Mathematik

für Physiker

Dr. Michael Kaplan <kaplan@ma.tum.de>

WS 2010/11

Teil I

Lineare Algebra

Inhaltsverzeichnis

1 Einleitung

1.1 Worum geht's?

Am Beginn des Studiums der Mathematik stehen traditionell zwei Vorlesungen, die *Analysis* und die *Lineare Algebra*

Analysis
Der zentrale Begriff der Analysis ist der *Grenzwert*. Er erlaubt den Übergang von der Sekante zur *Tangente*, vom Differenzenquotienten zur Ableitung, von der Summe zum *Integral*. Darauf baut die Differential- und Integralrechnung auf. Mit Hilfe von Differential- oder Integralgleichungen lassen sich Naturvorgänge hervorragend modellieren.

Lineare Algebra
Die lineare Algebra ist, wie der Name bereits sagt, ein Teilgebiet der *Algebra*[1].
In der Algebra geht es um die *grundlegenden Strukturen*, wie *Gruppen*, *Ringe* und *Körper*, die Grundrechenarten in diesen Strukturen, und die Auflösung der in diesem Zusammenhang entstehenden *Gleichungen*. In der linearen Algebra betrachtet man vorerst nur den einfachsten Fall linearer Gleichungen, in denen die zu berechnenden Unbestimmten nur linear vorkommen.

- Das Studium linearer Gleichungen zeigt, dass deren Lösungsmengen selbst auch wieder mathematische Struktur besitzen, was auf die Begriffe *Vektorraum* führt.
- Die Untersuchung von Vektorräumen und Abbildungen zwischen diesen zeigt, dass diese neuen Begriffe weit über das Lösen von Gleichungssystemen hinaus eine zentrale Bedeutung in vielen mathematischen Disziplinen haben.
- Bei vielen Naturgesetzen ist die Linearität grundlegende Eigenschaft (siehe einleitende Beispiele, erstes Übungsblatt etc.)
- Von Haus aus nichtlineare Probleme werden oft im ersten Schritt linearisiert und können so mit Linearer Algebra behandelt werden
- Viele numerische Verfahren beruhen ebenfalls auf linearer Approximation, lineare Optimierung ist weitgehend eine Anwendung der linearen Algebra
- Somit ist die Lineare Algebra für die Physik ein selbstverständliches Werkzeug zur Mathematisierung der Theorie
- Während in der Schule meist Geometrie sehr aus der Anschauung heraus betrieben wird, wird in der Wissenschaft *deduktiv* vorgegangen, d.h. die untersuchten Strukturen werden *axiomatisch* beschrieben und Folgerungen daraus werden untersucht.
- Ein *Axiom* (von griech. axiómata=Grundsatz oder Forderung) nennt man eine Aussage, die grundlegend ist und deshalb nicht innerhalb ihres Systems begründet werden kann bzw. muss. Sie dient als Grundlage für eine deduktive Theorie und kann deshalb nicht selber durch diese Theorie begründet werden. Wenn eine Theorie aus begründeten Sätzen bestehen soll, so muss es notwendigerweise solche Axiome geben, denn sonst würde die Argumentation nie enden: Jeder Satz, den ich zur Begründung anführte, bedürfte wieder einer Begründung usw. Daher ist ein Axiom etwas ganz anderes als eine Vermutung.

Beispiel 1.1.1. *„Jede natürliche Zahl n hat einen Nachfolger $n+1$“ ist ein Axiom der Arithmetik.*

- Ein *Axiomensystem* für eine mathematische Theorie muss *widerspruchsfrei* sein, d.h. aus ihm darf nicht gleichseitig ein Satz und sein logisches Gegenteil ableitbar sein.
- Weiterhin muss ein Axiomensystem *vollständig* sein, d.h. alle Sätze der Theorie müssen aus ihm logisch abzuleiten sein.
- Aus ökonomischen Gründen fordert man außerdem meist die *Unabhängigkeit* von Axiomensystem, d.h. keines der Axiome soll aus den anderen herleitbar sein.

[1] Das Wort *Algebra* kommt aus dem Arabischen. Es geht zurück auf ein Werk des Muhammad ibn Musa al-Chwarismi (aus der Region südöstlich des Kaspischen Meeres, heute Usbekistan), 780 bis ca. 850 n.Chr, der zur Zeit des Kalifen al-Ma'mun (813-833) in Bagdad arbeitete. Er verfasste ein Lehrbuch mit Titel „al-jabr wa-lmuqabala“ (frei übersetzt: „Regeln der Wiedereinsetzung und Reduktion“). Hier geht es um die Umformung und Lösung von linearen und quadratischen Gleichungen. Das Wort al-jabr (hier als „Wiedereinsetzung“ übersetzt) hat sich bis heute als „Algebra“ erhalten. Aus dem Namen des Autors al-Chwarismi wurde dann übrigens unser *Algorithmus* als Begriff für „exaktes Rechenverfahren“.

2 Lineare Gleichungssysteme

2.1 Grundlagen

Beispiel 2.1.1.

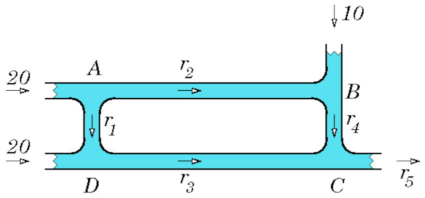


Abbildung 1: Verbundene Wasserrohre

Wasserrohre sind in den Punkten A, B, C und D wie skizziert verbunden. Bei A, B und D wird Wasser eingeleitet mit $20\frac{\ell}{min}$, $10\frac{\ell}{min}$ und $20\frac{\ell}{min}$. Der Fluss in Litern pro Minute in den anderen Rohren wird mit $r_1, \ldots, r_5$ benannt (Richtung beliebig festgelegt).

Da Wasser inkompressibel ist, führt dies auf 4 Gleichungen in den 5 Unbekannten $r_1, \ldots, r_5$:

$$\begin{aligned}
A &: r_1 + r_2 = 20 \\
B &: r_2 + 10 = r_4 \iff r_2 - r_4 = -10 \\
C &: r_3 + r_4 - r_5 = 0 \\
D &: r_1 + 20 = r_3 \iff r_1 - r_3 = -20
\end{aligned}$$

Im vorliegenden Fall wird man wohl recht schnell die Lösung $r_1 = 0$, $r_2 = 20$, $r_3 = 20$, $r_4 = 30$ und $r_1 = 50$ sehen, es stellen sich aber die Fragen,

- *wie man diese systematisch berechnen kann,*
- *ob es noch mehr (oder auch mal keine) Lösungen gibt*
- *und wie die Lösungen solch eines Systems allgemein aussehen.*

Definition 2.1.2. *Ein System von Gleichungen der Gestalt*

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\
 & & & & \vdots & & & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m
\end{array} \tag{2.1}$$

heißt lineares Gleichungssystem (LGS). Dabei sind $x_1, \ldots, x_n$ die zu berechnenden Unbestimmten, die Koeffizienten a_{ij} mit $1 \leq i \leq m$ und $1 \leq j \leq n$ und die auf der rechten Seite stehenden b_i mit $1 \leq i \leq m$ entstammen einem zugrunde gelegten Körper K. Mit Hilfe des Summenzeichens kann man auch kompakter

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \textit{ für } i = 1, \ldots, m$$

schreiben.

- Ein Körper ist dabei eine mathematische Struktur, die benötigt wird, um die entsprechenden Rechenoperationen zur Aufstellung und Lösung eines LGS auszuführen. Für den Anfang sei $K = \mathbb{R}$.
- Man ist an der Lösungsmenge des LGS interessiert, d.h. an
$$\mathbb{L} := \{(x_1, \ldots, x_n) \in K^n \text{ mit } (2.1)\}\,.$$
- Diese Lösungsmenge kann natürlich auch leer sein.
- Nachdem in vielen Anwendungen solche lineare Gleichungssysteme durchaus auch mal mit hunderten (oder viel mehr) von Variablen und Gleichungen auftreten und man dann solche Arbeiten gerne dem Computer überlässt, muss man also einen effektiven *Algorithmus* formulieren, um damit noch zurecht zu kommen.
- Zur Lösung des LGS nimmt man *Äquivalenzumformungen* vor, d.h. Umformungen, die die Lösungsmenge des Systems nicht verändern.
- Diese erkennt man daran, dass sie sich ohne Informationsverlust wieder vollständig rückgängig machen lassen

- Das sind genau die folgenden Umformungen:
  (i) Vertauschen zweier Gleichungen,
  (ii) Beidseitige Multiplikation einer Gleichung mit einer Zahl $\neq 0$,
  (iii) Addition eines beliebigen Vielfachen einer Gleichung zu einer anderen Gleichung.

2.2 Lösungsstrategie

- Wenn in (2.1) die Unbestimmte x_1 überhaupt vorkommt, so ist einer der Koeffizienten $a_{11}, \ldots, a_{m1} \neq 0$.
- Ist $a_{11} = 0$ und $a_{\ell 1} \neq 0$, so tauscht man im ersten Schritt die 1-te und die ℓ-te Gleichung (Umformung vom Typ (i)).
- Dies führt auf

$$\begin{array}{lccccccccc}
\text{1-te Zeile} \rightarrow & a_{\ell 1}x_1 & + & a_{\ell 2}x_2 & + & \ldots & + & a_{\ell n}x_n & = & b_\ell \\
\updownarrow & & & & & \vdots & & & & \vdots \\
\ell\text{-te Zeile} \rightarrow & & & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\
 & & & & & \vdots & & & & \vdots \\
 & a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m
\end{array}$$

- Um kein Durcheinander mit den Indizes zu bekommen, geht man weiterhin von der Gestalt (2.1) der Gleichung aus und nimmt an, dass gleich so nummeriert wurde, dass $a_{11} \neq 0$ ist.
- Wegen $a_{11} \neq 0$ darf man die erste Gleichung durch a_{11} dividieren (Umformung (ii)). Alle anderen Gleichungen bleiben gleich:

$$\begin{array}{ccccccccc}
x_1 & + & \frac{a_{12}}{a_{11}}x_2 & + & \ldots & + & \frac{a_{1n}}{a_{11}}x_n & = & \frac{b_1}{a_{11}} \\
a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\
 & & & & \vdots & & & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m
\end{array}$$

- Nun verwendet man mehrere Umformungen vom Typ (iii), um mit Hilfe der ersten Gleichung die Variable x_1 aus den anderen Gleichungen zu eliminieren. Dazu geht man so vor:
  - Subtrahiere das a_{21}-fache der 1.Zeile von der 2.Zeile
  - Subtrahiere das a_{31}-fache der 1.Zeile von der 3.Zeile
  - ...
  - Subtrahiere das a_{m1}-fache der 1.Zeile von der m.Zeile
- Dies führt auf die folgende Gestalt des LGS:

$$\begin{array}{ccccccccc}
x_1 & + & a'_{12}x_2 & + & \ldots & + & a'_{1n}x_n & = & b'_1 \\
 & & a'_{22}x_2 & + & \ldots & + & a'_{2n}x_n & = & b'_2 \\
 & & & & \vdots & & & & \vdots \\
 & & a'_{m2}x_2 & + & \ldots & + & a'_{mn}x_n & = & b'_m
\end{array} \tag{2.2}$$

mit $a'_{12} = \frac{a_{12}}{a_{11}}, \ldots, a'_{1n} = \frac{a_{1n}}{a_{11}}, b'_1 = \frac{b_1}{a_{11}}$ und $a'_{22} = a_{22} - a_{21} \cdot \frac{a_{12}}{a_{11}}, a'_{23} = a_{23} - a_{21} \cdot \frac{a_{23}}{a_{11}}, \ldots, b'_2 = b_2 - a_{21} \cdot \frac{b_1}{a_{11}}$ usw.

Beispiel 2.2.1 (Fortsetzung von 2.1.1). *Das LGS im Beispiel mit den Wasserrohren sieht etwas schöner geschrieben und in abgeänderter Reihenfolge so aus:*

$$\begin{array}{ccccccccccc}
r_1 & + & r_2 & & & & & & & = & 20 \\
r_1 & & & - & r_3 & & & & & = & -20 \\
 & & r_2 & & & - & r_4 & & & = & -10 \\
 & & & & r_3 & + & r_4 & - & r_5 & = & 0
\end{array}$$

Da links oben schon r_1 steht, ist in den Schritten 1 und 2 nichts zu tun. Für Schritt 2 subtrahiert man die 1.Gleichung von der 2.Gleichung:

$$\begin{array}{rcrcrcrcrcr} r_1 & + & r_2 & & & & & & & = & 20 \\ & - & r_2 & - & r_3 & & & & & = & -40 \\ & & r_2 & & & - & r_4 & & & = & -10 \\ & & & & r_3 & + & r_4 & - & r_5 & = & 0 \end{array}$$

- Ab jetzt lässt man die 1.Gleichung des LGS (und damit die 1.Variable r_1) 2.2 unverändert und rechnet sinngemäß mit dem Teilsystem aus Gleichungen 2 bis m genauso weiter.
- Aus
$$\begin{array}{ccccccccc} a'_{22}x_2 & + & a'_{23}x_3 & + & \dots & + & a'_{2n}x_n & = & b'_2 \\ & & & & \vdots & & & & \vdots \\ a'_{m2}x_2 & + & a'_{m3}x_3 & + & \dots & + & a'_{mn}x_n & = & b'_m \end{array}$$
- wird (falls x_2 überhaupt vorkommt, sonst analog mit x_3)
$$\begin{array}{ccccccccc} x_2 & + & a''_{23}x_3 & + & \dots & + & a''_{2n}x_n & = & b''_2 \\ & & & & \vdots & & & & \vdots \\ & & a''_{m3}x_3 & + & \dots & + & a''_{mn}x_n & = & b''_m \end{array}$$
- Damit hat man bisher insgesamt
$$\begin{array}{ccccccccccc} x_1 & + & a'_{12}x_2 & + & a'_{13}x_3 & + & \dots & + & a'_{1n}x_n & = & b'_1 \\ & & x_2 & + & a''_{23}x_3 & + & \dots & + & a''_{2n}x_n & = & b''_2 \\ & & & & & & \vdots & & & & \vdots \\ & & & & a''_{m3}x_3 & + & \dots & + & a''_{mn}x_n & = & b''_m \end{array}$$
- Setzt man dieses Verfahren entsprechend fort, so kommt man auf eine *Stufenform* (wird noch genau definiert) aus der man dann sehr schön das Ergebnis ablesen kann.

