix}$$

und

$$\mathrm{e}^{\mathsf{A}t} := \begin{pmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{pmatrix},$$

so haben wir die allgemeine Lösung

$$\mathbf{z}(t) = \mathrm{e}^{\mathsf{A}t}\mathbf{K}.$$

4.2.2. *Spezialfall 2: Trigonalmatrix.* Ist die Koeffizientenmatrix A eine Trigonalmatrix (auch bekannt als obere Dreiecksmatrix), d.h. von der Form

$$\mathsf{A} = \begin{pmatrix} \lambda_1 & \mu \\ 0 & \lambda_2 \end{pmatrix},$$

wobei $\lambda_1, \lambda_2, \mu$ gegebene Konstanten sind, dann kann man das System (4.8), jetzt in der Form

$$\dot{z}_1 = \lambda_1 z_1 + \mu z_2,$$
$$\dot{z}_2 = \lambda_2 z_2,$$

auch einfach auflösen. Nämlich löst man zunächst die zweite Gleichung auf als

$$z_2(t) = \mathrm{e}^{\lambda_2 t} K_2,$$

wobei K_2 eine Konstante ist. Dann setzt man diese Lösung in die erste Gleichung ein:

$$\dot{z}_1 = \lambda_1 z_1 + \mu \mathrm{e}^{\lambda_2 t} K_2.$$

Mit

$$x := z_1 \qquad a := \lambda_1 \qquad b(t) := \mu \mathrm{e}^{\lambda_2 t} K_2,$$

ist das ein Beispiel einer ihnhomogenen Gleichung, die wir bereits in Bemerkung 4.2, aufgelöst haben.

Wir überlassen dem Leser die Berechnung im allgemeinen Falle und konzentrieren uns auf dem wichtigen Fall, wenn $\lambda_1 = \lambda_2$ und $\mu = 1$ gelten. Setzen wir $\lambda := \lambda_1 = \lambda_2$, so hat die Koeffizientenmatrix die Form

$$\boxed{\mathsf{A} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}}. \tag{4.11}$$

Die Lösung zur zweiten Gleichung ist jetzt $z_2(t) = \mathrm{e}^{\lambda t} K_2$ und die erste Gleichung nimmt die Form[7] $\dot{z}_1 = \lambda z_1 + \mathrm{e}^{\lambda t} K_2$ an. Setzen wir, wie in (4.4), $X(t) = \mathrm{e}^{-\lambda t} z_1(t)$, so haben wir

$$\dot{X}(t) = K_2,$$

wenn z_1 eine Lösung ist. Deshalb

$$X(t) = K_2 t + K_1,$$

wobei K_1 eine neue Konstante ist. Schliesslich erhalten wir

$$\begin{aligned} z_1(t) &= \mathrm{e}^{\lambda t} K_1 + t\,\mathrm{e}^{\lambda t} K_2, \\ z_2(t) &= \mathrm{e}^{\lambda t} K_2, \end{aligned}$$

oder in Matrixform

$$\mathbf{z}(t) = \mathrm{e}^{\mathsf{A}t} \mathbf{K},$$

mit

$$\mathbf{K} := \begin{pmatrix} K_1 \\ K_2 \end{pmatrix}$$

und

$$\mathrm{e}^{\mathsf{A}t} := \begin{pmatrix} \mathrm{e}^{\lambda t} & t\,\mathrm{e}^{\lambda t} \\ 0 & \mathrm{e}^{\lambda t} \end{pmatrix}.$$

4.3. **Ähnliche Matrizen.** Unser nächstes Ziel ist zu zeigen, dass man immer zu einem der Spezialfälle (4.10) oder (4.11) (die sog. JORDANschen Normalformen) reduzieren kann, wenn man mit komplexen Matrizen arbeitet (d.h. Matrizen mit komplexen Einträgen, s. Anhang B). Damit werden wir auch für eine beliebige Matrix A das Exponential $\mathrm{e}^{\mathsf{A}t}$, das auch eine Matrix ist, definieren und zeigen, dass sich die Lösung des linearen Systems von Differenzialgleichungen erster Ordnung mit konstanten Koeffizienten (4.8) immer als

$$\mathbf{z}(t) = \mathrm{e}^{\mathsf{A}t} \mathbf{K},$$

schreiben lässt.

Die Idee ist, wie im Falle von linearen Gleichungssystemen, eine Neudefinition

$$\mathbf{y}(t) := \mathsf{B}^{-1} \mathbf{z}(t)$$

zu verwenden, wobei B eine invertierbare Matrix mit konstanten Einträgen ist. Aus $\dot{\mathbf{z}} = \mathsf{B}\dot{\mathbf{y}}$ erhalten wir

$$\dot{\mathbf{y}} = \mathsf{B}^{-1} \mathsf{A} \mathsf{B} \mathbf{y},$$

falls $\mathbf{z}$ (4.8) auflöst. In anderen Worten sollen wir jetzt das äquivalente System

$$\dot{\mathbf{y}} = \tilde{\mathsf{A}} \mathbf{y}$$

auflösen, wobei

$$\tilde{\mathsf{A}} := \mathsf{B}^{-1} \mathsf{A} \mathsf{B}$$

ist. Das motiviert folgende

Definition 4.5. Zwei Matrizen A und $\tilde{\mathsf{A}}$ heissen ähnlich, wenn es eine invertierbare Matrix B gibt mit

$$\boxed{\tilde{\mathsf{A}} = \mathsf{B}^{-1} \mathsf{A} \mathsf{B}}.$$

[7] Eigentlich ist das die Übung in Bemerkung 4.3.

Bemerkung 4.6. Die Ähnlichkeit von Matrizen ist auch eine Äquivalenzrelation, wie man einfach zeigt (s. den Beweis, dass die in 3.34 definierte Äquivalenz von Matrizen eine Äquivalenzrelation ist.)

Bemerkung 4.7. Zwei ähnliche Matrizen sind auch äquivalent im Sinne von Definition 3.34, aber die Umkehrung gilt nicht. I.A. können wir nicht durch Ähnlichkeit eine Matrix zur im Satz 3.36 betrachtete Normalform überführen. Wir werden aber zeigen, dass jede komplexe Matrix zu einer Matrix in JORDANscher Normalform — d.h. (4.10) oder (4.11) — ähnlich ist. Das wird uns erlauben, ein allgemeines lineares System von Differenzialgleichungen aufzulösen.

Bemerkung 4.8. Zwei ähnliche Matrizen haben die gleiche Determinante:

$$\boxed{\det \mathsf{A} = \det \tilde{\mathsf{A}}},$$

wenn A und $\tilde{\mathsf{A}}$ ähnlich sind. Eine zweite Invariante ist die Spur, die als die Summe der Diagonaleinträge definiert ist: d.h.[8]

$$\mathrm{Sp}\begin{pmatrix} a & b \\ c & d \end{pmatrix} := a + d.$$

Durch eine einfache Berechnung zeigt man, dass[9]

$$\boxed{\mathrm{Sp}\,\mathsf{AB} = \mathrm{Sp}\,\mathsf{BA}}$$

für je zwei Matrizen A und B. Daraus folgt

$$\boxed{\mathrm{Sp}\,\mathsf{A} = \mathrm{Sp}\,\tilde{\mathsf{A}}},$$

wenn A und $\tilde{\mathsf{A}}$ ähnlich sind. Äquivalente Matrizen haben i.A. nicht die gleiche Spur.

4.4. **Diagonalisierung.** Wir wollen zunächst den wichtigen Fall studieren, wenn eine Matrix in Diagonalform überführt werden kann.

Definition 4.9. Eine Matrix A heisst diagonalisierbar, wenn sie zu einer Diagonalmatrix ähnlich ist, d.h. wenn es eine invertierbare Matrix B gibt mit

$$\mathsf{A} = \mathsf{B}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\mathsf{B}^{-1}.$$

In diesem Falle können wir das System (4.8) sehr einfach auflösen. Wir führen nämlich $\mathbf{y}(t) := \mathsf{B}^{-1}\mathbf{z}(t)$ ein und bekommen

$$\dot{\mathbf{y}} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\mathbf{y}.$$

Aus dem Spezialfall 1, Abschnitt 4.2.1, kennen wir schon die allgemeine Lösung:

$$\mathbf{y}(t) = \begin{pmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{pmatrix}\mathbf{d},$$

[8] Die Spur wird auch mit Tr, aus dem englischen „trace“, gekennzeichnet.

[9] Eine zweite sehr nützliche Eigenschaft der Spur ist ihre Linearität: nämlich gelten

$$\mathrm{Sp}\,\lambda\mathsf{A} = \lambda\,\mathrm{Sp}\,\mathsf{A} \qquad \text{und} \qquad \mathrm{Sp}(\mathsf{A}+\mathsf{B}) = \mathrm{Sp}\,\mathsf{A} + \mathrm{Sp}\,\mathsf{B}$$

für alle Matrizen A und B und alle Zahlen λ.

wobei $\mathbf{d} = \binom{d_1}{d_2}$ und d_1, d_2 Konstanten sind. Da $\mathbf{z}(t) = \mathsf{B}\mathbf{y}(t)$ gilt, haben wir schliesslich

$$\mathbf{z}(t) = \mathsf{B}\begin{pmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{pmatrix}\mathbf{d}.$$

Da B invertierbar ist, können wir statt $\mathbf{d}$ auch $\mathbf{c} := \mathsf{B}\mathbf{d}$ betrachten. Dann lautet die allgemeine Lösung

$$\boxed{\mathbf{z}(t) = \mathsf{B}\begin{pmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{pmatrix}\mathsf{B}^{-1}\mathbf{c}}$$

mit $\mathbf{c} = \binom{c_1}{c_2}$, wobei c_1, c_2 Konstanten sind. Kennt man die Matrix B, so kann man sehr einfach die Lösung durch Ausmultiplizieren finden. Es ist das Ziel dieses Abschnitts, eine solche Matrix B zu finden (eigentlich ist B nicht eindeutig bestimmt).

Wir beginnen aber mit einer negativen Ergebnis: *Nicht jede Matrix ist diagonalisierbar.* Als Beispiel betrachten wir die Matrix

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

Sie entspricht dem Fall $\gamma = k = 0$ in (4.9). Da $\det \mathsf{A} = 0 = \operatorname{Sp} \mathsf{A}$ gilt, haben wir $\det \tilde{\mathsf{A}} = 0 = \operatorname{Sp} \tilde{\mathsf{A}}$ für jede zu A ähnliche Matrix $\tilde{\mathsf{A}}$. Wäre A diagonalisierbar, so könnten wir eine zu A ähnliche Matrix $\tilde{\mathsf{A}} = \left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)$ finden. Aber dann hätten wir $\lambda_1\lambda_1 = \det \tilde{\mathsf{A}} = 0$ und $\lambda_1 + \lambda_2 = \operatorname{Sp} \tilde{\mathsf{A}} = 0$, d.h. $\lambda_1 = \lambda_2 = 0$. Deshalb wäre A die Nullmatrix $\mathbf{0}$. Aber $\mathsf{B}\mathbf{0}\mathsf{B}^{-1} = \mathbf{0}$ für jede invertierbare Matrix B. Das ist ein Widerspruch, denn A ist nicht die Nullmatrix.

Von jetzt an nehmen wir an, dass A diagonalisierbar ist. Wir wollen B bestimmen. Damit werden wir auch Kriterien zur Diagonalisierbarkeit gewinnen: Zu lösen ist die Gleichung

$$\mathsf{A}\mathsf{B} = \mathsf{B}\tilde{\mathsf{A}} \tag{4.12}$$

mit

$$\tilde{\mathsf{A}} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Wir schreiben

$$\mathsf{B} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix},$$

wobei $\mathbf{v}_1$ und $\mathbf{v}_2$ Vektoren sind. Dann gilt

$$\mathsf{A}\mathsf{B} = \begin{pmatrix} \mathsf{A}\mathbf{v}_1 & \mathsf{A}\mathbf{v}_2 \end{pmatrix}.$$

Wir schreiben jetzt

$$\mathbf{v}_1 = \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix},$$

und erhalten

$$\mathsf{B}\tilde{\mathsf{A}} = \begin{pmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{21} \\ \lambda_1 v_{12} & \lambda_2 v_{22} \end{pmatrix} = \begin{pmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 \end{pmatrix}.$$

Deshalb kann (4.12) als

$$\begin{pmatrix} \mathsf{A}\mathbf{v}_1 & \mathsf{A}\mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 \end{pmatrix}$$

geschrieben werden, oder äquivalent

$$\mathsf{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1,$$
$$\mathsf{A}\mathbf{v}_2 = \lambda_2\mathbf{v}_2.$$

Da B invertierbar ist, sind $\mathbf{v}_1$ und $\mathbf{v}_2$ von Null verschieden (sonst wäre $\det\mathsf{B} = 0$). Das motiviert folgende

Definition 4.10. Ein *von Null verschiedener* Vektor $\mathbf{v}$ heisst Eigenvektor von A zum Eigenwert λ, wenn

$$\boxed{\mathsf{A}\mathbf{v} = \lambda\mathbf{v}}$$

gilt.

Das Problem der Diagonalisierung von A besteht dann darin, die Eigenwerte λ_1 und λ_2 von A und die enstprechenden Eigenvektoren $\mathbf{v}_1$ und $\mathbf{v}_2$ zu finden, sodass $\mathsf{B} = \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{pmatrix}$ invertierbar ist, d.h. $\det\mathsf{B} \neq 0$.

Definition 4.11. Zwei Vektoren $\mathbf{v}_1$ und $\mathbf{v}_2$ heissen linear unabhängig, wenn

$$\det\begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{pmatrix} \neq 0$$

gilt.

Bemerkung 4.12. Diese Determinante entspricht dem sog. Vektorprodukt von 2-Vektoren:

$$\mathbf{v}_1 \times \mathbf{v}_2 := \det\begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{pmatrix}.$$

D.h.

$$\begin{pmatrix}u\\v\end{pmatrix} \times \begin{pmatrix}w\\z\end{pmatrix} = uz - vw.$$

Hilfssatz 4.13. *Zwei Vektoren sind genau dann linear unabhängig, wenn*

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = 0 \Rightarrow x_1 = x_2 = 0.$$

Beweis. Wir setzen $\mathsf{B} := \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{pmatrix}$ und $\mathbf{x} := \left(\begin{smallmatrix}x_1\\x_2\end{smallmatrix}\right)$. Dann

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = \mathsf{B}\mathbf{x}.$$

Den Satz kann man dann wie folgt neu fassen:

$$\det\mathsf{B} \neq 0 \Leftrightarrow (\mathsf{B}\mathbf{x} = 0 \Rightarrow \mathbf{x} = 0)$$

Diese Aussage ist nach der Diskussion im Abschnitt 3 offensichtlich wahr. Gilt nämlich $\det\mathsf{B} \neq 0$, so hat die Gleichung $\mathsf{B}\mathbf{x} = 0$ nur die Lösung $\mathbf{x} = 0$. Ist andererseits $\det\mathsf{B} = 0$, so gibt es nichttriviale Lösungen und die Implikation $\mathsf{B}\mathbf{x} = 0 \Rightarrow \mathbf{x} = 0$ gilt nicht. □

Satz 4.14. *Seien $\mathbf{v}_1$ und $\mathbf{v}_2$ Eigenvektoren einer Matrix A zu den Eigenwerten λ_1 und λ_2. Wenn $\lambda_1 \neq \lambda_2$ gilt, sind $\mathbf{v}_1$ und $\mathbf{v}_2$ linear unabhängig.*

Beweis. Nehmen wir an, dass $\mathbf{v}_1$ und $\mathbf{v}_2$ linear abhängig sind, d.h. dass es x_1 und x_2 gibt mit

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = 0 \text{ und } \begin{pmatrix}x_1\\x_2\end{pmatrix} \neq \begin{pmatrix}0\\0\end{pmatrix}.$$

O.B.d.A. können wir dann $x_2 \neq 0$ annehmen. Das impliziert $\mathbf{v}_2 = -\frac{x_1}{x_2}\mathbf{v}_1$. Nach dem Definition vom Eigenvektor gilt $\mathbf{v}_2 \neq 0$ und deshalb haben wir auch $x_1 \neq 0$. Aber

$$\mathsf{A}\mathbf{v}_2 = \lambda_2\mathbf{v}_2 = -\lambda_2\frac{x_1}{x_2}\mathbf{v}_1$$

und

$$\mathsf{A}\mathbf{v}_2 = -\frac{x_1}{x_2}\mathsf{A}\mathbf{v}_1 = -\lambda_1\frac{x_1}{x_2}\mathbf{v}_1.$$

Da $\mathbf{v}_1 \neq 0$ gilt, erhalten wir

$$\lambda_2\frac{x_1}{x_2} = \lambda_1\frac{x_1}{x_2},$$

und da $x_1 \neq 0$ gilt, haben wir schliesslich $\lambda_2 = \lambda_1$, was der Voraussetzung widerspricht. □

Es folgt dann unmittelbar folgender

Satz 4.15. *Besitzt* A *zwei verschiedene Eigenwerte* λ_1 *und* λ_2*, so ist* A *diagonalisierbar: Nämlich gilt*

$$\mathsf{A} = \mathsf{B}\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}\mathsf{B}^{-1}$$

mit

$$\mathsf{B} = \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{pmatrix},$$

wobei $\mathbf{v}_1$ *bzw.* $\mathbf{v}_2$ *der Eigenvektor zu* λ_1 *bzw.* λ_2 *ist.*

4.4.1. *Die charakteristiche Gleichung.* Die Gleichung $\mathsf{A}\mathbf{v} = \lambda\mathbf{v}$ können wir auch als

$$(\mathsf{A} - \lambda\mathbf{1})\mathbf{v} = 0$$

schreiben.

Satz 4.16. λ *ist ein Eigenwert von* A *genau dann, wenn*

$$\det(\mathsf{A} - \lambda\mathbf{1}) = 0 \tag{4.13}$$

gilt.

Beweis. Ist λ ein Eigenwert, so gibt es einen Vektor $\mathbf{v} \neq 0$ mit $(\mathsf{A} - \lambda\mathbf{1})\mathbf{v} = 0$. Das impliziert, dass $\mathsf{A} - \lambda\mathbf{1}$ nicht invertierbar ist, und deshalb, dass ihre Determinante verschwindet. Verschwindet auf der anderen Seite die Determinante von $\mathsf{A} - \lambda\mathbf{1}$, so gibt es einen Vektor $\mathbf{v} \neq 0$ mit $(\mathsf{A} - \lambda\mathbf{1})\mathbf{v} = 0$, was impliziert, dass λ ein Eigenwert ist. □

Bemerkung 4.17. Das Polynom

$$\boxed{P_{\mathsf{A}}(x) := \det(\mathsf{A} - x\mathbf{1})}$$

heisst das charakteristische Polynom von A und $P_{\mathsf{A}}(\lambda) = 0$ heisst die charakteristiche Gleichung.