Beispiel 2.2.2 (Fortsetzung von 2.2.1)**.**

$$\begin{array}{rcrcrcrcrcr} r_1 & + & r_2 & & & & & & & = & 20 \\ & - & r_2 & - & r_3 & & & & & = & -40 \\ & & r_2 & & & - & r_4 & & & = & -10 \\ & & & & r_3 & + & r_4 & - & r_5 & = & 0 \end{array}$$

Nachdem die 1. Variable r_1 nur noch in der 1.Gleichung vorkommt, wendet man sich jetzt der nächsten Variablen r_2 zu. Diese kommt in der 2.Zeile auch vor (muss also nicht hochgetauscht werden). Da der Koeffizient von r_2 in der 2.Gleichung -1 ist, multipliziert man diese mit -1. Wahlweise könnte man auch die 2. und die 3. Gleichung vertauschen:

$$\begin{array}{rcrcrcrcrcr} r_1 & + & r_2 & & & & & & & = & 20 \\ & & r_2 & + & r_3 & & & & & = & 40 \\ & & r_2 & & & - & r_4 & & & = & -10 \\ & & & & r_3 & + & r_4 & - & r_5 & = & 0 \end{array}$$

Nun subtrahiert man die 2.Gleichung von der 3.Gleichung und erhält:

$$\begin{array}{rcrcrcrcrcr} r_1 & + & r_2 & & & & & & & = & 20 \\ & & r_2 & + & r_3 & & & & & = & 40 \\ & & & - & r_3 & - & r_4 & & & = & -50 \\ & & & & r_3 & + & r_4 & - & r_5 & = & 0 \end{array}$$

Entsprechend geht es nun mit r_3 in der 3.Gleichung weiter, d.h. man multipliziert mit -1 und subtrahiert dann die 3.Gleichung von der 4.Gleichung:

$$\begin{array}{rcrcrcrcrcr} r_1 & + & r_2 & & & & & & & = & 20 \\ & & r_2 & + & r_3 & & & & & = & 40 \\ & & & & r_3 & + & r_4 & & & = & 50 \\ & & & & & & & - & r_5 & = & -50 \end{array}$$

Damit sind die beschriebenen Eliminationsschritte erledigt, die angekündigte Stufenform ist erreicht. An der letzten Zeile kann man immerhin schon mal ablesen, dass r_5 nur die eindeutige Lösung 50 besitzt. Die Interpretation der anderen Zeilen gestaltet sich dagegen etwas schwieriger.

2.3 Matrixschreibweise

Definition 2.3.1. *Ein rechteckiges Zahlenschema aus m mal n Elementen eines Körpers K heißt $m \times n$-Matrix. Man schreibt*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in K^{m \times n},$$

oder kurz $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ (oder nur (a_{ij}), wenn m und n klar sind).
$K^{m \times n}$ ist die Menge aller $m \times n$-Matrizen mit Einträgen aus K (oft auch $M_{m,n}(K)$).

Definition 2.3.2. *Das LGS (2.1) schreibt man abkürzend als $Ax = b$. Dabei heißt $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in K^{m \times n}$ die Koeffizientenmatrix, $x = (x_j)_{1 \le j \le n} \in K^n$ der Vektor aus den Unbekannten und $b = (b_i)_{1 \le i \le m} \in K^m$ die rechte Seite oder Inhomogenität. Das LGS heißt homogen, wenn $b_1 = b_2 = \dots = b_m = 0$ ist (kurz $b = 0$), sonst inhomogen. Besonders ökonomisch für das Lösen des LGS (2.1) ist die Schreibweise mit der erweiterten Koeffizientenmatrix $(A|b) \in K^{m \times (n+1)}$.*

Beispiel 2.3.3 (Fortsetzung von 2.2.2)**.** *Das Wasserrohr-Beispiel ist ein inhomogenes LGS $Ar = b$ mit*

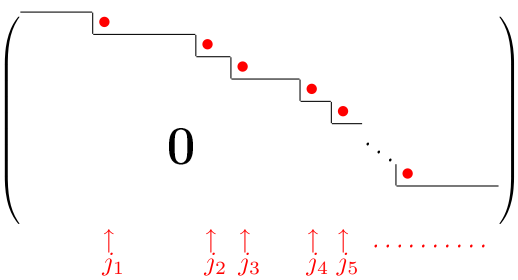
$$A = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & -1 \end{pmatrix} \in \mathbb{R}^{4 \times 5}, \; b = \begin{pmatrix} 20 \\ -20 \\ -10 \\ 0 \end{pmatrix} \in \mathbb{R}^4,$$

$$(A|b) = \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20 \\ 1 & 0 & -1 & 0 & 0 & -20 \\ 0 & 1 & 0 & -1 & 0 & -10 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right) \in \mathbb{R}^{4 \times 6}, \; r = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{pmatrix} \in \mathbb{R}^5$$

2.4 Elementare Zeilenumformungen und Zeilenstufenform

- Mit der erweiterten Koeffizientenmatrix kann man nun genauso rechnen, wie mit dem damit verbundenen LGS.
- Die drei bereits formulierten erlaubten Umformungen werden deshalb jetzt nochmal für (erweiterte) Koeffizientenmatrizen notiert.
- *Elementare Zeilenumformungen*
  Die Lösungsmenge eines LGS ändert sich nicht, wenn man an der erweiterten Koeffizientenmatrix die folgenden Umformungen in beliebiger Reihenfolge ausführt:
  (I) Vertauschen zweier Zeilen
  (II) Multiplikation einer Zeile mit einer Konstanten $\alpha \in K \setminus \{0\}$
  (III) Addition des α-fachen einer Zeile, $\alpha \in K$ beliebig, zu einer anderen.
- Mit diesen elementaren Zeilenumformungen kann man nun nach dem zuvor Gesagten jede Matrix (erweitert oder nicht) in Stufenform bringen.

Definition 2.4.1. *Eine $m \times n$-Matrix $A \in K^{m \times n}$ heißt in Zeilenstufenform (oder Staffelform), wenn sie von folgender Gestalt ist:*



Dabei sind die Elemente an den mit • markierten Stellen (diese heißen Pivotelemente, die zugehörigen Spalten Pivotspalten) ungleich Null und unterhalb der Treppenlinie stehen ausschließlich Nullen.

Definition 2.4.2. *Es sei K ein Körper. Eine Matrix $A \in K^{m\times n}$ heißt in Zeilenstufenform, wenn sie entweder gleich der Nullmatrix $O = (0)_{\substack{1\le i\le m\\ 1\le j\le n}}$ ist oder wenn gilt*

(1) Es gibt eine Zahl r (den Rang von A, i.Z. $\operatorname{Rang}(A) = r$) mit $1 \le r \le m$, so dass in den Zeilen mit den Indizes $1,\dots,r$ jeweils nicht nur Nullen und in den Zeilen mit den Indizes $r+1,\dots,m$ ausschließlich Nullen stehen.

(2) Für jeden Zeilenindex i mit $1 \le i \le r$ sei j_i der kleinste Index jener Spalte, in der ein Element ungleich Null steht, d.h. $j_i := \min\{j \mid a_{ij} \ne 0\}$ (also $1 \le j_i \le n$). Für diese Indizes gilt die Stufenbedingung $j_1 < j_2 < \ldots < j_r$.

- Es empfiehlt sich, die elementaren Zeilenumformungen auf dem Weg zur Zeilenstufenform jeweils mitzuschreiben.
- Notiert man etwa $Z_2 \leftarrow Z_2 - 4Z_1$, so heißt das: „Die neue zweite Zeile (Z_2) sei die alte zweite Zeile minus das 4-fache der ersten Zeile“, d.h. es handelt sich um eine elementare Zeilenumformung vom Typ (iii).
- Für den Übergang von einer Koeffizientenmatrix zur nächsten schreibt man meistens $\to$ oder $\rightsquigarrow$ (jedenfalls nicht $=$!).

Beispiel 2.4.3 (Fortsetzung von 2.3.3).

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20\\ 1 & 0 & -1 & 0 & 0 & -20\\ 0 & 1 & 0 & -1 & 0 & -10\\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right) \underset{Z_2\leftarrow Z_2-Z_1}{\rightsquigarrow} \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20\\ 0 & -1 & -1 & 0 & 0 & -40\\ 0 & 1 & 0 & -1 & 0 & -10\\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right) \underset{Z_3\leftarrow Z_3+Z_2}{\rightsquigarrow}$$

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20\\ 0 & -1 & -1 & 0 & 0 & -40\\ 0 & 0 & -1 & -1 & 0 & -50\\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right) \underset{Z_4\leftarrow Z_4+Z_3}{\rightsquigarrow} \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20\\ 0 & -1 & -1 & 0 & 0 & -40\\ 0 & 0 & -1 & -1 & 0 & -50\\ 0 & 0 & 0 & 0 & -1 & -50 \end{array}\right)$$

2.5 Von der Zeilenstufenform zur Lösung

Beispiel 2.5.1. *Es sei*

$$(A|c) = \left(\begin{array}{cccccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & 5 & 6 & 3\\ & & & 5 & 2 & 3 & 4 & 5 & 5\\ & \mathbf{0} & & & & & & 4 & 7\\ & & & & & & & & 9 \end{array}\right)$$

Aus dieser Zeilenstufenform liest man $m = 4$, $n = 8$ und $r = \operatorname{Rang}(A) = 3$ bzw. $\operatorname{Rang}(A|c) = 4$ ab.
Das Gleichungssystem $Ax = c$ ist nicht lösbar, denn seine vierte Gleichung lautet $0x_1 + \ldots + 0x_8 = 9$, d.h. sie ist durch keine Belegung der x_i erfüllbar, denn $0 \cdot x = 0$ für alle x in jedem Körper.
Dies sieht man sofort aus der Zeilenstufenform an $\underline{b_{r+1} = b_4 \ne 0}$ bzw. $\operatorname{Rang}(A|c) > \operatorname{Rang}(A)$. 21.10.10

Beispiel 2.5.2. *Es sei*

$$(A|b) = \left(\begin{array}{cccccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & 5 & 6 & 18\\ & & & 5 & 2 & 3 & 4 & 5 & 10\\ & \mathbf{0} & & & & & & 4 & 12\\ & & & & & & & & 0 \end{array}\right)$$

Es ist $m = 4$, $n = 8$, $r = \operatorname{Rang}(A) = 3 = \operatorname{Rang}(A|b)$, d.h. es ergibt sich nicht wie im letzten Beispiel ein direkter Widerspruch.
Das LGS ist lösbar, man kann jeweils nach den Pivotelementen auflösen.
Pivotspalten sind $j_1 = 1$, $j_2 = 4$ und $j_3 = 8$.
Die zugehörigen Variablen x_1, x_4, x_8 nennt man auch gebundene Variable, während die restlichen Variablen x_2, x_3, x_5, x_6 und x_7 so genannte freie Variable sind.
Aus der Zeilenstufenform ergibt sich (von unten nach oben)

$$\begin{array}{rcl} x_8 &=& 3\\ x_4 &=& 2 - \frac{2}{5}x_5 - \frac{3}{5}x_6 - \frac{4}{5}x_7 - x_8 =\\ &=& -1 - \frac{2}{5}x_5 - \frac{3}{5}x_6 - \frac{4}{5}x_7\\ x_1 &=& 18 - 5x_7 - 6x_8 = -5x_7 \end{array}$$

$$
\begin{array}{lrcrcrcrcr}
 & x_1 & = & & & & & & & -5x_7 \\
\textit{bzw.} & x_4 & = & (-1) & + & (-\tfrac{2}{5}x_5) & + & (-\tfrac{3}{5}x_6) & + & (-\tfrac{4}{5}x_7) \\
 & x_8 & = & 3 & & & & & &
\end{array}
$$

Ausführlich ist das

$$
\begin{array}{rclllllll}
x_1 & = & & & & & & & -5x_7 \\
x_2 & = & & & x_2 & & & & \\
x_3 & = & & & & x_3 & & & \\
x_4 & = & (-1) & + & & & (-\tfrac{2}{5}x_5) \; + & (-\tfrac{3}{5}x_6) \; + & (-\tfrac{4}{5}x_7) \\
x_5 & = & & & & & x_5 & & \\
x_6 & = & & & & & & x_6 & \\
x_7 & = & & & & & & & x_7 \\
x_8 & = & 3 & & & & & &
\end{array}
$$