Das charakteristische Polynom kann man sehr einfach berechnen. Sei nämlich $\mathsf{A} = \left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)$. Dann gilt

$$P_{\mathsf{A}}(x) = \det\begin{pmatrix}a - x & b\\ c & d - x\end{pmatrix} = (a - x)(d - x) - bc,$$

d.h.,

$$P_{\mathsf{A}}(x) = x^2 - (a + d)x + ad - bc.$$

Das kann man auch wie folgt schreiben:

$$\boxed{P_{\mathsf{A}}(x) = x^2 - \operatorname{Sp}\mathsf{A}\, x + \det\mathsf{A}}. \tag{4.14}$$

Übung 4.18 (Satz von Cayley–Hamilton)**.** Sei B eine Matrix. Man definiert

$$P_{\mathsf{A}}(\mathsf{B}) := \mathsf{B}^2 - (\operatorname{Sp}\mathsf{A})\,\mathsf{B} + (\det\mathsf{A})\,\mathbf{1}.$$

Man zeige, dass

$$P_{\mathsf{A}}(\mathsf{A}) = \mathbf{0}$$

gilt. Das ist der Satz von Cayley–Hamilton im Falle von 2×2-Matrizen.

Bemerkung 4.19. Sind A und $\tilde{\mathsf{A}}$ ähnliche Matrizen, so gilt

$$\boxed{P_{\mathsf{A}} = P_{\tilde{\mathsf{A}}}}.$$

Nämlich folgt aus $\tilde{\mathsf{A}} = \mathsf{BAB}^{-1}$, dass

$$\begin{aligned} P_{\tilde{\mathsf{A}}}(x) &= \det(\tilde{\mathsf{A}} - x\mathbf{1}) = \det(\mathsf{BAB}^{-1} - x\mathsf{B}\mathbf{1}\mathsf{B}^{-1}) = \\ &= \det\mathsf{B}(\mathsf{A} - x\mathbf{1})\mathsf{B}^{-1} = \det\mathsf{B}\det(\mathsf{A} - x\mathbf{1})(\det\mathsf{B})^{-1} = P_{\mathsf{A}}(x). \end{aligned}$$

Der erste Schritt zur Diagonalisierung einer Matrix A ist dann die Auflösung ihrer charakterischen Gleichung. Hat sie keine reellen Lösungen, so ist die Matrix nicht diagonalisierbar im reellen Sinne. Um dieses Problem zu umgehen, werden wir von jetzt an, nur komplexe Matrizen und Vektoren betrachten. (S. Anhang B für weitere Informationen über die komplexen Zahlen.) Das ist auch für die Auflösung des Systems (4.7) mit reellen Koeffizienten nötig.

Bemerkung 4.20. Hat eine Matrix A reelle Eigenwerte λ_1 und λ_2, so haben die Gleichungssysteme $(\mathsf{A} - \lambda_1\mathbf{1})\mathbf{v}_1 = 0$ und $(\mathsf{A} - \lambda_2\mathbf{1})\mathbf{v}_2 = 0$ reelle Koeffizienten und deshalb kann man reelle Lösungen $\mathbf{v}_1$ und $\mathbf{v}_2$ finden. Die Matrix $\mathsf{B} = \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{pmatrix}$ ist deshalb reell. Die Matrix A ist also diagonalisierbar im reellen Sinne.

Bemerkung 4.21. Die Koeffizientenmatrix für den harmonischen Oszillator (4.9) ohne Reibung ist $\mathsf{A} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{pmatrix}$ Ihr charakteristisches Polynom ist dann $P_{\mathsf{A}}(x) = x^2 + \frac{k}{m}$, das keine reellen Wurzeln im physikalischen Falle $k > 0, m > 0$ besitzt.

Wir wollen zunächst die Differenzialgleichung $\dot{z} = az$ auflösen, wobei a eine komplexe Zahl ist und die unbekannte Funktion z komplexe Werte annimmt.

Bemerkung 4.22 (Das komplexe Exponential)*.* Für u und v reell führt man das komplexe Exponential

$$\mathrm{e}^{u+\mathrm{i}v} := \mathrm{e}^u(\cos v + \mathrm{i}\sin v)$$

ein. Dank der Regel für die Ableitung eines Produktes erhält man

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{(u+\mathrm{i}v)t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{ut}(\cos vt + \mathrm{i}\sin vt)\right) = \\ &= u\,\mathrm{e}^{ut}(\cos vt + \mathrm{i}\sin vt) + \mathrm{e}^{ut}(-v\sin vt + \mathrm{i}v\cos vt) = \\ &= (u + \mathrm{i}v)\,\mathrm{e}^{ut}(\cos vt + \mathrm{i}\sin vt) = (u + \mathrm{i}v)\,\mathrm{e}^{(u+\mathrm{i}v)t}. \end{aligned}$$

Deshalb hat die Differenzialgleichung $\dot{z} = az$ auch im komplexen Fall die Lösung $K\mathrm{e}^{at}$, wobei jetzt K eine komplexe Konstante ist.

Bemerkung 4.23 (Auflösung von linearen Systemen von Differenzialgleichungen durch komplexe Diagonalisierung). Sei $\dot{z} = \mathsf{A}z$ unser System, wobei A eine reelle Matrix ist. Wir nehmen an, A ist diagonalisierbar im komplexen Sinne:

$$\mathsf{A} = \mathsf{B}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\mathsf{B}^{-1},$$

wobei λ_1, λ_2 und B komplex sein könenn (es gilt aber $\bar{\lambda}_1 = \lambda_2$, denn das charakteristische Polynom von A besitzt reelle Koeffizienten). Damit bekommen wir die allgemeine komplexwertige Lösung

$$z(t) = \mathsf{B}\begin{pmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{pmatrix}\mathsf{B}^{-1}\mathbf{c},$$

wobei $\mathbf{c}$ ein komplexer Vektor ist.

Satz 4.24. *Sei A eine reelle Matrix, die diagonalisierbar im komplexen Sinne ist. Die allgemeine Lösung zur reellen Differenzialgleichung $\dot{z} = \mathsf{A}z$ ist dann*

$$z(t) = \mathsf{B}\begin{pmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{pmatrix}\mathsf{B}^{-1}\mathbf{K},$$

wobei $\mathbf{K}$ ein reeller Vektor ist.

In anderen Worten ist $\mathsf{B}\left(\begin{smallmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{smallmatrix}\right)\mathsf{B}^{-1}$ eine reelle Matrix, obwohl B und $\left(\begin{smallmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{smallmatrix}\right)$ nicht reell sind.

Beweis. Da A reell ist, haben wir $\dot{\bar{z}} = \mathsf{A}\bar{z}$. Ist $z(t)$ eine Lösung, dann ist $\bar{z}(t)$ auch eine Lösung. Da

$$\bar{z}(t) = \bar{\mathsf{B}}\begin{pmatrix} \mathrm{e}^{\bar{\lambda}_1 t} & 0 \\ 0 & \mathrm{e}^{\bar{\lambda}_2 t} \end{pmatrix}\bar{\mathsf{B}}^{-1}\bar{\mathbf{c}},$$

eine Lösung ist, gibt es ein komplexer Vektor $\mathbf{d}$, sodass

$$\bar{z}(t) = \mathsf{B}\begin{pmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{pmatrix}\mathsf{B}^{-1}\mathbf{d}$$

gilt. Setzen wir $t = 0$ ein, so erhalten wir $\bar{\mathbf{c}} = \mathbf{d}$. Suchen wir nach einer reellen Lösung, $\bar{z}(t) = z(t)$, so muss $\mathbf{c} = \mathbf{d}$ gelten. Daraus folgt, dass die Lösung $z(t)$ genau dann reell ist, wenn $\bar{\mathbf{c}} = \mathbf{c}$ gilt, d.h. wenn $\mathbf{c}$ reell ist. □

4.5. **Trigonalisierung.** Nehmen wir jetzt an, dass das charakteristische Polynom von A zwei gleiche Wurzeln λ besizt. Kann A diagonalisierbar sein? Wäre das der Fall, so hätten wir, für irgendeine Matrix B,

$$\mathsf{A} = \mathsf{B}\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\mathsf{B}^{-1} = \lambda\mathsf{B}\mathbf{1}\mathsf{B}^{-1} = \lambda\mathbf{1} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}.$$

D.h. A hätte bereits Diagonal sein müssen:

Satz 4.25. *Eine Matrix mit gleichen Eigenwerten ist genau dann diagonalisierbar, wenn sie diagonal ist.*

Der nichttriviale Fall ist im folgenden Satz beschrieben:

Satz 4.26. *Sei* A *eine nichtdiagonale Matrix mit gleichen Eigenwerten* λ*. Dann ist* A *zur Matrix*
$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$
ähnlich.