Dies kann man auch spaltenweise zusammenfassen

$$
\begin{pmatrix} x_1\\x_2\\x_3\\x_4\\x_5\\x_6\\x_7\\x_8 \end{pmatrix} = \begin{pmatrix} 0\\0\\0\\-1\\0\\0\\0\\3 \end{pmatrix} + x_2 \begin{pmatrix} 0\\1\\0\\0\\0\\0\\0\\0 \end{pmatrix} + x_3 \begin{pmatrix} 0\\0\\1\\0\\0\\0\\0\\0 \end{pmatrix} + x_5 \begin{pmatrix} 0\\0\\0\\-\frac{2}{5}\\1\\0\\0\\0 \end{pmatrix} + x_6 \begin{pmatrix} 0\\0\\0\\-\frac{3}{5}\\0\\1\\0\\0 \end{pmatrix} + x_7 \begin{pmatrix} -5\\0\\0\\-\frac{4}{5}\\0\\0\\1\\0 \end{pmatrix}
$$

Üblicherweise benennt man die verbleibenden Variablen um

$$
\begin{pmatrix} x_1\\x_2\\x_3\\x_4\\x_5\\x_6\\x_7\\x_8 \end{pmatrix} = \begin{pmatrix} 0\\0\\0\\-1\\0\\0\\0\\3 \end{pmatrix} + \lambda_1 \begin{pmatrix} 0\\1\\0\\0\\0\\0\\0\\0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0\\0\\1\\0\\0\\0\\0\\0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0\\0\\0\\-\frac{2}{5}\\1\\0\\0\\0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0\\0\\0\\-\frac{3}{5}\\0\\1\\0\\0 \end{pmatrix} + \lambda_5 \begin{pmatrix} -5\\0\\0\\-\frac{4}{5}\\0\\0\\1\\0 \end{pmatrix}
$$

mit $\lambda_i \in K$. (hier z.B. $K = \mathbb{Q}$). *Man schreibt deshalb auch*

$$
\begin{pmatrix} x_1\\x_2\\x_3\\x_4\\x_5\\x_6\\x_7\\x_8 \end{pmatrix} = \begin{pmatrix} 0\\0\\0\\-1\\0\\0\\0\\3 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} 0\\1\\0\\0\\0\\0\\0\\0 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} 0\\0\\1\\0\\0\\0\\0\\0 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} 0\\0\\0\\-\frac{2}{5}\\1\\0\\0\\0 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} 0\\0\\0\\-\frac{3}{5}\\0\\1\\0\\0 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} -5\\0\\0\\-\frac{4}{5}\\0\\0\\1\\0 \end{pmatrix}
$$

- Den Prozess elementarer Zeilenumformungen bis hin zur Zeilenstufenform nennt man meist *Vorwärtselimination*
- den Vorgang des Ablesens und Einsetzens von der letzten Gleichung bis hin zur ersten *Rückwärtssubstitution*
- und beides zusammen wird meist *Gauß-Algorithmus* (nach *Johann Carl Friedrich Gauß (1777-1855)* zur Lösung linearer Gleichungssysteme genannt.

Abbildung 2: Carl Friedrich Gauß

Beispiel 2.5.3 (Fortsetzung von 2.4.3)**.** *In 2.4.3 ist die Zeilenstufenform*

$$
\left(\begin{array}{ccccc|c}
1 & 1 & 0 & 0 & 0 & 20 \\
0 & -1 & -1 & 0 & 0 & -40 \\
0 & 0 & -1 & -1 & 0 & -50 \\
0 & 0 & 0 & 0 & -1 & -50
\end{array}\right)
$$

berechnet worden. Es ist $m = 4$, $n = 5$.
Wegen $\text{Rang}(A) = 4 = \text{Rang}(A|b)$ *ist das LGS lösbar.*
Man liest ab

$$r_5 = 50\,,\; r_3 = 50 - r_4\,,\; r_2 = 40 - r_3\,,\; r_1 = 20 - r_2\,.$$

Rückwärtssubstitution liefert

$$\begin{aligned} r_5 &= 50\,,\; r_3 = 50 - r_4\,,\; r_2 = 40 - (50 - r_4) = -10 + r_4\,,\\ r_1 &= 20 - r_2 = 20 - (-10 + r_4) = 30 - r_4\,, \end{aligned}$$

oder mit $\lambda := r_4 \in \mathbb{R}$ zusammen

$$\begin{pmatrix} r_1\\ r_2\\ r_3\\ r_4\\ r_5 \end{pmatrix} = \begin{pmatrix} 30\\ -10\\ 50\\ 0\\ 50 \end{pmatrix} + \lambda \begin{pmatrix} -1\\ 1\\ -1\\ 1\\ 0 \end{pmatrix}$$

Für $\lambda = 30$ erhält man z.B. die bereits in 2.1.1 erwähnte Lösung $r_1 = 0$, $r_2 = 20$, $r_3 = 20$, $r_4 = 30$ und $r_1 = 50$. Da $\lambda \in \mathbb{R}$ unendlich viele Werte annehmen kann, gibt es neben dieser Lösung unendlich viele andere Lösungen dieses LGS.

- Für einige Anwendungen ist es sinnvoll, nach der Vorwärtselimination noch eine *Rückwärtselimination* durchzuführen, d.h. dafür zu sorgen, dass auch in den Spalten über den Pivotelementen nur Nullen stehen.
- Dies erreicht man durch elementare Zeilenumformungen startend mit dem letzten Pivotelement a_{r,j_r}, ohne die bisher erreichte Stufenform zu zerstören
- Die Matrix aus 2.5.2 lässt sich z.B. so noch weiter vereinfachen, indem man mit der 4 in der 8.Spalte die 5 und die 6 darüber eliminiert
- Das Endprodukt nach abgeschlossener Rückwärtselimination heißt auch *reduzierte Zeilenstufenform* der Matrix

Beispiel 2.5.4 (Fortsetzung von 2.5.3)**.** *In 2.4.3 ist die Zeilenstufenform*

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20\\ 0 & -1 & -1 & 0 & 0 & -40\\ 0 & 0 & -1 & -1 & 0 & -50\\ 0 & 0 & 0 & 0 & -1 & -50 \end{array}\right)$$

berechnet worden. Mit weiteren elementaren Zeilenumformungen erreicht man die reduzierte Zeilenstufenform:

$$\underset{\rightsquigarrow}{Z_2 \leftarrow Z_2 - Z_3} \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20\\ 0 & -1 & 0 & 1 & 0 & 10\\ 0 & 0 & -1 & -1 & 0 & -50\\ 0 & 0 & 0 & 0 & -1 & -50 \end{array}\right) \underset{\rightsquigarrow}{Z_1 \leftarrow Z_1 + Z_2} \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & 0 & 30\\ 0 & -1 & 0 & 1 & 0 & 10\\ 0 & 0 & -1 & -1 & 0 & -50\\ 0 & 0 & 0 & 0 & -1 & -50 \end{array}\right)$$

An der reduzierten Zeilenstufenform kann man das Ergebnis ohne Rückwärtssubstitution direkt ablesen!

2.6 Rang, Lösbarkeit und Struktur

Satz 2.6.1. *Ein homogenes LGS $Ax = 0$ ist immer lösbar.*
Es hat zumindest immer die so genannte triviale Lösung $x = (0, \ldots, 0)$.
Ein inhomogenes LGS kann dagegen unlösbar sein, und zwar genau dann, wenn der Rang der erweiterten Matrix $(A|b)$ größer als der Rang der Matrix A ist, i.Z. $\;Ax = b$ *ist nicht lösbar* $\iff \text{Rang}(A|b) > \text{Rang}(A)$

Beweis: Die Aussage über das homogene LGS ist durch Einsetzen zu überprüfen.
Es gilt für alle A und b, dass $\text{Rang}(A|b) \geq \text{Rang}(A) := r$. Ist $\text{Rang}(A|b) > \text{Rang}(A)$, so liest man an der $r+1$-ten Zeile der Zeilenstufenform einen Widerspruch ab (vgl. 2.5.1), d.h. das LGS ist nicht lösbar.
Ist $\text{Rang}(A|b) = \text{Rang}(A)$, so kann man sofort ein Ergebnis ablesen (vgl. 2.5.2), d.h. im Widerspruch zur Annahme ist das LGS lösbar. $\square$

Satz 2.6.2. *Gilt für das homogene LGS* $Ax = 0 \iff \sum_{j=1}^n a_{ij}x_j = 0\,,\ i = 1, \ldots, m$ *die Bedingung* $m < n$, *d.h. gibt es weniger Gleichungen als Unbekannte, so gibt es eine Lösung mit mindestens* $n - m$ *freien Parametern. Lässt sich die Anzahl der Gleichungen mittels elementarer Zeilenumformungen auf eine Zeilenstufenform mit* $\mathrm{Rang}(A) = r \leq m$ *Gleichungen umformen, so ist die Anzahl der Parameter im Ergebnis gleich* $n - r$.

Beweis: Die Anzahl der Stufen in einer Zeilenstufenform, und somit auch die Anzahl der Pivotelemente, in einer $m \times n$-Matrix ist $\mathrm{Rang}(A) = r \leq m$. Somit gibt es $n - r \geq n - m$ Spalten ohne Pivotelemente, zu denen die Variablen frei gewählt werden können. □

Folgerung 2.6.3. *Aus den beiden vorhergehenden Sätzen fasst man die folgenden Lösungskriterien für ein LGS*

$$Ax = b \iff \sum_{j=1}^{n} a_{ij}x_j = b_i\,,\ i = 1, \ldots, m$$

zusammen:

$$\boxed{\begin{array}{rcl} & & Ax = b \ \ \textit{hat} \\ \mathrm{Rang}(A) < \mathrm{Rang}(A|\,b) & \iff & \textit{keine Lösung} \\ r = \mathrm{Rang}(A) = \mathrm{Rang}(A|\,b) < n & \iff & \textit{eine } n-r\textit{-parametrige Lösung} \\ r = \mathrm{Rang}(A) = \mathrm{Rang}(A|\,b) = n & \iff & \textit{eine eindeutige Lösung} \end{array}}$$

Der letzte Fall ist wegen $r \leq \min\{m, n\}$ *nur möglich, wenn* $r = n \leq m$ *ist, d.h. wenn* A *eine so genannte quadratische Matrix mit* $m = n$ *ist oder mehr Zeilen als Spalten hat.*

Satz 2.6.4. *Es seien* K *ein Körper,* $A = (a_{ij}) \in K^{m \times n}$ *und* $0 \neq b = (b_1, \ldots, b_m) \in K^m$. *Dann gilt für die Lösungen des inhomogenen LGS* $Ax = b$ *und des zugehörigen homogenen LGS* $Ax = 0$:

$$\boxed{\begin{array}{c}\textit{Spezielle Lösung} \\ \textit{des inhomogenen LGS}\end{array} + \begin{array}{c}\textit{Allgemeine Lösung} \\ \textit{des homogenen LGS}\end{array} = \begin{array}{c}\textit{Allgemeine Lösung} \\ \textit{des inhomogenen LGS}\end{array}}$$

Beweis: Es sei $(y_1, \ldots, y_n) \in K^n$ eine spezielle Lösung des inhomogenen LGS, d.h.

$$\sum_{j=1}^{n} a_{ij}y_j = b_i\,,\ i = 1, \ldots, m\,.$$

Ist $(z_1, \ldots, z_n) \in K^n$ eine beliebige Lösung des inhomogenen LGS, also

$$\sum_{j=1}^{n} a_{ij}z_j = b_i\,,\ i = 1, \ldots, m\,,$$

so folgt $\sum_{j=1}^n a_{ij}(y_j - z_j) = 0\,,\ i = 1, \ldots, m\,,$
d.h. $(y_1 - z_1, \ldots, y_n - z_n) \in K^n$ ist eine Lösung der homogenen Gleichung. Ist dagegen $(w_1, \ldots, w_n) \in K^n$ eine beliebige Lösung des homogenen LGS, also

$$\sum_{j=1}^{n} a_{ij}w_j = 0\,,\ i = 1, \ldots, m\,,$$

so ist $(w_1 + y_1, \ldots, w_n + y_n) \in K^n$ eine Lösung des inhomogenen LGS, denn

$$\sum_{j=1}^{n} a_{ij}(w_j + y_j) = \sum_{j=1}^{n} a_{ij}w_j + \sum_{j=1}^{n} a_{ij}y_j = 0 + b_i = b_i\,,\ i = 1, \ldots, m\,.$$

□

3 Naïve Mengenlehre

3.1 Grundbegriffe

Georg Cantor (1845-1918) gab die folgende

Definition 3.1.1. *Eine Menge ist jede Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens – welche die Elemente dieser Menge genannt werden – zu einem Ganzen.*

- Das Wort „Definition“ sollte hier in Anführungszeichen stehen, da der zu erklärende Begriff mittels weiterer undefinierte Begriffe erläutert wird. Eine echte Definition dagegen darf nur bereits definierte Begriffe zu Grunde legt.
- Der so eingeführte Mengenbegriff führt schnell zu Unklarheiten und logischen Widersprüchen.
- Später wurde von *Ernst Zermelo (1871-1953)* und *Adolf Fraenkel (1891-1965)* eine axiomatische Begründung der Mengenlehre geliefert, die bisher noch zu keinen Widersprüchen geführt hat (siehe etwa http://mathworld.wolfram.com/Zermelo-FraenkelAxioms.html)
- Dies soll hier nicht weiter vertieft werden. Mit einigen Vorsichtsmaßnahmen kann man ganz gut mit der Cantorschen Auffassung leben.

Eine Menge M wird meist auf eine der 3 folgenden Arten gegeben:

(1) Durch eine unmissverständliche *verbale Formulierung*:
Beispiel: M sei die Menge der im WS 2010/2011 an der TU München eingeschriebenen Studenten.

(2) Durch *Aufzählen* (falls möglich, bei $\mathbb{R}$ geht das z.B. nicht!) der Elemente $a_1, a_2, \dots$ der Menge M.
Schreibweise: $M = \{a_1, a_2, \dots\}$

(3) Durch eine *charakteristische Eigenschaft* E für die Objekte aus der Grundmenge G, die genau den Elementen von M zukommt:
Schreibweise: $M := \{x \mid x \in G \text{ erfüllt } E\}$ oder $M := \{x \in G \mid x \text{ erfüllt } E\}$.
Beispiel: $M := \{x \in \mathbb{N} \mid 3 \text{ teilt } x\}$.

Nach *Bertrand Russel (1872-1970)* ist das folgende Beispiel (das so genannte *Russelsche Paradoxon*) benannt

- Es sei
$$M := \{x \mid x \text{ ist eine Menge und } x \notin x\}$$
d.h. die Menge aller Mengen, die sich nicht selbst enthalten. Das ist eine Menge nach der Cantorschen „Definition“.
- Aber es gilt:
$$M \in M \Rightarrow M \notin M \qquad M \notin M \Rightarrow M \in M\,.$$
- Die Bezeichnung „Paradoxon“ ist also etwas irreführend, denn es handelt sich um einen dicken logischen *Widerspruch!* Deshalb wurde die Zermelo-Fraenkel-Theorie nötig.

3.2 Beziehungen zwischen Mengen

Definition 3.2.1. *Es seien A und B Mengen. A heißt Teilmenge von B oder B Obermenge von A, in Zeichen $A \subseteq B$, wenn gilt*
$$A \subseteq B :\Longleftrightarrow \text{ für alle } x \in A \text{ gilt } x \in B\,.$$
A und B heißen gleich, i.Z. $A = B$, wenn gilt
$$A = B :\Longleftrightarrow A \subseteq B \text{ und } B \subseteq A\,.$$

- Das Zeichen $:\Longleftrightarrow$ steht abkürzend für „wird definiert als…“
- Mit Hilfe von so genannten Quantoren und Junktoren (s. entsprechender Abschnitt) lassen sich solche Aussagen kompakter, aber auch erst mal schwieriger lesbar, notieren.

- Vorsicht: Verschiedene Schreibweisen je nach Autor:
  bei vielen (bei mir auch): $A \subset B \iff A \subseteq B$
  bei anderen: $A \subset B \iff A \subseteq B \wedge A \neq B$
  Ich schreibe, falls ich einmal $A \neq B$ habe, zur Vermeidung von Verwechslungen:
  $A \subsetneq B \iff A \subseteq B \wedge A \neq B$.

Definition 3.2.2. *Für eine endliche Menge M bezeichnet man die Anzahl der Elemente mit $|M|$. $|M|$ heißt die Mächtigkeit oder Kardinalität von M. Für Mengen mit einer nicht endlichen Anzahl von Elementen schreibt man $|M| = \infty$.*

- Diese Definition ist noch recht schwammig und kann erst mit Hilfe von Abbildungen genauer gefasst werden.

- Der Begriff der Kardinalität lässt sich so verfeinern, dass man noch verschiedene „Typen von Unendlich" unterscheiden kann, etwa $|\mathbb{N}| = |\mathbb{Q}|$ oder $|\mathbb{N}| \neq |\mathbb{R}|$. Das wird hier aber nicht vertieft.

4 Aussagen

4.1 Junktoren und Wahrheitstafeln

Definition 4.1.1. *Eine Aussage ist ein Element der Menge* $\{$*wahr, falsch*$\}$ *(auch* $\{w, f\}$, $\{$*true, false*$\}$ *oder* $\{$*1, 0*$\}$ *etc. geschrieben).*
Eine Aussageform ist eine Abbildung einer (oft nicht explizit angegebenen) Menge in diese zweielementige Menge.

Beispiel 4.1.2. *Die Aussage „Die Quersumme von* 28 *ist durch* 3 *teilbar“ ist falsch.*
Die Aussage „Die Quersumme von 27 *ist durch* 3 *teilbar“ ist wahr.*
$A(x) :=$*„Die Quersumme von* x *ist durch* 3 *teilbar“ ist eine Aussageform, die z.B. für* $x = 28$ *eine falsche, für* $x = 27$ *eine wahre Aussage liefert.*

Sind A und B zwei Aussagen, so lassen sich mittels so genannter *Junktoren* (=Verbinder) neue Aussagen zusammenstellen:

Junktor	Bedeutung	Zeichen
Negation	nicht A	$\neg A$
Konjugation	A und B	$A \wedge B$
Disjunktion	A oder B	$A \vee B$
Implikation	wenn A, dann B	$A \Rightarrow B$
Äquivalenz	A genau dann, wenn B	$A \iff B$

Die Wahrheitswerte der so zusammengesetzten Aussagen werden mit Hilfe einer *Wahrheitstabelle* angegeben, in der zu jeder wahr-falsch-Kombination der beteiligten Aussagen der Wahrheitswert der zusammengesetzten Aussage angegeben wird.

Definition 4.1.3.

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \iff B$
w	w	f	w	w	w	w
w	f	f	f	w	f	f
f	w	w	f	w	w	f
f	f	w	f	f	w	w

Definition 4.1.4. *Bei einer Implikation* $A \Rightarrow B$ *heißt* A *Voraussetzung und* B *Behauptung. Ist die Implikation wahr, so nennt man* A *hinreichend für* B *und* B *notwendig für* A.

- Über Wahrheitstafeln kann man (→Übungen) nun einfache Rechenregeln für diese Junktoren beweisen, z.B.:
- $A \Rightarrow B$ lässt sich auch durch die davor eingeführten Junktoren definieren als $\neg A \vee B$, d.h. man hätte definieren können
$$(A \Rightarrow B) :\iff (\neg A \vee B)$$
- $A \iff B$ steht für $A \Rightarrow B \wedge B \Rightarrow A$ (letzteres auch als $A \Leftarrow B$ geschrieben), also
$$(A \iff B) :\iff (A \Rightarrow B \wedge B \Rightarrow A)$$

Satz 4.1.5 (De Morgansche Regeln). *Für Aussagen* A *und* B *gilt*

$$(\neg(A \wedge B)) \iff (\neg A \vee \neg B) \tag{4.1}$$
$$(\neg(A \vee B)) \iff (\neg A \wedge \neg B) \tag{4.2}$$

Beweis:

A	B	$\neg A$	$A \wedge B$	$\neg(A \wedge B)$	$\neg B$	$\neg A \vee \neg B$
w	w	f	w	f	f	f
w	f	f	f	w	w	w
f	w	w	f	w	f	w
f	f	w	f	w	w	w

□

4.2 Quantoren

Aus Aussageformen kann man durch so genannte *Quantoren* Aussagen machen.

Definition 4.2.1. *Der „für alle“-Quantor* $\forall$*:*

$$(\forall x \in M : A(x)) := \begin{cases} w & \text{falls } f \notin A(M), \\ f & \text{falls } f \in A(M). \end{cases}$$

Beispiel 4.2.2. *Es sei* $A(x)$ *die Aussageform „Die Quersumme der Zahl* x *ist durch* 3 *teilbar“ und* M *die Menge aller Vielfachen von* 3*. Dann ist* $A(x)$ *für alle* $x \in M$ *wahr, d.h.* $A(M) = \{w\}$*. Die Aussage* $\forall x \in M : A(x)$ *(in Worten: „Für alle Vielfachen von* 3 *gilt, dass ihre Quersumme durch* 3 *teilbar ist“), ist also wahr.*
Für die Grundmenge $M = \{27, 28, 29\}$ *ist dagegen* $A(M) = \{w, f\}$ *(nämlich wahr für* 27 *und falsch für* 28*,* 29*), also die Aussage* $\forall x \in M : A(x)$ *falsch.*

Definition 4.2.3. *Der „es existiert“-Quantor* $\exists$*:*

$$(\exists x \in M : A(x)) := \begin{cases} w & \text{falls } w \in A(M), \\ f & \text{falls } w \notin A(M). \end{cases}$$

Beispiel 4.2.4. *Für die Grundmenge* $M = \{27, 28, 29\}$ *und die Aussageform* $A(x)$ *„Die Quersumme der Zahl* x *ist durch* 3 *teilbar“ ist* $A(M) = \{w, f\}$*, also die Aussage* $\exists x \in M : A(x)$ *wahr.*

Dieses Beispiel zeigt insbesondere, dass das mathematische „es gibt ein“ bedeutet „es gibt *mindestens* ein“. Möchte man ausdrücken, dass es *genau ein* Element mit der angegebenen Eigenschaft gibt, so versieht man den Existenzquantor $\exists$ mit einer kleinen tiefgestellten 1:

Beispiel 4.2.5. $\exists_1\, p$ *prim:* $19 < p \leq 28$*.*
Folgerung 4.2.6. *Aus der Definition der beiden Quantoren folgt:*

$$\neg(\forall x \in M : A(x)) = (\exists x \in M : \neg A(x))$$
$$\neg(\exists x \in M : A(x)) = (\forall x \in M : \neg A(x))$$

Beispiel 4.2.7. *Die Verneinung der (falschen) Aussage*
„Für alle $x \in \{27, 28, 29\}$ *ist die Quersumme durch* 3 *teilbar“*
ist die (wahre) Aussage
„Es gibt ein $x \in \{27, 28, 29\}$*, dessen Quersumme nicht durch* 3 *teilbar ist“.*
Die Verneinung der (falschen) Aussage
„Es gibt eine natürliche Zahl n *mit* $n < -4$*“*
ist die (wahre) Aussage
„Für alle natürlichen Zahlen n *ist* $n \geq -4$*“.*

Satz 4.2.8.

$$(A \Rightarrow B) \iff (\neg B \Rightarrow \neg A)$$

Beweis: mit einfacher Tabelle wie zuvor oder so:

$$\begin{aligned} (A \Rightarrow B) &\iff (\neg A \vee B) \iff (\neg A \vee \neg\neg B) \iff \\ &\iff (\neg\neg B \vee \neg A) \overset{(4.2)}{\iff} (\neg(\neg B) \vee (\neg A)) \iff \\ &\iff (\neg B \Rightarrow \neg A) \end{aligned}$$

□

4.3 Beweistechniken

- Die folgenden Beweistechniken werden unterschieden:
  (1) Direkter Beweis: $A \Rightarrow B$ (*modus ponens*),
  (2) Indirekter Beweis: $\neg B \Rightarrow \neg A$ (*modus tollens*),

(3) Widerspruchsbeweis: $\neg A$ ist falsch, also gilt A,

(4) vollständige Induktion,

(5) Konstruktiver Beweis (bei Existenzaussagen).

- Eine axiomatische Definition der natürlichen Zahlen ist nach *Giuseppe Peano (1858-1932)* benannt, der diese Axiome aufbauend auf Arbeiten von *Richard Dedekind (1831-1916)* erstmals formuliert und benutzt hat:
- Eines der Peano-Axiome ist das so genannte *Induktionsprinzip*, mit dem die Induktion begründet wird:

Satz 4.3.1. *Es sei $A(n)$ eine Aussage über alle natürlichen Zahlen $n \geq a$, wobei a eine fest gewählte natürliche Zahl sei. $A(n)$ ist für alle $n \geq a$ richtig, wenn die folgenden 2 Bedingungen erfüllt sind:*

(I_1) *$A(a)$ ist richtig.*

(I_2) *Aus der Annahme der Gültigkeit von $A(n)$ für alle n mit $a \leq n \leq k$ folgt stets die Gültigkeit von $A(k+1)$.*

- Das Prinzip der vollständigen Induktion findet Anwendung, wenn Behauptungen des Typs „Für alle $n \in \mathbb{N}$ gilt ...“ bewiesen werden.
- Genauso wie bei der Begründung der vollständigen Induktion kann man aber auch vorgehen, um mathematische Objekte $O(n)$ *rekursiv* zu definieren:
- Man gibt $O(a), \dots, O(b)$ an und ein Verfahren, wie man jeweils $O(n)$ aus den $O(k)$ mit $k < n$ erhält. Wegen des eben gezeigten Prinzips der vollständigen Induktion ist damit $O(n)$ für alle natürlichen Zahlen $n \geq a$ definiert.