Beweis. Das charakteristische Polynom von A ist nach Annahme $P_{\mathsf{A}}(x) = (x-\lambda)^2$. Wegen (4.14) gelten dann $\operatorname{Sp}\mathsf{A} = 2\lambda$ und $\det\mathsf{A} = \lambda^2$. Dank Übung 4.18 (Satz von Cayley–Hamilton) gilt aber dann
$$\mathsf{A}^2 - 2\lambda\mathsf{A} + \lambda^2\mathbf{1} = 0,$$
d.h.
$$(\mathsf{A} - \lambda\mathbf{1})^2 = 0.$$
Sei jetzt $\mathbf{v}$ ein Eigenvektor zu λ. Da nach Annahme $\mathsf{A} - \lambda\mathbf{1} \neq \mathbf{0}$ gilt, gibt es ein Vektor $\mathbf{w}$ so, dass $(\mathsf{A} - \lambda\mathbf{1})\mathbf{w} \neq 0$. D.h. $(\mathsf{A} - \lambda\mathbf{1})\mathbf{w}$ ist auch ein Eigenvektor zu λ. Deshalb muss es zu $\mathbf{v}$ proportionell sein (sonst wäre A diagonalisierbar, was wir nach Annahme ausgeschlossen haben). Bis auf Skalierung von $\mathbf{w}$ können wir dann annehmen, dass
$$(\mathsf{A} - \lambda\mathbf{1})\mathbf{w} = \mathbf{v} \tag{4.15}$$
gilt. Mit $\mathsf{B} := \begin{pmatrix} \mathbf{v} & \mathbf{w} \end{pmatrix}$ bekommen wir dann,
$$\mathsf{AB} = \mathsf{B}\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Noch zu zeigen ist, dass B invertierbar ist, oder äquivalent, dass $\mathbf{v}$ und $\mathbf{w}$ linear unabhängig sind (Definition 4.11). Das zeigen wir mit der Hilfe von Hilfssatz 4.13. Seien nämlich x_1 und x_2 Skalare, sodass
$$x_1\mathbf{v} + x_2\mathbf{w} = 0 \tag{4.16}$$
gilt. Man bemerkt zunächst, dass $\mathbf{v}$ nach der Definition vom Eigenvektor ungleich Null ist und dass auch $\mathbf{w}$ nicht verschwinden kann, sonst wäre aufgrund von (4.15) $\mathbf{v}$ gleich Null. Dass bedeutet, dass entweder $x_1 = x_2 = 0$ gilt, was zu zeigen ist, oder beide x_1 und x_2 von Null verschieden sind. Um den zweiteren Fall auszuschliessen, nehmen wir an, dass $x_1 \neq 0$ gilt. Dividieren wir (4.16) durch x_1 und verwenden wir (4.15), so erhalten wir
$$\left((\mathsf{A} - \lambda\mathbf{1}) + \frac{x_2}{x_1}\mathbf{1}\right)\mathbf{w} = 0.$$
Da $\mathbf{w}$ von Null verschieden ist, ist die Matrix $\mathsf{A} - \lambda\mathbf{1} + \frac{x_2}{x_1}\mathbf{1}$ nicht invertierbar. Deshalb verschwindet ihre Determinante, was impliziert, dass $\lambda - \frac{x_2}{x_1}$ ein Eigenwert von A ist. Da aber λ der einzige Eigenwert ist, muss x_2 verschwinden, was ein Widerspruch ist. □

Satz 4.27. *Ist die Matrix* A *im Satz 4.26 reell, so gilt*
$$\mathsf{A} = \mathsf{B}\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}\mathsf{B}^{-1}$$
mit λ *und* B *reell.*

Beweis. Sind die Wurzeln einer quadratischen Gleichung mit reellen koeffizienten gleich, so müssen sie reell sein. Deshalb ist λ reell. Sei $\mathbf{v}$ ein reeller Eigenvektor zu λ. Er existiert als Lösung des Gleichungssystem mit reellen Koeffizienten $(\mathsf{A} - \lambda\mathbf{1})\mathbf{v} = 0$. Wir suchen dann nach einem $\mathbf{w}$, sodass $(\mathsf{A} - \lambda\mathbf{1})\mathbf{w} = \mathbf{v}$ gilt. Da diese Gleichung reelle erweiterte Koeffizienten besitzt, können wir eine reelle Lösung finden. □

Bemerkung 4.28 (Auflösung von linearen Systemen von Differenzialgleichungen durch Trigonalisierung)*.* Ist die Matrix A nicht diagonalisierbar im komplexen Sinne, dann kann sie zur Form (4.11) überführt werden und das System $\dot{z} = \mathsf{A}z$ kann man wie in (4.2.2) auflösen.

4.6. **Zusammenfassung.** Wir können Sätze 4.15, 4.16, 4.25 und 4.26 zusammenfassen:

Lehrsatz 4.29. *Jede komplexe Matrix ist trigonalisierbar, d.h. ähnlich zu einer oberen Dreiecksmatrix. Sind ihre Eigenwerte voneinander verschieden, so ist sie sogar diagonalisierbar. Die Normalformen bez. Ähnlichkeit, die sog.* JORDAN*schen Normalformen, sind*

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \qquad \textit{und} \qquad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Eine reelle Matrix ist genau dann trigonalisierbar, wenn sie reelle Eigenwerte besitzt. Eine reelle Matrix, die nicht diagonalisierbar im komplexen Sinne ist, ist trigonalisierbar im reellen Sinne.

Bemerkung 4.30*.* Sind λ_1 und λ_2 voneinander verschieden, so sind die Matrizen $\left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{smallmatrix}\right)$ auch voneiander verschieden. Sie sind aber zueinander ähnlich:

$$\begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Andererseits ist eine Diagonalmatrix $\left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)$ zu einer Trigonalmatrix $\left(\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right)$ nicht ähnlich.

4.6.1. *Schritte zur Trigonalisierung.* Wir fassen die Schritte zusammen, die zur Trigonalisierung einer gegebenen 2×2-Matrix A führen:

(1) Man schreibt das charakterische Polynom P_{A} von A, (4.14).
(2) Man findet die Eigenwerte von A, d.h. die Wurzeln des charakteristichen Polynoms (Auflösung einer quadratischen Gleichung).
(3) Sind die Eigenwerte λ_1 und λ_2 voneinander verschieden, so sucht man nach entsprechenden Eigenvektoren, d.h. nach von Null verschiedenen Lösungen $\mathbf{v}_1$ und $\mathbf{v}_2$ der Systeme $\mathsf{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ und $\mathsf{A}\mathbf{v}_2 = \lambda_1\mathbf{v}_2$. Die Matrix A ist dann zur Diagonalmatrix $\left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)$ ähnlich:

$$\mathsf{A} = \mathsf{B} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathsf{B}^{-1}$$

mit $\mathsf{B} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix}$. Ist die Matrix A reell und sind ihre verschiedenen Eigenwerte reell, so kann man reelle Eigenvektoren finden.
(4) Hat A zwei gleiche Eigenwerte λ, so sucht man nach einem entsprechenden Eigenvektor $\mathbf{v}$, d.h. nach einer von Null verschiedenen Lösung $\mathbf{v}$ des Systems

$\mathsf{A}\mathbf{v} = \lambda\mathbf{v}$, und nach einer Lösung $\mathbf{w}$ des Systems $(\mathsf{A}-\lambda\mathbf{1})\mathbf{w} = \mathbf{v}$. Die Matrix A ist dann zur Trigonalmatrix $\left(\begin{smallmatrix}\lambda & 1\\ 0 & \lambda\end{smallmatrix}\right)$ ähnlich:

$$\mathsf{A} = \mathsf{B}\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix}\mathsf{B}^{-1}$$

mit $\mathsf{B} = \begin{pmatrix}\mathbf{v} & \mathbf{w}\end{pmatrix}$. Ist die Matrix A reell, so ist λ auch reell und man kann reelle $\mathbf{v}$ und $\mathbf{w}$ finden.

4.7. **Die Exponentialabbildung.** Sei A eine (reelle oder komplexe) Matrix. Mit $\mathrm{e}^{\mathsf{A}t}$ bezeichnet man die Lösung an der Zeit t zum folgenden Matrixanfangswertproblem:

$$\begin{aligned}\dot{\Phi} &= \mathsf{A}\Phi,\\ \Phi(0) &= \mathbf{1},\end{aligned} \tag{4.17}$$

wobei die Unbekannte $\Phi(t)$ eine matrixwertige Funktion ist. Durch die so definierte[10] Exponentialbbildung $t \mapsto \mathrm{e}^{\mathsf{A}t}$ können wir jedes Anfangswertproblem

$$\begin{aligned}\dot{\mathbf{z}} &= \mathsf{A}\mathbf{z},\\ \mathbf{z}(t_0) &= \mathbf{z}_0,\end{aligned}$$

wobei t_0 eine gegebene Zahl und $\mathbf{z}_0$ ein gegebener Vektor ist, als $\mathbf{z}(t) = \mathrm{e}^{\mathsf{A}(t-t_0)}\mathbf{z}_0$ auflösen.

Das Problem (4.17) löst man einfach auf, indem man A zur Jordanschen Normalform überführt: Wir schreiben $\mathsf{A} = \mathsf{B}\mathsf{R}\mathsf{B}^{-1}$, mit B invertierbar, wobei

$$\mathsf{R} = \begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix},$$

falls A diagonalisierbar ist, und

$$\mathsf{R} = \begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix},$$

andernfalls. Wir setzen $\Psi(t) := \mathsf{B}^{-1}\Phi(t)\mathsf{B}$. Ist Φ die Lösung zum Problem (4.17), so erfüllt Ψ das Problem

$$\begin{aligned}\dot{\Psi} &= \mathsf{R}\Psi,\\ \Psi(0) &= \mathbf{1}.\end{aligned}$$

Schreiben wir $\Psi = \begin{pmatrix}\mathbf{z}_1 & \mathbf{z}_2\end{pmatrix}$, so erhalten wir die Anfangswertprobleme

$$\begin{aligned}\dot{\mathbf{z}}_1 &= \mathsf{R}\mathbf{z}_1, & \dot{\mathbf{z}}_2 &= \mathsf{R}\mathbf{z}_2,\\ \mathbf{z}_1(0) &= \begin{pmatrix}1\\0\end{pmatrix}, & \mathbf{z}_2(0) &= \begin{pmatrix}0\\1\end{pmatrix}.\end{aligned}$$

[10]Man kann auch explizit zeigen, dass

$$\mathrm{e}^{\mathsf{A}t} = \sum_{n=0}^{\infty}\frac{t^n}{n!}\mathsf{A}^n$$

gilt, wobei die Reihe absolut $\forall t$ und gleichmässig auf jedem kompakten Teilintervall von $\mathbb{R}$ konvergiert. Dazu braucht man aber die entsprechenden Begriffe aus der Analysis, die wir hier nicht voraussetzen wollen.