5 Rechnen mit Mengen

5.1 Grundlagen

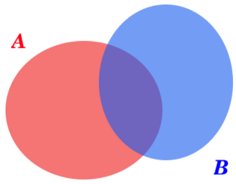


Abbildung 3: Venn-Diagramm

- Im Folgenden seien A und B zwei Mengen. Die eingeführten Operationen zwischen diesen Mengen werden verdeutlicht durch so genannte Venn-Diagramme (nach *John Venn (1834-1923)*), die auch eulersche Kreise genannt werden (nach *Leonhard Euler (1701-1783)*).
- Diese Darstellung kann beim Nachweis von Rechengesetzen etwas die Vorstellung unterstützen, nicht aber den formalen Beweis ersetzen, der mit Hilfe der eingeführten Junktoren oder mit Wahrheitstabellen geführt wird!
- Mit den eingeführten Quantoren könnte man 3.2.1 jetzt auch in der folgenden Form schreiben:

$$A \subseteq B :\Longleftrightarrow \forall x \in A \ : \ x \in B\,,$$
$$A = B :\Longleftrightarrow A \subseteq B \wedge B \subseteq A\,.$$

- Daraus folgt

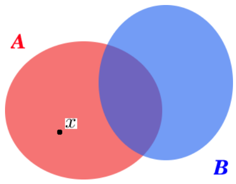


Abbildung 4: Ungleiche Mengen

$$\begin{aligned} A \neq B \iff \neg(A = B) \iff \neg(A \subseteq B \wedge B \subseteq A) &\overset{4.2.6}{\iff} \\ \iff \neg(A \subseteq B) \vee \neg(B \subseteq A) &\overset{(4.1)}{\iff} \\ \iff (\exists x \in A \ : \ x \notin B) \vee (\exists x \in B \ : \ x \notin A) & \end{aligned}$$

- Einige wichtige Mengen, die im folgenden Stoff noch mit sehr viel mehr mathematischem Leben gefüllt werden, sind

$$\begin{aligned} \mathbb{N} :=& \{1, 2, 3, \dots\} \ \text{ Menge der } \textit{natürlichen Zahlen}, \\ \mathbb{N}_0 :=& \{0, 1, 2, \dots\} \ \text{ Menge der } \textit{natürlichen Zahlen und } 0, \\ \mathbb{Z} :=& \{0, -1, 1, -2, 2, \dots\} \ \text{ Menge der } \textit{ganzen Zahlen}, \\ \mathbb{Q} :=& \{x \mid \exists a \in \mathbb{Z}\ \exists b \in \mathbb{N} \ : \ x = \frac{a}{b}\} \ \text{ Menge der } \textit{rationalen Zahlen}, \\ \mathbb{R} :=& \{x \mid x \ \text{ ist eine Dezimalbruch (endlich oder unendlich)}\} \\ & \text{Menge der } \textit{reellen Zahlen}, \\ \mathbb{C} :=& \{z \mid \exists a \in \mathbb{R}\ \exists b \in \mathbb{R} \ : \ z = a + i \cdot b \wedge i^2 = -1\} \\ & \text{Menge der } \textit{komplexen Zahlen}, \\ \emptyset :=& \{x \mid x \neq x\} \ \textit{ leere Menge.} \end{aligned}$$

- Jede Menge enthält zwei *triviale Teilmengen*, nämlich die leere Menge und sich selbst.
- Man beachte:

$$\{a\} \notin \{a\}\,,\ \{a\} \subset \{a\}\,,\ a \in \{a\}\,,\ \{a\} \neq \{\{a\}\}\,.$$

- $\subset$ bzw. $\subseteq$ sind *transitiv*:

$$A \subset B \wedge B \subset C \Rightarrow A \subset C \ \text{ bzw. } \ A \subseteq B \wedge B \subseteq C \Rightarrow A \subseteq C$$

Definition 5.1.1. *Es sei A eine Menge. Die Menge*

$$\mathfrak{P}(A) := \{M \mid M \subseteq A\}$$

aller Teilmengen von A heißt Potenzmenge von A.

Beispiel 5.1.2. $A = \{0, 1\} \Rightarrow \mathfrak{P}(A) = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$.

5.2 Operationen

Definition 5.2.1.

Der Durchschnitt oder einfach Schnitt zweier Mengen A und B ist

$$A \cap B := \{x \mid x \in A \wedge x \in B\}$$

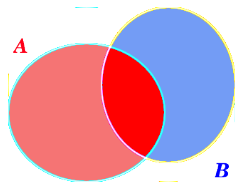


Abbildung 5: Mengendurchschnitt

Definition 5.2.2.

Der Vereinigung zweier Mengen A und B ist

$$A \cup B := \{x \mid x \in A \vee x \in B\}$$

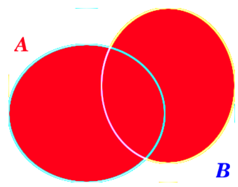


Abbildung 6: Mengenvereinigung

Definition 5.2.3.

Die Differenz der beiden Mengen A und B ist

$$A \setminus B := \{x \mid x \in A \wedge x \notin B\}$$

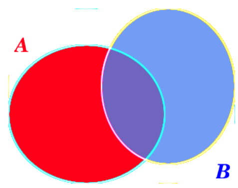


Abbildung 7: Mengendifferenz

Definition 5.2.4.

Ist A Teilmenge einer Menge G (oft Grundmenge genannt), also $A \subseteq G$, so nennt man die Differenz $G \setminus A$ auch das Komplement von A in G, in Zeichen:

$$\overline{A} := G \setminus A = \{x \in G \mid x \notin A\}\,.$$

Oft findet man auch die Schreibweise $\complement A$ für $\overline{A}$.

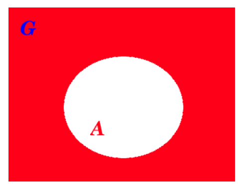


Abbildung 8: Mengenkomplement

Definition 5.2.5.

Die symmetrische Differenz der beiden Mengen A und B ist

$$A \triangle B := (A \setminus B) \cup (B \setminus A)$$

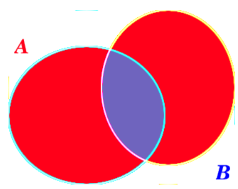


Abbildung 9: Symmetrische Differenz

- Die Mengenoperationen kann man auch mit Hilfe von Wahrheitstabellen recht kompakt darstellen: Man untersucht die Wahrheitswerte der Aussage $x \in M$.

- Statt $\begin{array}{|c|}\hline x \in M \\ \hline w \\ f \\ \hline\end{array}$ schreibt man kürzer $\begin{array}{|c|}\hline M \\ \hline w \\ f \\ \hline\end{array}$ und kann so alle Operationen in einer Tabelle darstellen:

-

A	B	$A\cap B$	$A\cup B$	$A\setminus B$	$A\Delta B$
f	f	f	f	f	f
f	w	f	w	f	w
w	f	f	w	w	w
w	w	w	w	f	f

Satz 5.2.6. *Für beliebige Teilmengen A, B und C eines Grundbereichs G gilt:*

$$\left.\begin{aligned}(A\cap B)\cap C &= A\cap(B\cap C)\\ (A\cup B)\cup C &= A\cup(B\cup C)\\ (A\Delta B)\Delta C &= A\Delta(B\Delta C)\end{aligned}\right\}\ \textit{Assoziativgesetze}$$

Satz 5.2.7. *Für beliebige Teilmengen A, B und C eines Grundbereichs G gilt:*

$$\left.\begin{aligned}A\cap B &= B\cap A\\ A\cup B &= B\cup A\\ A\Delta B &= B\Delta A\end{aligned}\right\}\ \textit{Kommutativgesetze}$$

Satz 5.2.8. *Für beliebige Teilmengen A, B und C eines Grundbereichs G gilt:*

$$\left.\begin{aligned}A\cap(B\cup C) &= (A\cap B)\cup(A\cap C)\\ A\cup(B\cap C) &= (A\cup B)\cap(A\cup C)\end{aligned}\right\}\ \textit{Distributivgesetze}$$

Satz 5.2.9. *Für beliebige Teilmengen A, B und C eines Grundbereichs G gilt:*

$$\left.\begin{aligned}A\cup A &= A\\ A\cap A &= A\\ \complement(\complement A) = \overline{\overline{A}} &= A\end{aligned}\right\}\ \textit{Idempotenzgesetze}$$

Satz 5.2.10. *Für beliebige Teilmengen A, B und C eines Grundbereichs G gilt:*

$$\left.\begin{aligned}(A\setminus B)\cap C &= (A\cap C)\setminus(B\cap C)\\ A\setminus(B\cap C) &= (A\setminus B)\cup(A\setminus C)\\ (A\setminus B)\cup B &= A\cup B\\ A\setminus(B\cup C) &= (A\setminus B)\cap(A\setminus C)\end{aligned}\right\}\ \textit{Rechengesetze für Mengendifferenzen}$$

Satz 5.2.11. *Für beliebige Teilmengen A, B und C eines Grundbereichs G gilt:*

$$\left.\begin{aligned}\complement(A\cap B) = \overline{A\cap B} &= \overline{A}\cup\overline{B} = \complement A\cup\complement B\\ \complement(A\cup B) = \overline{A\cup B} &= \overline{A}\cap\overline{B} = \complement A\cap\complement B\end{aligned}\right\}\ \textit{De Morgansche Regeln}$$

- Praktisch alle aufgezählten Rechengesetze lassen sich durch einfaches Nachrechnen über Junktoren beweisen oder mit Hilfe einer geeigneten Wahrheitstabelle.
- Einige Rechengesetze werden in den Übungen gezeigt.

5.3 Das kartesische Produkt

- Es seien A und B nichtleere Mengen, $a, a' \in A$ und $b, b' \in B$.
- Man kürzt ab $(a, b) := \{\{a\}, \{a, b\}\}$.
- Mit Hilfe der Axiome der Mengenlehre (Zermelo-Fraenkel) kann man dann zeigen: $(a, b) = (a', b') \iff a = a' \wedge b = b'$.
- Das neu definierte Objekt (a, b) nennt man ein *geordnetes Paar*, denn im Gegensatz zu Mengen ($\{a, b\} = \{b, a\}$) kommt es hier auf die Reihenfolge an ($(a, b) \neq (b, a)$).
- Da wir nicht in die axiomatische Mengenlehre einsteigen wollen, sollte man diese Definition eines geordneten Paares gleich wieder vergessen.

- Mit der folgenden etwas schwammigen Definition kann man ganz gut zurecht kommen; es ist aber gut zu wissen, dass sich das auch axiomatisch korrekt einführen lässt.

Definition 5.3.1. *Es seien A und B nichtleere Mengen, $a, a' \in A$ und $b, b' \in B$. Ein geordnetes Paar oder Tupel von Elementen $a \in A$ und $b \in B$ wird (a, b) geschrieben. Zwei solche Tupel heißen gleich, wenn sie elementweise übereinstimmen, i.Z.*

$$(a, b) = (a', b') :\iff a = a' \wedge b = b' .$$

Definition 5.3.2. *Es seien A und B nichtleere Mengen. Die Menge aller geordneten Paare (a, b) mit $a \in A$ und $b \in B$ heißt kartesisches Produkt oder Kreuzprodukt der Mengen A und B, i.Z.*

$$A \times B := \{(a, b) \mid a \in A \wedge b \in B\} .$$

Definition 5.3.3. *Es seien $A_1, A_2, \dots A_n$ mit $n \in \mathbb{N}$ und $n \geq 2$ nichtleere Mengen. Man betrachtet n-Tupel $(a_1, a_2, \dots, a_n)$ von Elementen $a_i \in A_i$ für $i = 1, \dots, n$. Zwei solche n-Tupel heißen gleich, wenn sie elementweise übereinstimmen, i.Z.*

$$(a_1, a_2, \dots, a_n) = (a'_1, a'_2, \dots, a'_n) :\iff a_i = a'_i \quad \forall i \in \{1, 2, \dots, n\}$$

Definition 5.3.4. *Die Menge aller geordneten n-Tupel $(a_1, a_2, \dots, a_n)$ heißt n-faches kartesisches Produkt oder n-faches Kreuzprodukt der Mengen $A_1, A_2, \dots, A_n$, i.Z.*

$$A_1 \times A_2 \times \cdots \times A_n := \{(a_1, a_2, \dots, a_n) \mid a_i \in A_i \forall i \in \{1, 2, \dots, n\}\} .$$

- Bei n-fachen kartesischen Produkten einer Menge mit sich selbst benutzt man auch die Potenzschreibweise, also
$$A^n := \underbrace{A \times \cdots \times A}_{n\text{-mal}} .$$
- Die Schreibweise $\mathbb{R}^3$ für alle 3-Tupel reeller Zahlen ist ja bereits aus der Schule bekannt.
- Eine $n \times 1$-Matrix$\in K^{n\times 1}$ oder eine $1 \times n$-Matrix$\in K^{1\times n}$ können prinzipiell mit einem n-Tupel des K^n identifiziert werden.
- Für das spätere Rechnen mit Matrizen kommt es aber auf die Schreibweise als Spalte oder als Zeile an.

Beispiel 5.3.5. *Für $A = \{1, 2, 3\}$ und $B = \{a, b\}$ ist*

$$\begin{aligned} A \times B =& \{(1, a), (1, b), (2, a), (2, b), (3, a), (3, b)\} , \\ B^2 =& \{(a, a), (a, b), (b, a), (b, b)\} . \end{aligned}$$

Satz 5.3.6. *Sind $A_1, A_2, \dots A_n$ nichtleere endliche Mengen, so gilt*

$$|A_1 \times A_2 \times \cdots \times A_n| = |A_1| \cdot |A_2| \cdot \cdots \cdot |A_n| .$$

Beweis: durch einfaches Abzählen. □

6 Relationen und Funktionen

6.1 Binäre Relationen

Definition 6.1.1. *Es seien $k \in \mathbb{N}$ und $A_1, \ldots, A_k$ Mengen. Eine Teilmenge $R \subseteq A_1 \times \cdots \times A_k$ heißt k-stellige (oder k-näre) Relation auf $A_1 \times \cdots \times A_k$.*

Da diese Definition ziemlich abstrakt ist, schreibt man bei den am häufigsten verwendeten *binären Relationen* oft statt $(a, b) \in R$ lieber aRb (lies: *a steht in Relation zu b*) und verwendet für R meist ein spezielles Zeichen (etwa $=$, $>$, $\leq$, $||$ etc.)