Im Diagonalfall haben wir dann (s. Abschnitt 4.2.1)

$$\Psi(t) = \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix}$$

und deshalb

$$\boxed{e^{\mathsf{A}t} = \mathsf{B} \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} \mathsf{B}^{-1}}. \tag{4.18}$$

Andernfalls bekommen wir (s. Abschnitt 4.2.2)

$$\Psi(t) = \begin{pmatrix} e^{\lambda t} & t\,e^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}$$

und deshalb

$$\boxed{e^{\mathsf{A}t} = \mathsf{B} \begin{pmatrix} e^{\lambda t} & t\,e^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix} \mathsf{B}^{-1}}. \tag{4.19}$$

Bemerkung 4.31. Ist die Matrix A reell, dann ist $e^{\mathsf{A}t}$, als Lösung zum Problem (4.17), auch reell.

Satz 4.32. *Für alle Matrizen A, C, mit C invertierbar, und alle Zahlen s, t gelten folgende Aussagen:*

$$\begin{aligned} e^{\mathsf{A}(t+s)} &= e^{\mathsf{A}t}\, e^{\mathsf{A}s}, \\ e^{\mathsf{CAC}^{-1}t} &= \mathsf{C}\, e^{\mathsf{A}t}\, \mathsf{C}^{-1}, \\ \det e^{\mathsf{A}t} &= e^{\operatorname{Sp} \mathsf{A}t}. \end{aligned}$$

Der Beweis anhand von (4.18) und (4.19) ist dem Leser überlassen.

Übung 4.33. Seien

$$\mathsf{A} := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{und} \qquad \mathsf{B} := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Man berechne $e^{\mathsf{A}t}$, $e^{\mathsf{B}t}$ und $e^{(\mathsf{A}+\mathsf{B})t}$, und man zeige, dass

$$e^{(\mathsf{A}+\mathsf{B})t} \neq e^{\mathsf{A}t}\, e^{\mathsf{B}t}.$$

5. Bilinearformen

Wir kommen jetzt langsam zum Schluss, indem wir das Skalarprodukt von Vektoren

$$\mathbf{v} \cdot \mathbf{x}$$

verallgemeinern.

Zur Erinnerung: Das Skalarprodukt von $\mathbf{v} = \left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right)$ und $\mathbf{x} = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ ist durch

$$\mathbf{v} \cdot \mathbf{x} = vx + wy$$

definiert.

Das Skalaprodukt kann man als Komposition der $\mathsf{Transposition}$

$$\boxed{\begin{pmatrix} v \\ w \end{pmatrix}^t := \begin{pmatrix} v & w \end{pmatrix}}$$

und der Multiplikation von Zeilen- und Spaltenvektoren

$$\boxed{\begin{pmatrix} v & w \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} := vx + wy}$$

auffassen:

$$\mathbf{v} \cdot \mathbf{x} = \mathbf{v}^t \mathbf{x}.$$

Die Transposition kann man zu Matrizen erweitern

$$\boxed{\begin{pmatrix} a & b \\ c & d \end{pmatrix}^t := \begin{pmatrix} a & c \\ b & d \end{pmatrix}},$$

sowie die Multiplikation von Zeilenvektoren und Matrizen:

$$\boxed{\begin{pmatrix} v & w \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} := \begin{pmatrix} va + wb & vc + wd \end{pmatrix}}.$$

Man kann einfach folgende Eigenschaften zeigen:

$$\begin{aligned} (\mathsf{A}\mathbf{v})^t &= \mathbf{v}^t \mathsf{A}^t, \\ (\mathsf{AB})^t &= \mathsf{B}^t \mathsf{A}^t, \\ \det \mathsf{A}^t &= \det \mathsf{A}, \\ \operatorname{Sp} \mathsf{A}^t &= \operatorname{Sp} \mathsf{A}, \end{aligned} \tag{5.1}$$

für alle Matrizen A, B und alle Vektoren $\mathbf{v}$.

Ist B eine Matrix, so definiert man

$$\boxed{\phi_{\mathsf{B}}(\mathbf{v}, \mathbf{w}) := \mathbf{v}^t \mathsf{B} \mathbf{w}}.$$

Z.B. ist das übliche Skalarprodukt $\mathbf{v} \cdot \mathbf{w} = \phi_{\mathbf{1}}(\mathbf{v}, \mathbf{w})$. Wegen (5.1) gilt auch

$$\phi_{\mathsf{B}}(\mathbf{v}, \mathbf{w}) := \mathbf{w}^t \mathsf{B}^t \mathbf{v}. \tag{5.2}$$

Nach der Diskussion in Abschnitt 3.6 folgt, dass

$$\phi_{\mathsf{B}}(\mathbf{v}, \lambda\mathbf{w} + \mu\mathbf{z}) = \lambda\phi_{\mathsf{B}}(\mathbf{v}, \mathbf{w}) + \mu\phi_{\mathsf{B}}(\mathbf{v}, \mathbf{w})$$

für alle Vektoren $\mathbf{v}, \mathbf{w}, \mathbf{z}$ und alle Skalare λ, μ. D.h. $\phi_{\mathsf{B}}(\mathbf{v}, \mathbf{w})$ ist linear bez. des zweiten Arguments. Dank (5.1) ist sie linear auch bez. des ersten Arguments:

$$\phi_{\mathsf{B}}(\lambda\mathbf{v} + \mu\mathbf{z}, \mathbf{w}) = \lambda\phi_{\mathsf{B}}(\mathbf{v}, \mathbf{w}) + \mu\phi_{\mathsf{B}}(\mathbf{z}, \mathbf{w})$$

für alle Vektoren $\mathbf{v}, \mathbf{w}, \mathbf{z}$ und alle Skalare λ, μ. Man sagt dann, ϕ_{B} sei eine Bilinearform.

Durch eine invertierbare Matrix S dürfen wir die Argumente einer Bilinearform umbenennen: $\mathbf{v}' := \mathsf{S}\mathbf{v}$ und $\mathbf{w}' := \mathsf{S}\mathbf{w}$. Wir haben

$$\phi_{\mathsf{B}}(\mathbf{v}', \mathbf{w}') = \phi_{\mathsf{S}^t\mathsf{B}\mathsf{S}}(\mathbf{v}, \mathbf{w}).$$

Das motiviert die Einführung der Äquivalenzrelation

$$\boxed{\tilde{\mathsf{B}} \sim \mathsf{B} \overset{\text{Def}}{\Leftrightarrow} \exists \mathsf{S} \in \mathrm{GL}_2 : \tilde{\mathsf{B}} = \mathsf{S}^t \mathsf{B} \mathsf{S}}. \tag{5.3}$$

Bemerkung 5.1. Dank (5.1) gilt $\det \tilde{\mathsf{B}} = \det \mathsf{B} (\det \mathsf{S})^2$. Arbeiten wir mit reellen Matrizen, so erhält die bereits eingeführte Äquivalenzrelation das Vorzeichen der Determinante.

Bemerkung 5.2. Die neue Äquivalenzrelation (5.3) ist anders von der in Definition 4.5 eingeführten Ähnlichkeit, wo die Inverse statt der transponierten Matrix vorkam.

Definition 5.3. Eine reelle Matrix S heisst orthogonal, wenn

$$\mathsf{S}^t = \mathsf{S}^{-1}.$$

Bemerkung 5.4. Sind B und $\tilde{\mathsf{B}}$ durch eine orthogonale Matrix äquivalent, so sind sie auch ähnlich.

Bemerkung 5.5. $\mathsf{S} = \begin{pmatrix}\mathbf{v} & \mathbf{w}\end{pmatrix}$ ist genau dann orthogonal, wenn

$$\mathbf{v}\cdot\mathbf{v} = 1, \quad \mathbf{v}\cdot\mathbf{w} = 0, \qquad \mathbf{w}\cdot\mathbf{w} = 1,$$

was die Bezeichnung „orthogonal“ erklärt.

Bemerkung 5.6. Wir können einen Vektor $\mathbf{v}$ von Norm 1 (d.h. $\mathbf{v}\cdot\mathbf{v} = 1$) durch

$$\mathbf{v} = \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}$$

parametrisieren. Ein zu $\mathbf{v}$ orthogonaler Vektor $\mathbf{w}$ (d.h. $\mathbf{v}\cdot\mathbf{w} = 0$) hat dann die Form

$$\mathbf{w} = a\begin{pmatrix}-\sin\theta \\ \cos\theta\end{pmatrix}.$$

Ist $\mathbf{w}$ auch von Norm 1, so ist $a = \pm 1$. Da die Determinante der orthogonalen Matrix $\mathsf{S} = \begin{pmatrix}\mathbf{v} & \mathbf{w}\end{pmatrix}$ gleich zu a ist, haben zwei Klassen von orthogonalen Matrizen. Eine orthogonale Matrix S, deren Determinante gleich 1 ist, heisst speziell ortogonal und hat notwendigerweise die Form

$$\mathsf{S} = \begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}.$$

Sie stellt eine Drehung im Gegenuhrzeigersinne mit Winkel θ dar. Ist auf der anderen Weise die Determinante einer orthogonalen Matrix S gleich -1, so können wir

$$\mathsf{S} = \begin{pmatrix}\cos\theta & \sin\theta \\ \sin\theta & -\cos\theta\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & -1\end{pmatrix}$$

schreiben, d.h. als Produkt einer Drehung mit einer Spiegelung an der y-Achse.

Definition 5.7. Eine reelle Matrix B heisst symmetrisch, wenn

$$\mathsf{B}^t = \mathsf{B}.$$

Die entsprechende Bilinearform ϕ_{B} erfüllt dann dank (5.2)

$$\phi_{\mathsf{B}}(\mathbf{w},\mathbf{v}) = \phi_{\mathsf{B}}(\mathbf{v},\mathbf{w}), \qquad \forall \mathbf{v},\mathbf{w},$$

und heisst auch symmetrisch.

Bemerkung 5.8. Ist ϕ_{B} symmetrisch, so ist man oft am geometrischen Ort

$$\phi_{\mathsf{B}}(\mathbf{x},\mathbf{x}) = a \tag{5.4}$$

interessiert, wobei a eine gegebene Konstante ist. Z.B. beschreibt (5.4) im Falle $\mathsf{B} = \mathbf{1}$ und $a = R^2$,

$$\mathbf{x}\cdot\mathbf{x} = R^2,$$

den Kreis mit Radius R und Zentrum im Ursprung.

Wir wollen jetzt die Normalformen einer symmetrischen Bilinearform finden und damit das Problem lösen, den geometrischen Ort (5.4) zu beschreiben.

Satz 5.9 (Hauptachsentransformation)**.** *Eine symmetrische Matrix ist immer diagonalisierbar und ihre Diagonalisierung kann durch eine speziell orthogonale Matrix durchgeführt werden. D.h.: ist* B *symmetrisch, so gibt es eine speziell orthogonale Matrix* S *mit*

$$\mathsf{B} = \mathsf{S}^{-1} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathsf{S} = \mathsf{S}^t \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathsf{S}.$$

Deshalb ist B *auch als Bilinearform diagonalisierbar. Die Eigenvektoren von* B *heissen ihre* Hauptachsen *und die speziell orthogonale Matrix* S *ist die Drehung, die die ursprünglichen Cartesischen Achsen auf die Hauptachsen abbildet.*

Beweis. Wir schreiben

$$\mathsf{B} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}.$$

Man bemerke, dass die Einträge ausserhalb der Diagonale gleich sein müssen, da B symmetrisch ist. Die charakterisische Gleichung ist dann

$$\lambda^2 - (a+d)\lambda + ad - b^2 = 0.$$

Ihre Diskriminante

$$\Delta = (a+d)^2 - 4ad + 4b^2 = (a-d)^2 + 4b^2$$

ist nicht negativ. Deshalb hat die Gleichung reelle Lösungen. Die sind genau dann gleich, wenn $a = d$ und $b = 0$ gelten; d.h., wenn B diagonal ist. In diesem Falle nehmen wir $\mathsf{S} = \mathbf{1}$.

Andernfalls hat B zwei verschiedene reelle Eigenwerte λ_1 und λ_2. Seien $\mathbf{v}_1$ und $\mathbf{v}_2$ entsprechende Eigenvektoren. Bis auf Skalierung könnnen wir

$$\mathbf{v}_1 \cdot \mathbf{v}_1 = 1 = \mathbf{v}_2 \cdot \mathbf{v}_2$$

annehmen. Dann gilt

$$\mathsf{B} = \mathsf{T} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathsf{T}^{-1}$$

mit $\mathsf{T} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix}$. Es gilt

$$\phi_{\mathsf{B}}(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_1^t \mathsf{B} \mathbf{v}_2 = \lambda_2 \mathbf{v}_1^t \mathbf{1} \mathbf{v}_1 = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2.$$

Ähnlicherweise haben wir

$$\phi_{\mathsf{B}}(\mathbf{v}_2, \mathbf{v}_1) = \lambda_1 \mathbf{v}_2 \cdot \mathbf{v}_1.$$

Aus der Symmetrie bekommen wir

$$\lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2 = \lambda_1 \mathbf{v}_2 \cdot \mathbf{v}_1,$$

d.h.

$$(\lambda_2 - \lambda_1)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0.$$

Da $\lambda_2 \neq \lambda_1$ gilt, bekommen wir

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = 0.$$

Das zeigt dank Bemerkung 5.5, dass T orthogonal ist. Wir können o.B.d.A. annehmen, dass $\det \mathsf{T} = 1$ gilt (sonst ersetzen wir $\mathbf{v}_2$ durch $-\mathbf{v}_2$), d.h. dass T speziell orthogonal ist. Schliesslich definieren wir $\mathsf{S} := \mathsf{T}^t = \mathsf{T}^{-1}$. □

Den geometrischen Ort (5.4) kann man dann so verstehen: Durch die Drehung S, $\binom{x}{y} := \mathsf{S}\mathbf{x}$, erhalten wir die Gleichung

$$\lambda_1 x^2 + \lambda_2 y^2 = a.$$

Sind λ_1, λ_2, a von Null verschieden, so beschreibt die Gleichung eine Ellipse, wenn alle drei das gleiche Vorzeichen haben, und eine Hyperbel, wenn $\lambda_1\lambda_2 < 0$ gilt. Die anderen Fälle überlassen wir dem Leser.

Schliesslich können wir die Normalformen einer symmetrischen Matrix klassifizieren:

Satz 5.10. *Eine symmetrische Matrix ist bis auf die Äquivalenzrelation* (5.3) *im reellen Sinne zu einer der folgenden zueinander nichtäquivalenten sechs Normalformen äquivalent:*

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

*In den ersten drei Fällen heisst die entsprechende symmetrische Bilinearform **entartet**, im vierten Fall **positiv definit**, im Fünften **nicht definit** und im Sechsten **negativ definit**.*

Beweis. Laut Satz (5.9) ist jede symmetrische Matrix zu einer Diagonalmatrix im Sinne von Bilinearformen äquivalent. Es ist noch zu zeigen, dass jede Diagonalmatrix äquivalent zu einer der sechs Normalformen ist. Sei also

$$\mathsf{B} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Durch

$$\mathsf{R} = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}, \qquad \alpha_1\alpha_2 \neq 0,$$

führen wir B zu

$$\tilde{\mathsf{B}} = \mathsf{R}^t\mathsf{B}\mathsf{R} = \begin{pmatrix} \alpha_1^2\lambda_1 & 0 \\ 0 & \alpha_2^2\lambda_2 \end{pmatrix}$$

über. Ist $\lambda_i > 0$, so können wir $\alpha_i^2\lambda_i = 1$ erhalten, indem wir $\alpha_i = 1/\sqrt{\lambda_i}$ definieren. Ist $\lambda_i < 0$, so können wir $\alpha_i^2\lambda_i = -1$ durch die Definition $\alpha_i = 1/\sqrt{-\lambda_i}$ erhalten. Ist schliesslich $\lambda_i = 0$, se verändern wir es nicht mit $\alpha_i = 1$.

Damit haben wir bewiesen, dass jede symmetrische Matrix im Sinne von Bilinearformen zu einer Diagonalmatrix mit Einträgen gleich zu $\pm 1, 0$ äquivalent ist. Durch

$$\mathsf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

können wir die Diagonaleinträge vertauschen:

$$\mathsf{P}^t \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathsf{P} = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix}.$$

Damit führen wir schliesslich B in eine der sechs Normalformen über.

Zu zeigen ist nur noch, dass die sechs Normalformen zueinander nicht äquivalent sind. Die erste Matrix aus der Liste ist die Nullmatrix und sie ist zu keiner anderen Matrix äquivalent, denn $\mathsf{S}^t\mathbf{0}\mathsf{S} = \mathbf{0}\ \forall\mathsf{S}$. Dank Bemerkung 5.1 wissen wir, dass das Vorzeichen der Determinante sich nicht verändern kann. Das zeigt schon, dass die fünfte Matrix aus der Liste (Determinante = -1) zu keiner anderen äquivalent ist. Die Zweite und die Dritte Matrix haben Determinante gleich Null, während die Vierte und die Sechste Determinante gleich $+1$ haben. Deshalb gehören diese Gruppen zu verschiedenen Äquivalenzklassen. Sei jetzt

$$\mathsf{B} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

eine von diesen vier übrigen Matrizen. Durch das allgemeinste Transformation

$$\mathsf{S} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \qquad \alpha\delta - \beta\gamma \neq 0$$

erhalten wir

$$\mathsf{S}^t\mathsf{BS} = \begin{pmatrix} \alpha^2\lambda_1 + \gamma^2\lambda_2 & \alpha\beta\lambda_1 + \gamma\delta\lambda_2 \\ \alpha\beta\lambda_1 + \gamma\delta\lambda_2 & \beta^2\lambda_1 + \delta^2\lambda_2 \end{pmatrix}.$$

Gehört B zur Klasse von

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix},$$

so ist $\lambda_2 = 0$. Ist $\mathsf{S}^t\mathsf{BS}$ wieder eine von diesen zwei Matrizen, so ist ihr erster Eintrag $\alpha^2\lambda_1$, der das gleiche Vorzeichen wie λ_1 besitzt.

Gehört schliesslich B zur Klasse von

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

so ist $\lambda_2 = \lambda_1$, und deshalb ist der erste Eintrag von $\mathsf{S}^t\mathsf{BS}$ gleich $(\alpha^2 + \beta^2)\lambda_1$, der das gleiche Vorzeichen wie λ_1 besitzt. □

Bemerkung 5.11. Geometrisch entsprechen die Trasformationen R und P im obigen Beweis zu Skalierungen der Achsen und zur Spiegelung an der Diagonale $x = y$.

Anhang A. Gruppen

Eine Verknüpfung auf einer Menge G ist eine Abbildung

$$m\colon G \times G \to G.$$

Normalerweise schreibt man gh statt $m(g, h)$ für das Ergebnis der Verknüpfung.