Beispiel 6.1.2. *Für $A = A_1 = A_2 = \{1, 2, 3\}$ ist z.B.*

$$R = \{(1, 1), (2, 2), (3, 3)\}$$

eine binäre Relation. Da zwei Elemente $a, b \in A$ offenbar genau dann in Relation stehen, wenn sie gleich sind, schreibt man statt $(a, b) \in R$ oder aRb in diesem Fall wohl suggestiver gleich $a = b$.

Beispiel 6.1.3. *Für $A = A_1 = A_2 = A_3 = \{1, 2, 3, 4, 5, 6\}$ ist*

$$R := \{(a, b, c) \in A^3 \mid a^2 + b^2 = c^2\} = \{(3, 4, 5), (4, 3, 5)\}$$

die ternäre Relation, die gerade die pythagoreischen Tripel enthält.

Definition 6.1.4. *Es sei A eine Menge und $R \subseteq A^2$ eine binäre Relation. Es werden die folgenden Zusatzbezeichnungen für R vereinbart:*

$$\begin{aligned}
\textit{reflexiv} :&\Longleftrightarrow \forall a \in A : \ (a, a) \in R \\
\textit{irreflexiv} :&\Longleftrightarrow \forall a \in A : \ (a, a) \notin R \\
\textit{symmetrisch} :&\Longleftrightarrow \forall a, b \in A : \ (a, b) \in R \Rightarrow (b, a) \in R \\
\textit{antisymmetrisch} :&\Longleftrightarrow \forall a, b \in A : \ (a, b) \in R \Rightarrow (b, a) \notin R \vee a = b \\
\textit{asymmetrisch} :&\Longleftrightarrow \forall a, b \in A : \ (a, b) \in R \Rightarrow (b, a) \notin R \\
\textit{transitiv} :&\Longleftrightarrow \forall a, b, c \in A : \ (a, b) \in R \wedge (b, c) \in R \Rightarrow (a, c) \in R
\end{aligned}$$

25.10.10

6.2 Besondere binäre Relationen

Definition 6.2.1. *Eine binäre Relation $R \subseteq A^2$ heißt*

$$\begin{aligned}
\textit{reflexive teilweise Ordnung auf } A :&\Longleftrightarrow R \textit{ ist reflexiv, transitiv,} \\
&\qquad \textit{und antisymmetrisch,} \\
\textit{irreflexive teilweise Ordnung auf } A :&\Longleftrightarrow R \textit{ ist irreflexiv und transitiv,} \\
\textit{Äquivalenzrelation auf } A :&\Longleftrightarrow R \textit{ ist reflexiv, transitiv} \\
&\qquad \textit{und symmetrisch.}
\end{aligned}$$

Beispiel 6.2.2. *Die teilweise reflexive Ordnung auf A ist eine Verallgemeinerungen der üblichen $\leq$-Beziehungen (etwa mit $A = \mathbb{N}$):*

$$\begin{array}{rll}
a \leq a & \forall a \in A & \text{d.h. } \leq \text{ ist reflexiv} \\
a \leq b \Rightarrow [b \not\leq a \vee a = b] & \forall a, b \in A & \text{d.h. } \leq \text{ ist antisymmetrisch} \\
a \leq b \wedge b \leq c \Rightarrow a \leq c & \forall a, b, c \in A & \text{d.h. } \leq \text{ ist transitiv}
\end{array}$$

Beispiel 6.2.3. *Die teilweise irreflexive Ordnung auf A ist eine Verallgemeinerungen der üblichen $<$-Beziehungen (etwa mit $A = \mathbb{N}$):*

$$\begin{array}{rll}
a < b \wedge b < c \Rightarrow a < c & \forall a, b, c \in A & \text{d.h. } < \text{ ist transitiv,} \\
a < b \Rightarrow a \neq b & \forall a, b \in A & \text{d.h. } < \text{ ist irreflexiv.}
\end{array}$$

Beispiel 6.2.4. *Es sei $M := \{a, b\}$ und $A := \mathfrak{P}(M)$. Die Relation $R = \subseteq$, oder ausführlich*

$$R := \{(X, Y) \in A^2 \mid X \subseteq Y\},$$

auf A ist eine reflexive Halbordnung.
Ausgeschrieben sieht R so aus:

$$\begin{aligned} R = \{&(\emptyset,\emptyset), (\emptyset, \{a\}), (\emptyset, \{b\}), (\emptyset, \{a, b\}), (\{a\}, \{a\}),\\ &(\{a\}, \{a, b\}), (\{b\}, \{b\}), (\{b\}, \{a, b\}), (\{a, b\}, \{a, b\}), \}.\end{aligned}$$

Dies kann man auch graphisch über ein Pfeildiagramm veranschaulichen. Dabei steht jeweils $\to$ für R:

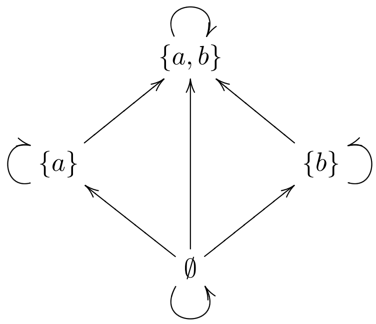


Abbildung 10: Diagramm der Relation $\subseteq$

- Da insbesondere das bekannte Gleichheitszeichen für eine Menge A von Zahlen die drei Eigenschaften *reflexiv*, *symmetrisch* und *transitiv* hat, also eine Äquivalenzrelation ist, kann man Äquivalenzrelationen als Verallgemeinerungen des Gleichheitszeichens auffassen.

- Auf jeder Menge A kann man stets die folgenden zwei *trivialen Äquivalenzrelationen* definieren:

$$\begin{aligned} R_0 :=& \{(a, a) \mid a \in A\},\\ R_1 :=& A \times A \end{aligned}$$

Definition 6.2.5. *Es sei $a \in A$ und R eine Äquivalenzrelation auf A. Die Menge*

$$[a]_R := \{b \in A \mid (a, b) \in R\}$$

(oder kurz $[a]$ geschrieben, wenn R klar ist) heißt Äquivalenzklasse von a bezüglich R.

Beispiel 6.2.6. *Für die beiden trivialen Äquivalenzrelationen gilt*

$$[a]_{R_0} = \{a\}, \; [a]_{R_1} = A \,\forall a \in A.$$

Satz 6.2.7. *Es sei R eine Äquivalenzrelation auf einer Menge A. Dann gilt*

(1) $\displaystyle\bigcup_{a\in A} [a]_R = A$,

(2) Für alle $a, b \in A$ ist $[a]_R = [b]_R \iff (a, b) \in R$,

(3) Für alle $a, b \in A$ ist $[a]_R \cap [b]_R = \emptyset \vee [a]_R = [b]_R$.

Beweis: (1) „$\subseteq$“ ist klar laut Definition 6.2.5.
(1) „$\supseteq$“ Wegen $aRa\ \forall a \in A$ gilt $a \in [a]_R\ \forall a \in A$ und somit die Behauptung.
(2) „$\Rightarrow$“: Ist $[a]_R = [b]_R$, so ist wegen $b \in [b]_R$ auch $b \in [a]_R$ und somit $(a, b) \in R$.
(2) „$\Leftarrow$“: Nun sei $(a, b) \in R$ und $c \in [b]_R$. Dann ist $(b, c) \in R$. Zusammen mit $(a, b) \in R$ und der Transitivität von R folgt $(a, c) \in R$, und damit $c \in [a]_R$, d.h. $[b]_R \subseteq [a]_R$.
Da R symmetrisch ist, ist auch $(b, a) \in R$ und man folgert völlig analog, dass $[a]_R \subseteq [b]_R$ ist, also insgesamt die Gleichheit der beiden Mengen.
(3) Angenommen, es existieren $a, b \in A$ mit $[a]_R \cap [b]_R \neq \emptyset$ und $[a]_R \neq [b]_R$. dann gibt es ein c aus dem Schnitt von $[a]_R$ und $[b]_R$ und für dieses c gilt $(a, c) \in R$ und $(b, c) \in R$ bzw. (wegen der Symmetrie) $(c, b) \in R$.
Dann folgt wegen der Transitivität $(a, b) \in R$, was aber nach (2) $[a]_R = [b]_R$ zur Folge hat, also einen Widerspruch zur Annahme. $\square$

Definition 6.2.8. *Es sei A eine nichtleere Menge. Ein $P \subseteq \mathfrak{P}(A) \setminus \{\emptyset\}$ heißt Partition von A, wenn gilt*

$$A = \bigcup_{X \in P} X \quad \wedge \tag{6.1}$$

$$\forall X, Y \in P : X \neq Y \Rightarrow X \cap Y = \emptyset \tag{6.2}$$

Oft findet man für Partitionen auch die abkürzende, an (6.1) angelehnte, Schreibweise mit einem Punkt über dem Vereinigungszeichen als Hinweis auf (6.2):

$$A = \dot{\bigcup_{X \in P}} X$$

(lies: A ist die disjunkte Vereinigung der Mengen $X \in P$).

6.3 Restklassen

- Oft ist es sinnvoll, jeweils eine ganze Äquivalenzklasse als ein einziges Objekt zu betrachten und nur mit einem besonders einfachen Vertreter aus jeder Äquivalenzklasse zu arbeiten.
- Dazu definiert man

Definition 6.3.1. *Es sei R eine Äquivalenzrelation auf einer Menge A. Die Menge*

$$A/_R := \{[a]_R \mid a \in A\}$$

heißt Faktormenge von A nach R.

6.4 Abbildungen

Definition 6.4.1. *Es seien A und B Mengen. Eine binäre Relation $f \subseteq A \times B$ heißt Abbildung oder Funktion von A nach B, wenn gilt*

(1) Für jedes $a \in A$ gibt es ein $b \in B$ mit $(a, b) \in f$,

(2) Ist $(a, b_1) \in f$ und $(a, b_2) \in f$ so ist $b_1 = b_2$.

- Die Menge $D(f) := A$ ist der so genannte *Definitionsbereich* der Abbildung f.
- Für $(a, b) \in f$ schreibt man im Fall einer Abbildung $f(a) = b$, denn durch (1) wird garantiert, dass es zu jedem $a \in A$ mindestens ein $b \in B$ mit $(a, b) \in f$ gibt. Wegen (2) ist dies sogar eindeutig, d.h. die Schreibweise $f(a)$ kann auch keine Verwirrung stiften.
- Die Menge B heißt *Bildbereich* und

  $$W(f) := \{b \in B \mid \exists a \in A : (a, b) \in f\}$$

  der *Wertebereich* von f.

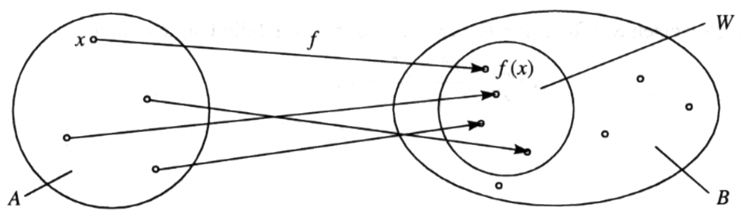


Abbildung 11: Abbildung f von A in B

Definition 6.4.2. *Eine Funktion f heißt injektiv, wenn aus $f(a_1) = b$ und $f(a_2) = b$ stets folgt, dass $a_1 = a_2$ ist.*

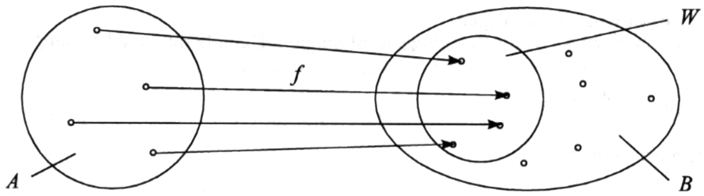


Abbildung 12: Injektive Abbildung

Definition 6.4.3. *Eine Funktion f heißt surjektiv, wenn*

$$W(f) = \{b \in B \mid \exists a \in A : (a, b) \in f\} = B$$

ist.

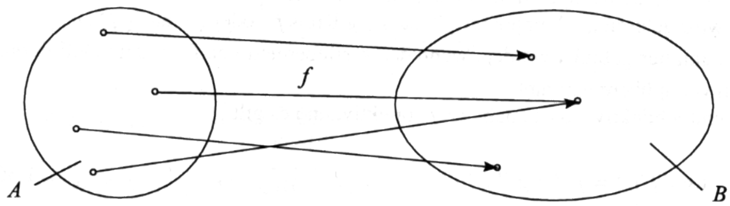


Abbildung 13: Surjektive Abbildung

Definition 6.4.4. *Eine Funktion f heißt bijektiv, wenn sie injektiv und surjektiv ist.*

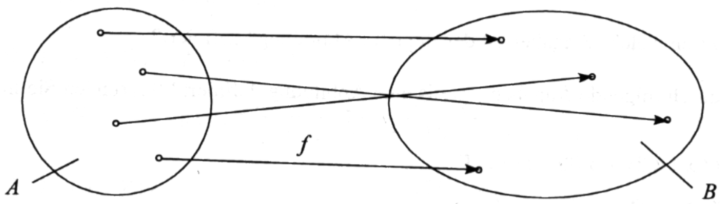


Abbildung 14: Bijektive Abbildung

- Interpretiert man eine Abbildung gemäß ihrer ursprünglichen Definition als Menge von Tupeln, so nennt man diese auch den *Graphen* der Abbildung.
- Handelt es sich um eine Abbildung $\mathbb{R} \to \mathbb{R}$ oder von Teilmengen, so kann man diese Tupel in einem entsprechend gewählten kartesischen Koordinatensystem visualisieren.
- Aus der Definition der Abbildung folgt, dass jeder Schnitt parallel zur y-Achse den Graphen einer Abbildung höchstens einmal schneiden kann.