Definition A.1. Eine Menge G mit einer Verknüpfung $(g, h) \mapsto gh$ und einem ausgewählten Element $e \in G$ heisst Gruppe, falls folgende Bedingungen erfüllt sind:

- **G1:** Die Verknüpfung ist assoziativ: $(gh)k = g(hk)\ \forall g, h, k \in G$.
- **G2:** Die Verknüpfung mit e ist die Identität: $ge = g\ \forall g \in G$.
- **G3:** Jedes Element besitzt ein Inverses: $\forall g \in G\ \exists h \in G : gh = e$.

Bemerkung A.2. Die Verknüpfung auf einer Gruppe heisst Multiplikation, das ausgewählte Element e heisst neutrales Element. Wegen **G1** schreibt man normalerweise einfach ghk für die Multiplikation dreier Elemente.

Am Ende von 3.4 haben wir das Beispiel der Gruppe GL_2 gesehen. Ein einfacheres Beispiel ist $G = \mathbb{R} \setminus \{0\}$ mit der üblichen Multiplikation und $e = 1$. Ein weiteres Beispiel ist $G = \mathbb{R}$, wobei die Verknüpfung die übliche Addition ist und $e = 0$ gilt (in solch einem Beispiel schreibt man $+$ für die Verknüpfung: $(g, h) \mapsto g + h$). Im ersten Beispiel ist die Multiplikation nichtkommutativ (d.h. i.A. $gh \neq hg$), in den anderen ist sie kommutativ (d.h. $gh = hg \; \forall g, h \in G$).

Satz A.3. *Sei G eine Gruppe. Dann:*

(1) *Ist h ein Inverses zu g (d.h. $gh = e$), dann gilt auch $hg = e$.*
(2) *Das neutrale Element operiert auch von links trivial: $eg = g \; \forall g \in G$.*
(3) *Das neutrale Element ist eindeutig bestimmt: $g\tilde{e} = g \; \forall g \in G \Rightarrow \tilde{e} = e$.*
(4) *Jedes Element besitzt* genau *ein Inverses: $\forall g \in G \; \exists! h \in G : gh = e$*

Beweis. Zu (1): Sei k ein Inverses zu h (d.h. $hk = e$). Dann

$$e = hk = hek = h(gh)k = (hg)(hk) = (hg)e = hg.$$

Zu (2): Sei $g \in G$ gegeben und sei h ein Inverses zu g (d.h. $gh = e$). Deshalb

$$eg = (gh)g = g(hg) = ge = e,$$

wobei wir (1) in der dritten Gleichung verwendet haben.

Zu (3): Sei $\tilde{e} \in G$ mit $g\tilde{e} = g \; \forall g \in G$. Insbesondere für $g = e$ haben wir $e\tilde{e} = e$, was bedeutet, dass $\tilde{e}$ ein Inverses zu e ist. Aus (1) folgt es aber, dass auch $\tilde{e}e = e$ gilt. Aber aus **G2** ist $\tilde{e}e = \tilde{e}$. Deshalb gilt $\tilde{e} = e$.

Zu (4): Seien h und $\tilde{h}$ Inverse zu g (d.h. $gh = e = g\tilde{h}$). Aus (1) folgt $\tilde{h}g = e$. Deshalb

$$\tilde{h} = \tilde{h}e = \tilde{h}(gh) = (\tilde{h}g)h = eh = h.$$

□

Bemerkung A.4. Wegen (3) spricht man dann von *dem* neutralen Element e und von *dem* Inversen h zu einem Element g. *Das* Inverse zu g wird üblicherweise mit g^{-1} bezeichnet. Zur Erinnerung:

$$ge = eg = g, \qquad gg^{-1} = g^{-1}g = e.$$

Das Inverse zu e ist e selbst. *Das* Inverse zu einem Produkt lässt sich einfach berechnen:

$$(g_1 g_2 \cdots g_n)^{-1} = g_n^{-1} \cdots g_2^{-1} g_1^{-1}.$$

Bemerkung A.5. Was wir hier gesehen haben, ist ein typisches Beispiel der Vorstellung einer mathematischen Struktur: Man führt die Struktur (in diesem Beispiel die Gruppe) durch eine Liste von minimalen Axiomen (hier **G1**, **G2** und **G3**) ein und man folgert daraus (hier im Satz A.3) die weiteren Eigenschaften (hier (1), (2), (3) und (4)). Es wäre nicht nur weniger elegant, alle Eigenschaften am Anfang (zusammen mit den Axiomen) aufzulisten, sondern auch weniger nützlich: Um eine Struktur zu anerkennen, genügt es die Axiomen zu verifizieren, und es ist ein Vorteil, wenn sie minimal sind.

Anhang B. Komplexe Zahlen

Die imaginäre Einheit i führ man ein, um formal die sonst unmögliche Gleichung $x^2 = -1$ aufzulösen. (Ist man mit diesem formalen Verfahren nicht zufrieden, dann gehe man direkt zu Abschnitt B.1.) Nämlich ist i ein Symbol mit der Eigenschaft

$$\mathrm{i}^2 = -1. \tag{B.1}$$

Eine komplexe Zahl z ist definitionsgemäss eine Linearkombination

$$z = a + \mathrm{i}b,$$

wobei a und b reelle Zahlen sind. Die reelle Zahlen selber sieht man als komplexe Zahlen mit i-Koeffizienten $b = 0$. Führt man die Konjugation $^-$ ein,

$$\overline{a + \mathrm{i}b} := a - \mathrm{i}b,$$

so kann man elegant die reellen Zahlen als die komplexen Zahlen z charakterisieren, die

$$\bar{z} = z$$

erfüllen.

Addition und Multiplikation von komplexen Zahlen genügen den üblichen Regeln von Ausdrücken mit reellen Koeffizienten zusammen mit der Rechenregel (B.1):

$$\begin{aligned}(a + \mathrm{i}b) + (c + \mathrm{i}d) &= (a + c) + \mathrm{i}(b + d),\\ (a + \mathrm{i}b)(c + \mathrm{i}d) &= (ac - bd) + \mathrm{i}(bc + ad).\end{aligned}$$

Addition und Multiplikation komplexer Zahlen haben dann die üblichen Eigenschaften (Assoziativität, Kommutativität, Distributivität); sie sind aber auch mit der Konjugation verträglich:

$$\begin{aligned}\overline{z + w} &= \bar{z} + \bar{w},\\ \overline{zw} &= \bar{z}\bar{w}.\end{aligned}$$

Daraus folgt, dass jede von Null verschiedene komplexe Zahl z invertierbar ist, denn man hat die explizite Formel

$$z^{-1} = \frac{\bar{z}}{|z|^2},$$

wobei

$$|z| := \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}, \qquad z = a + \mathrm{i}b,$$

die Norm von z ist (man sieht einfach, dass $|z| = 0$ genau dann gilt, wenn z Null ist).

Bemerkung B.1. Man kann auch Matrizen mit komplexen Einträgen, oder kürzer gesagt komplexe Matrizen, betrachten. Ihre Multiplikation, Skalarmultiplikation und Addition definiert man durch (3.4) und (3.8), wobei man jetzt Addition und Multiplikation von komplexen Zahlen benützt. Die Konjugation erweitert man auch zu Matrizen:

$$\overline{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} := \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix}.$$

Eine quadratische Gleichung

$$ax^2 + bx + c = 0, \qquad a \neq 0,$$

kann man im komplexen Sinne *immer* auflösen. Explizit hat man die übliche Formel

$$x_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

wobei jetzt alle Operation im komplexen Sinne zu verstehen sind. Ist insbesondere die Diskriminante

$$\Delta := b^2 - 4ac$$

reell, was z.B. passiert, wenn alle Koeffizienten a, b, c reell sind, so haben wir einfach

$$x_\pm = \frac{-b \pm \sqrt{\Delta}}{2a}, \qquad \Delta \geq 0,$$
$$x_\pm = \frac{-b \pm \mathrm{i}\sqrt{-\Delta}}{2a}, \qquad \Delta < 0,$$

wobei wir die Quadratwurzel einer positiven reellen Zahl (Δ bzw. $-\Delta$) berechnen sollen.

Im allgemeinen Fall müssen wir lernen, wie man die Quadratwurzel einer komplexen Zahl berechnet. Das tut man am einfachsten durch die Polarkoordinatendarstellung. Wir schreiben nämlich

$$z = a + \mathrm{i}b$$

als

$$z = \rho(\cos\theta + \mathrm{i}\sin\theta)$$

mit

$$\rho = |z| = \sqrt{a^2 + b^2}, \qquad a = \rho\cos\theta, \qquad b = \rho\sin\theta.$$

Dank der üblichen trigonometrischen Identitäten erhält man

$$\rho(\cos\theta + \mathrm{i}\sin\theta)\ \sigma(\cos\eta + \mathrm{i}\sin\eta) = \rho\sigma(\cos(\theta+\eta) + \mathrm{i}\sin(\theta+\eta)).$$

Daraus folgt

$$\sqrt{z} = \pm\sqrt{\rho}\left(\cos\frac{\theta}{2} + \mathrm{i}\sin\frac{\theta}{2}\right).$$

B.1. **Konkrete Darstellungen.** Komplexe Zahlen kann man konkret als Punkte der Ebene betrachten. Das ist die sog. GAUSS'sche Zahlenebene. Punkte dieser Ebene werden eigentlich als Vektoren betrachtet mit der üblichen Addition. Die Konjugation definiert man als Spiegelung an der x-Achse. Schliesslich definiert man die Multiplikation explizit durch die Formel

$$\begin{pmatrix} a \\ b \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} := \begin{pmatrix} ac - bd \\ bc + ad \end{pmatrix}.$$

Jeder Vektor in der Ebene kann man als Linearkombination der Vektoren $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ schreiben:

$$\begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Den formalen Gesichtspunkt bekommt man wieder, indem man den Vektor $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ wegfallen lässt und $\mathrm{i} := \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ setzt.