Beispiel 6.4.5.
Die Abbildung

$$f : \begin{cases} x \mapsto x^2 \\ \mathbb{R} \to \mathbb{R} \end{cases}$$

ist weder injektiv, noch surjektiv.
Der zugehörige Wertebereich ist

$$W(f) = \mathbb{R}_+ \cup \{0\} = \{x \in \mathbb{R} \mid x \geq 0\}$$

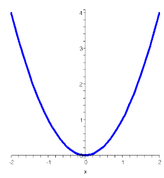


Abbildung 15: Graph von f

Beispiel 6.4.6.
Die Abbildung

$$f : \begin{cases} x \mapsto x^3 \\ \mathbb{R} \to \mathbb{R} \end{cases}$$

ist bijektiv (also $W(f) = \mathbb{R}$).

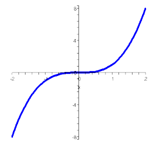


Abbildung 16: Graph von f

Beispiel 6.4.7.

Die Abbildung

$$f : \begin{cases} x \mapsto x^3 - x \\ \mathbb{R} \to \mathbb{R} \end{cases}$$

ist surjektiv (also $W(f) = \mathbb{R}$), nicht aber injektiv oder bijektiv.

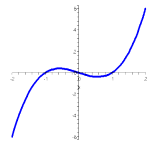


Abbildung 17: Graph von f

Beispiel 6.4.8.
Die Abbildung

$$f : \begin{cases} x \mapsto e^x = \exp(x) \\ \mathbb{R} \to \mathbb{R} \end{cases}$$

ist injektiv, nicht aber surjektiv, mit

$$W(f) = \mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$$

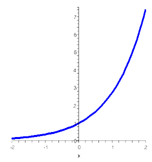


Abbildung 18: Graph von f

Definition 6.4.9. *Ist $f : M \to N$ eine Abbildung und $U \subseteq M$, so heißt*

$$f(U) := \{f(x) \mid x \in U\} \subseteq N$$

die Bildmenge von U unter der Abbildung f. Insbesondere ist $f(M) = W(f)$.

- Für alle $M_1, M_2 \subseteq M$ gilt $f(M_1 \cup M_2) = f(M_1) \cup f(M_2)$.
- Es gilt dagegen nicht allgemein das Entsprechende für $\cap$, sondern nur $f(M_1 \cap M_2) \subseteq f(M_1) \cap f(M_2)$.

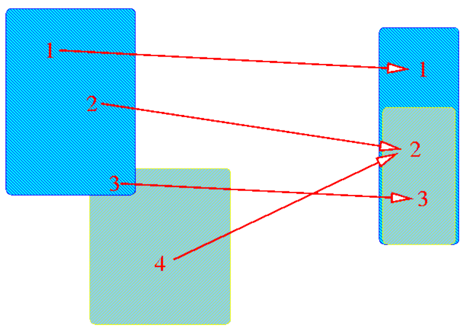


Abbildung 19: Beispiel für $f(M_1 \cap M_2) \subsetneq f(M_1) \cap f(M_2)$

Definition 6.4.10. *Für zwei Abbildungen $f : M \to N$ und $g : N \to R$ heißt*

$$g \circ f : \begin{cases} m \mapsto g(f(m)) \\ M \to R \end{cases}$$

die Produktabbildung oder Komposition von f und g. Man beachte die Reihenfolge der Ausführung von rechts nach links $g \circ f(m) = g(f(m))$, d.h. erst f, dann g!

Definition 6.4.11. *Eine bijektive Abbildung*

$$f : \begin{cases} M \to N \\ m \mapsto n \end{cases}$$

kann man umkehren, d.h es gibt die so genannte inverse Abbildung

$$f^{-1} : \begin{cases} N \to M \\ f(m) \mapsto m \end{cases}$$

Es gilt

$$f^{-1} \circ f = \mathrm{id}_M \, , \; f \circ f^{-1} = \mathrm{id}_N \, .$$

wobei id *die identische Abbildung auf der jeweiligen Menge bedeutet, also* $\mathrm{id}_M(m) = m \; \forall m \in M$.

- Ist die Abbildung $f : M \to N$ injektiv, so ist $\tilde{f} : M \to f(M)$ mit $\tilde{f}(m) = f(m)$ für alle $m \in M$ sogar bijektiv und somit invertierbar, d.h. injektive Abbildungen kann man auf Ihrem Bild umkehren.

Definition 6.4.12. *Die Einschränkung oder Restriktion einer Abbildung $f : M \to N$ auf eine Teilmenge $U \subset M$ ist die Abbildung*

$$f|_U : \begin{cases} U \to N \\ u \mapsto f(u) \end{cases}$$

Definition 6.4.13. *Es sei $f : M \to N$ eine Abbildung und $P \subseteq N$. Das Urbild von P unter der Abbildung f ist die Menge*

$$f^{-1}(P) := \{m \in M \mid f(m) \in P\} \, .$$

- Man beachte, dass bei dieser Schreibweise das Symbol f^{-1} verwendet wird, obwohl f nicht notwendig bijektiv ist, also möglicherweise gar keine Umkehrabbildung besitzt.
- Der Unterschied wird durch die Schreibweise deutlich, etwa bei $f^{-1}(a)$ und $f^{-1}(\{a\})$.
- Mit Hilfe bijektiver Abbildungen kann man jetzt noch genauere Aussagen über die Mächtigkeit von Mengen machen:

Definition 6.4.14. *Eine Menge N hat die Mächtigkeit oder Kardinalität $n \in \mathbb{N}$, i.Z. $|N| = n$, wenn es eine Bijektion*

$$f : \{1, \ldots, n\} \to N$$

gibt. Allgemeiner haben zwei beliebige Mengen M und N die gleiche Mächtigkeit, falls es eine Bijektion

$$f : M \to N$$

gibt.

7 Gruppen

7.1 Grundbegriffe

Definition 7.1.1. *Eine nichtleere Menge G zusammen mit einer Abbildung $\circ : G \times G \to G$ heißt Gruppe, wenn sie die folgenden Eigenschaften hat* [2]

(1) $\forall a, b, c \in G : \quad a \circ (b \circ c) = (a \circ b) \circ c$ *Assoziativität*

(2) $\exists e \in G\, \forall a \in G : \quad e \circ a = a$ *Neutrales Element*

(3) $\forall a \in G\, \exists a' \in G : \quad a' \circ a = e$ *Inverses Element*

Man schreibt oft $(G, \circ)$, um diesen Zusammenhang zwischen Menge und Abbildung auszudrücken. Wenn $\circ$ aus dem Zusammenhang klar ist, spricht man aber auch einfach von der Gruppe G. Wegen $\circ : G \times G \to G$ nennt man $\circ$ eine innere Verknüpfung von G.
$(G, \circ)$ heißt abelsch oder kommutativ, wenn gilt

(4) $\forall a, b \in G : \quad a \circ b = b \circ a$ *Kommutativität*

Beispiel 7.1.2. $(\mathbb{Z}, +)$ *ist eine abelsche Gruppe,* $(\mathbb{N}, +)$ *oder* $(\mathbb{N}_0, +)$ *dagegen nicht.*

Beispiel 7.1.3. $(\mathbb{Q}, +)$ *ist eine abelsche Gruppe, aber auch* $(\mathbb{Q} \setminus \{0\}, \cdot)$.
Das Gleiche gilt auch bei den anderen bekannten Körpern, wie $\mathbb{R}$ oder $\mathbb{C}$. Mehr dazu im Abschnitt über Körper.

Beispiel 7.1.4. *Im endlichen Fall kann man die Wirkung einer inneren Verknüpfung auf einer Menge mit Hilfe einer Verknüpfungstafel darstellen. So ist z.B. $(\{-1, 1\}, \cdot)$ eine endliche (multiplikative) kommutative Gruppe mit der Verknüpfungstafel*

$\cdot$	1	-1
1	1	-1
-1	-1	1

- Praktisch für das Verständnis von Verknüpfungstafeln ist der folgende kleine Satz, der hier ohne Beweis angegeben wird:

Satz 7.1.5. *Die Verknüpfungstafel einer endlichen Menge mit einer assoziativen inneren Verknüpfung ist genau dann eine Gruppentafel, wenn in jeder Zeile und jeder Spalte der Tafel jedes Element der Menge höchstens einmal vorkommt.*

Satz 7.1.6. *Es sei $(G, \circ)$ eine nicht notwendig abelsche Gruppe. Dann gilt:*

(1) Für jedes neutrale Element $e \in G$ gilt $a \circ e = a \;\; \forall a \in G$,
d.h. jedes linksneutrale Element e ist auch rechtsneutral. Deshalb spricht man auch einfach von einem neutralen Element.

(2) Aus $a' \circ a = e$ folgt jeweils auch $a \circ a' = e$,
d.h. jedes linksinverse Element a' ist auch rechtsinvers. Deshalb spricht man auch einfach von einem inversen Element.

(3) Es gibt genau ein neutrales Element $e \in G$.
Bereits aus $x \circ a = a$ oder $a \circ x = a$ für ein $a \in G$ folgt $x = e$.

(4) Zu jedem $a \in G$ gibt es genau ein inverses Element $a' \in G$.
Deshalb ist es möglich, diesem Inversen ein eigenes Symbol zu geben: in additiven Gruppen schreibt man $-a$, sonst meist a^{-1}.

- Die Beweise kann man fast in jedem Buch über Algebra oder Lineare Algebra nachlesen oder als einfache Übung selber machen.

- Hier wird nur als Beispiel für die Vorgehensweise Punkt (2) des Satzes bewiesen. Über den Gleichheitszeichen werden die Nummern der verwendeten Axiome vermerkt.

[2] Das Symbol $\circ$ ist hier beliebig gewählt und hat nicht unbedingt etwas mit der Komposition von Abbildungen zu tun

Beweis: Nach (3) gibt es zu jedem $a \in G$ ein $a' \in G$ mit $a' \circ a = e$. Zu diesem $a' \in G$ gibt es nach dem gleichen Axiom ein $a'' \in G$ mit $a'' \circ a' = e$. Damit gilt nun

$$\begin{aligned} a \circ a' &\overset{(2)}{=} e \circ (a \circ a') \overset{(3)}{=} (a'' \circ a') \circ (a \circ a') = \\ &\overset{(1)}{=} a'' \circ (a' \circ a) \circ a' \overset{(3)}{=} a'' \circ e \circ a' = a'' \circ (e \circ a') = \\ &\overset{(2)}{=} a'' \circ a' \overset{(3)}{=} e \end{aligned}$$

□

- Fasst man die ursprüngliche Definition der Gruppe und den soeben zitierten Satz zusammen, so erhält man:

Satz 7.1.7. *Eine Menge G zusammen mit einer Abbildung $\circ : G \times G \to G$ ist genau dann eine Gruppe, wenn sie die folgenden Eigenschaften hat*

(1) $\forall a, b, c \in G: \quad a \circ (b \circ c) = (a \circ b) \circ c \quad$ *Assoziativität*

(2) $\exists_1 e \in G \, \forall a \in G: \quad e \circ a = a \circ e = a \quad$ *Neutrales Element*

(3) $\forall a \in G \; \exists_1 a^{-1} \in G: \quad a^{-1} \circ a = a \circ a^{-1} = e \quad$ *Inverses Element*

26.10.10

- Die Tatsache, dass sich das neutrale Element und das Inverse jeweils so schön abelsch verhalten, darf aber nicht dazu verleiten, das für alle Element anzunehmen. Bei nicht-abelschen Gruppen gilt immer noch $a \circ b \neq b \circ a$ für die meisten $a, b \in G$.

- Für das Rechnen in Gruppen sind die folgenden Regeln nützlich:

Satz 7.1.8. *Es sei $(G, \circ)$ eine nicht notwendig abelsche Gruppe. Dann gilt:*

(1) $\forall a \in G: \quad \left(a^{-1}\right)^{-1} = a.$

(2) $\forall a, b \in G: \quad (a \circ b)^{-1} = b^{-1} \circ a^{-1}.$

(3) $\forall a, b \in G \; \exists_1 x \in G: \quad x \circ a = b.$

(4) $\forall a, b \in G \; \exists_1 y \in G: \quad a \circ y = b.$

- Die Beweise sind einfache Übungen unter Verwendung der Gruppengesetze.

- So folgt etwa die erste Behauptung aus der Eindeutigkeit des inversen Elements: $\left(a^{-1}\right)^{-1}$ ist das Inverse von a^{-1}. Von a^{-1} kennt man aber schon das Inverse a, also sind die beiden gleich.