In der Polarkoordinatendarstellung hat die Multiplikation eine geometrischere Bedeutung: Das Produkt zweier komplexen Zahlen, die man als Vektoren betrachtet, ist der Vektor, dessen Länge das Produkt der Längen ist und dessen (im Gegenuhrzeigersinne orientierten) Winkel mit der x-Achse die Summe der Winkel ist.

Die komplexen Zahlen kann man auch als 2×2-reelle Matrizen der Form

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

mit den üblichen Matrixaddition und Matrixmultiplikation definieren. Die Konjugation definiert man als die Transposition. Zu bemerken ist, dass jede solche Matrix eine Linearkombination der Matrizen

$$\mathbf{1} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{und} \qquad \mathsf{J} := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

ist:

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = a\mathbf{1} + b\mathsf{J}.$$

Man sieht einfach, dass

$$\mathsf{J}^2 = -\mathbf{1}$$

gilt. Den formalen Gesichtspunkt bekommt man wieder, indem man die Matrix $\mathbf{1}$ wegfallen lässt und $\mathrm{i} := \mathsf{J}$ setzt.

Anhang C. Beweismethoden

Wir werden hier sehr informell die wichtigsten Beweismethoden erklären. Eine ausführlichere Einführung gibt es z.B. in [1, I.1].

Ein Satz in der Mathematik ist eine logisch bewiesene Aussage. Genauer gesagt besteht ein Satz normalerweise aus einer Implikation

$$A \Rightarrow B,$$

wobei A (die Voraussetzung) und B (die Behauptung) Aussagen sind, und aus einem Beweis, in dem man A zusammen mit Axiomen und bereits bewiesenen Aussagen verwendet, um logisch die Gültigkeit von B zu zeigen.

Eine der angenommenen Regeln ist, dass eine Aussage entweder wahr oder falsch ist. Das ist der sog. Satz vom ausgeschlossenen Dritten (lat. tertium non datur), der von den meisten Logiksystemen als Grundprinzip angenommen wird. Das bedeutet, dass nicht jeder Ausdruck eine Aussage ist. Z.B. sind „2+2 = 4“ und „Jede ungerade Zahl ist eine Primzahl“ Aussagen (die erste ist wahr und die zweite ist falsch).

Wir führen ein paar wichtigen Notationen aus der Logik ein: die Konjunktion $\wedge$, die Disjunktion $\vee$ und die Negation $\neg$. Die Aussage $A \wedge B$ ist definitionsgemäss wahr, wenn die Aussage A *und* die Aussage B gelten, sonst ist sie falsch. Die Aussage $A \vee B$ ist definitionsgemäss wahr, wenn die Aussage A *oder* die Aussage B gilt, sonst ist sie falsch. Die Aussage $\neg A$ ist definitionsgemäss wahr, wenn die Aussage A falsch ist und umgekehrt. Mit diesen Notationen besagt der Satz vom ausgeschlossenen Dritten, dass $A \wedge \neg A$ für jede Aussage A falsch ist. Das Gesetz der doppelten Negation ist die Regel, dass $\neg\neg A$ genau dann gilt, wenn A gilt.

Man muss Acht geben, um die Negation einer Aussage richtig zu formulieren. Z.B. ist die Negation von „$2 + 2 = 4$“ einfach „$2 + 2 \neq 4$“, aber die Negation von „Jede ungerade Zahl ist eine Primzahl“ ist die kompliziertere Aussage „Es gibt eine ungerade Zahl, die keine Primzahl ist“. Wir beziehen uns z.B. auf [1, I.1] für die genauen Rechenregeln.

Wollen wir zeigen, dass A falsch ist, können wir deshalb zeigen, dass $\neg A$ wahr ist. Das entspricht oft dem Finden eines Gegenbeispiels. Man kann z.B. beweisen, dass die Aussage „Jede ungerade Zahl ist eine Primzahl“ falsch ist, indem man bemerkt, dass die ungerade Zahl 9 keine Primzahl ist.

Die logische Implikation definiert man dann einfach als

$$A \Rightarrow B := (\neg A) \vee B.$$

Das bedeutet, dass die Implikation definitionsgemäss genau dann gilt, wenn A falsch ist oder wenn A und B wahr sind. Die Umkehrung der Implikation $A \Rightarrow B$ ist die Implikation $B \Rightarrow A$. Wenn eine Implikation und ihre Umkehrung gelten, spricht man von einer Äquivalenz:

$$A \Leftrightarrow B := (A \Rightarrow B) \wedge (B \Rightarrow A).$$

Wir kommen jetzt zu drei wichtigen Regeln über das Implikationsverfahren:

Satz C.1.

(1) *Gelten die Implikationen $A \Rightarrow C$ und $C \Rightarrow B$, so gilt auch die Implikation $A \Rightarrow B$.*
(2) *Jede Implikation $A \Rightarrow B$ ist zu ihrer* Kontraposition *$\neg B \Rightarrow \neg A$ äquivalent.*
(3) *Die Aussage $(A \wedge \neg B) \Rightarrow C$ ist zur Aussage $(A \Rightarrow B) \vee C$ äquivalent.*

Beweis. Zu (1): Ist A falsch, so gilt definitionsgemäss die Aussage $A \Rightarrow B$. Ist andererseits A wahr, so sagt $A \Rightarrow C$, dass C wahr ist; aber $C \Rightarrow B$ sagt dann, dass B auch wahr ist.

Zu (2): Definitionsgemäss ist $\neg B \Rightarrow \neg A$ zu $(\neg\neg B) \vee (\neg A)$ äquivalent. Wegen der Kommutativität der Konjunktion ist das zu $(\neg A) \vee (\neg\neg B)$ äquivalent. Dank dem Gesetz der doppelten Negation ist dies schliesslich zu $(\neg A) \vee B$ äquivalent, was die Definition von $A \Rightarrow B$ ist.

Zu (3): Definitionsgemäss ist $(A \wedge \neg B) \Rightarrow C$ zu $\neg(A \wedge \neg B) \vee C$ äquivalent, d.h. zu $(\neg A \vee B) \vee C$. □

Diese Regeln liegen den drei Beweisverfahren zugrunde, um $A \Rightarrow B$ zu zeigen:

Der direkte Beweis: Man zeigt die endlich vielen Implikationen $A \Rightarrow C_1$, $C_1 \Rightarrow C_2$, $C_2 \Rightarrow C_3, \ldots, C_{n-1} \Rightarrow C_n$, $C_n \Rightarrow B$. Aus iterierter Anwendung von (1) folgt $A \Rightarrow B$.

Der indirekte Beweis: Man zeigt (durch einen direkten Beweis) die Kontraposition $\neg B \Rightarrow \neg A$, die dank (2) zu $A \Rightarrow B$ äquivalent ist.

Der Widerspruchsbeweis: Man zeigt (durch einen direkten oder indirekten Beweis) die Implikation $(A \wedge \neg B) \Rightarrow C$, wobei C eine falsche Aussage ist (oft eine Aussage der Form $D \wedge \neg D$). Dank (3) ist dann $A \Rightarrow B$ wahr.

Der Leser ist eingeladen zu versuchen, die verschiedenen Verfahren in den Beweisen dieser Notizen zu erkennen.

Bemerkung C.2. Die Beweisverfahren können verschachtelt sein: Nämlich wird jede der Implikationen in einem direkten Beweis durch einen direkten, einen indirekten oder durch einen Widerspruchsbeweis gezeigt. Wenn eine von diesen Implikationen einen langen Beweis braucht, ist es zur Lesbarkeit einer mathematischen Arbeit besser, sie als unabhängigen Satz (oft Hilfssatz genannt) vorzustellen. Normalerweise wird der Hilfssatz *vor* dem Satz, in dem er gebraucht wird, im Text vorgestellt.

Bemerkung C.3. Zum Beweis einer Äquivalenz von Aussagen $A \Leftrightarrow B$ verwendet man oft einen direkten Beweis für die eine Richtung und einen indirekten Beweis für die andere. Z.B. zeigt man $\neg A \Rightarrow \neg B$ statt $B \Rightarrow A$. D.h., um $A \Leftrightarrow B$ zu zeigen, zeigt man

$$(A \Rightarrow B) \wedge (\neg A \Rightarrow \neg B).$$

Bemerkung C.4. Um die Lesbarkeit einer mathematischen Arbeit zu verbessern, verwendet man verschiedene Bezeichnungen: ein Hilfssatz (oder Lemma) ist ein Satz, der zum Beweis von einem oder mehreren Sätzen verwendet werden wird, aber sonst von geringer Bedeutung ist.[11] Ein Folgesatz (oder Korollar) ist ein Satz, der fast unmittelbar aus einem gerade bewiesenen Satz folgt. Ein Lehrsatz (oder Theorem) ist ein besonders wichtiger Satz. Schliesslich nennt man eine Aussage, für die man keinen Beweis, sondern nur eine Begründung hat, Vermutung. Eine Vermutung ist deshalb kein Satz, solange kein Beweis gefunden wird, und oft werden Vermutungen widerlegt.

Literatur

[1] H. Amann und J. Escher, *Analysis I*, Birkhäuser.
[2] G. Fischer, *Lineare Algebra*, Vieweg Studium.

[11] Das ist eine Entscheidung des Verfassers. Es gibt viele Beispiele von Hilfssätzen, die nachträglich als sehr wichtig anerkannt wurden und heute berühmter sind als viele andere Sätze.

Index