- Die Punkte (3) und (4) ermöglichen das eindeutige Lösen von Gleichungen über Gruppen.

7.2 Die Symmetrische Gruppe

- Ist X eine Menge und $A(X)$ die Menge aller Abbildungen $\varphi : X \to X$, so ist die Komposition $\circ$ eine *innere Verknüpfung* von $A(X)$.

- Es gibt ein *neutrales Element*, nämlich id_X.

- Die Komposition $\circ$ ist *assoziativ*. Dazu ist nichts zu zeigen, denn $\circ$ ist gerade so definiert:

$$\begin{aligned} (f \circ (g \circ h))(x) =& f(g \circ h)(x) = f(g(h(x))) \\ =& (f \circ g)(h(x)) = ((f \circ g) \circ h)(x) \;\; \forall x \end{aligned}$$

- Trotzdem ist $(A(X), \circ)$ *keine Gruppe*, denn es sind ja nicht alle Abbildungen invertierbar!

- Ist dagegen X eine Menge und $S(X)$ die Menge aller bijektiven Abbildungen $\varphi : X \to X$, so ist $(S(X), \circ)$ nach 7.1.8(2) und dem soeben Gesagten eine Gruppe, denn jetzt sind ja alle Elemente invertierbar.

Definition 7.2.1. *Es sei X eine Menge und $S(X)$ die Menge aller bijektiven Abbildungen $\varphi : X \to X$. $(S(X), \circ)$ heißt die Gruppe der Permutationen von X oder die symmetrische Gruppe von X.*
Im wichtigen Spezialfall $X = \{1, 2, \ldots, n\}$ schreibt man statt $S(\{1, 2, \ldots, n\})$ kürzer S_n.

- Statt der üblichen Schreibweise für Relationen $\{(1,\varphi(1)),\dots,(n,\varphi(n))\}$ hat sich im Fall von Elementen der S_n die folgende Schreibweise eingebürgert:

$$\varphi = \begin{pmatrix} 1 & 2 & \dots & n \\ \varphi(1) & \varphi(2) & \dots & \varphi(n) \end{pmatrix}$$

Beispiel 7.2.2. *Mit Hilfe dieser Schreibweise kann man z.B. die Komposition von Abbildungen schnell berechnen oder die Inverse sofort angeben:*

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

- Für endliches X ist $|S(X)| = |X|!$ (mit $n! := \prod_{k=1}^{n} k$, lies: „*n Fakultät*“; Beweis mit Induktion.)

Beispiel 7.2.3. *In S_3 löst man z.B.*

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \circ y = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$

indem man beidseitig von links (die Gruppe ist nicht abelsch!) mit dem Inversen $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{smallmatrix}\right)$ von $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right)$ (s. 7.2.2) „multipliziert“:

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \circ y = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \iff y = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

7.3 Gruppen-Homomorphismen

Definition 7.3.1. *Eine Abbildung $f : G_1 \to G_2$ von der Gruppe $(G_1, \circ)$ in die Gruppe $(G_2, \bullet)$ heißt Homomorphismus, wenn gilt*

$$\forall a, b \in G_1 \; : \quad f(a \circ b) = f(a) \bullet f(b).$$

Ein surjektiver Homomorphismus heißt Epimorphismus, ein injektiver Monomorphismus und ein bijektiver Isomorphismus. Ist $G_1 = G_2$, so nennt man einen bijektiven Homomorphismus auch Automorphismus.

Beispiel 7.3.2. *Es ist $f_n(x) := n \cdot x$ für jedes $n \in \mathbb{N}$ ein Monomorphismus von $(\mathbb{Z}, +)$ in sich, denn für beliebige $a, b \in \mathbb{Z}$ gilt*

$$f_n(a+b) = n \cdot (a+b) = n \cdot a + n \cdot b = f_n(a) + f_n(b)$$

- Homomorphismen, die wir hier erstmals für Gruppen kennengelernt haben, sind ein wesentliches mathematisches Hilfsmittel.
- Ein Homomorphismus ist generell eine strukturerhaltende Abbildung.
- Neben Gruppen-Homomorphismen werden wir in diesem Kurs vor allem mit Vektorraum-Homomorphismen zu tun haben.

Definition 7.3.3. *Eine Permutation aus S_n heißt Transposition, wenn sie zwei der Zahlen $1, \dots, n$ vertauscht, die anderen aber elementweise festhält.*

Beispiel 7.3.4. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 5 & 4 & 3 \end{pmatrix} \in S_5$ *ist eine Transposition.*

Satz 7.3.5. *Jede Permutation aus S_n lässt sich als Produkt (=Komposition) von Transpositionen schreiben. Diese Transpositionen und ihre Anzahl m sind nicht eindeutig, wohl aber die Zahl $(-1)^m$.*

- Statt eines Beweises wird hier nur ein Beispiel gegeben:

Beispiel 7.3.6.

$$\begin{pmatrix}1&2&3&4\\3&2&4&1\end{pmatrix} = \begin{pmatrix}1&2&3&4\\4&2&3&1\end{pmatrix} \circ \begin{pmatrix}1&2&3&4\\3&2&1&4\end{pmatrix}$$
$$= \begin{pmatrix}1&2&3&4\\1&3&2&4\end{pmatrix} \circ \begin{pmatrix}1&2&3&4\\1&2&4&3\end{pmatrix} \circ \begin{pmatrix}1&2&3&4\\2&1&3&4\end{pmatrix} \circ \begin{pmatrix}1&2&3&4\\1&4&3&2\end{pmatrix}$$

d.h. es handelt sich um ein so genannte gerade Permutation mit $m = 2, 4, 6, \ldots$, *also* $(-1)^m = 1$.

28.10.10

Definition 7.3.7. *Die folgende Abbildung heißt Signum:*

$$\mathrm{sgn} : \begin{cases} S_n \to \{-1, 1\} \\ \varphi \mapsto \prod_{i<j} \frac{\varphi(j)-\varphi(i)}{j-i} \end{cases}$$

- Da φ eine bijektive Abbildung ist, stehen nicht nur im Nenner von $\mathrm{sgn}(\varphi)$ die Differenzen aller verschiedenen Zahlen von 1 bis n (jeweils größere minus kleinere Zahl), sondern im Zähler stehen genau die gleichen Differenzen (diese allerdings durch Vertauschungen manchmal auch mit negativem Ergebnis). Deshalb ist klar, dass sich der auf den ersten Blick kompliziert aussehende Bruch für $\mathrm{sgn}(\varphi)$ immer vollständig zu 1 oder -1 kürzen lässt.

Definition 7.3.8. *Es seien* $i, j, n \in \mathbb{N}$ *mit* $1 \leq i, j \leq n$, $\varphi \in S_n$. *Ein Tupel* (i, j) *mit* $j > i$ *und* $\varphi(j) < \varphi(i)$ *heißt Fehlstand der Permutation* φ.

- Man kann sich die Berechnung von $\mathrm{sgn}(\varphi)$ deutlich vereinfachen, wenn man einfach die Fehlstände von φ zählt, denn genau die führen jeweils zu einer Vorzeichenänderung bei $\mathrm{sgn}(\varphi)$:

Folgerung 7.3.9.

$$\mathrm{sgn}(\varphi) = (-1)^{\#\ \textit{Fehlstände von } \varphi}\,.$$

Beispiel 7.3.10. *Für die Permutation*

$$\sigma = \begin{pmatrix}1&2&3&4&5\\4&1&5&2&3\end{pmatrix}$$

ergibt sich durch Einsetzen in die Formel

$$\mathrm{sgn}(\sigma) = \frac{(3-4)(3-1)(3-5)(3-2)(2-4)(2-1)(2-5)(5-4)(5-1)(1-4)}{(5-1)(5-2)(5-3)(5-4)(4-1)(4-2)(4-3)(3-1)(3-2)(2-1)} = -1$$

Die Fehlstände dieser Permutation sind $(1,5)$, $(3,5)$, $(1,4)$, $(3,4)$, $(1,2)$.

Da die Anzahl dieser Fehlstände ungerade ist, sieht man auch so: $\mathrm{sgn}(\sigma) = -1$.

Satz 7.3.11. *Die Abbildung* $\mathrm{sgn} : S_n \to \{-1, 1\}$ *ist ein Gruppenhomomorphismus (für* $n \geq 2$ *sogar surjektiv, also Epimorphismus) von* $(S_n, \circ)$ *auf* $(\{-1, 1\}, \cdot)$, *d.h. es gilt*

$$\forall \varphi, \psi \in S_n \; : \; \mathrm{sgn}(\varphi \circ \psi) = \mathrm{sgn}(\varphi) \cdot \mathrm{sgn}(\psi)\,.$$

Beweis:

$$\begin{aligned}\mathrm{sgn}(\varphi \circ \psi) &= \prod_{i<j} \frac{(\varphi\circ\psi)(j) - (\varphi\circ\psi)(i)}{j-i} = \\ &= \prod_{i<j} \frac{(\varphi\circ\psi)(j) - (\varphi\circ\psi)(i)}{\psi(j)-\psi(i)} \cdot \frac{\psi(j)-\psi(i)}{j-i} = \\ &= \prod_{i<j} \frac{\varphi(\psi(j)) - \varphi(\psi(i))}{\psi(j)-\psi(i)} \cdot \prod_{i<j} \frac{\psi(j)-\psi(i)}{j-i} = \mathrm{sgn}(\varphi)\cdot\mathrm{sgn}(\psi)\end{aligned}$$

□

7.4 Untergruppen

Definition 7.4.1. *Eine nichtleere Teilmenge $U \subseteq G$ einer Gruppe $(G, \circ)$, die auch mit der Verknüpfung $\circ$ eine Gruppe bildet, heißt Untergruppe von G.*

Satz 7.4.2. *Eine nichtleere Teilmenge $U \subseteq G$ ist genau dann eine Untergruppe von $(G, \circ)$, wenn gilt $u \circ v^{-1} \in U \quad \forall u, v \in U$.*

Beweis: Wenn allgemein $u \circ v^{-1} \in U$ ist, so folgt auch $u \circ u^{-1} = e \in U$.
Setzt man in $u \circ v^{-1} \in U$ für u dieses neutrale Element e ein, so folgt $v^{-1} \in U$ für jedes $v \in U$.
Für $w \in U$ ist somit auch $w^{-1} \in U$ und deshalb auch $u \circ (w^{-1})^{-1} = u \circ w \in U$, d.h. $\circ$ ist eine innere Verknüpfung von U.
Diese ist natürlich assoziativ, da sie auf ganz G assoziativ ist. □

- Wahlweise kann man für diesen Untergruppentest auch zeigen:
  $u, v \in U \Rightarrow u \circ v \in U$ und $u^{-1} \in U$.

Beispiel 7.4.3.

$$\left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}$$

ist eine Untergruppe der S_3, denn $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ und $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ sind invers zueinander und $e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$ ist das neutrale Element. Somit klappt der Untergruppentest.

Definition 7.4.4. *Es sei $f : G_1 \to G_2$ ein Homomorphismus von der Gruppe $(G_1, \circ)$ in die Gruppe $(G_2, \bullet)$. Das Urbild des neutralen Elements $e_2 \in G_2$ heißt der Kern des Homomorphismus*

$$\mathrm{Kern}(f) := f^{-1}(\{e_2\})$$

Satz 7.4.5. *Der Kern eines Homomorphismus $f : G_1 \to G_2$ ist Untergruppe von G_1.*

Beweis: Es sei $e_1 \in G_1$ das neutrale Element von G_1. Wegen

$$f(e_1) = f(e_1 \circ e_1) = f(e_1) \bullet f(e_1) \Rightarrow f(e_1) = e_2$$

folgt $e_1 \in f^{-1}(\{e_2\})$. [3] Der Kern von f ist also sicher nicht leer.
Sind nun $a, b \in \mathrm{Kern}(f)$, also $f(a) = e_2 = f(b)$, so folgt

$$f(a \circ b) = f(a) \bullet f(b) = e_2 \bullet e_2 = e_2 \Rightarrow a \circ b \in \mathrm{Kern}(f)$$

Weiterhin folgt

$$e_2 = f(e_1) = f(a \circ a^{-1}) = f(a) \bullet f(a^{-1}) = e_2 \bullet f(a^{-1}) = f(a^{-1})$$

und somit gilt auch $a^{-1} \in \mathrm{Kern}(f)$. Der Untergruppentest zeigt also, dass es sich um eine Untergruppe handelt. □

[3] Hier sieht man schon einen ersten Hinweis auf die strukturerhaltende Eigenschaft eines Homomorphismus. Das neutrale Element wird immer auf das neutrale Element abgebildet!

8 Ringe und Körper

8.1 Grundbegriffe

Definition 8.1.1. *Eine Menge R mit zwei inneren Verknüpfungen*

$$+ : R \times R \to R \,, \; \cdot : R \times R \to R$$

heißt Ring, falls gilt

(1) $(R, +)$ *ist eine abelsche Gruppe*

(2) $\forall a, b, c \in R: \quad a \cdot (b \cdot c) = (a \cdot b) \cdot c \qquad$ *Assoziativität bzgl.* $\cdot$

(3) $\forall a, b, c \in R: \quad a \cdot (b + c) = a \cdot b + a \cdot c \;$ *und* $\; (a + b) \cdot c = a \cdot c + b \cdot c \qquad$ *Distributivität; mit „Punkt" vor „Strich"*

(4) *Wenn zusätzlich gilt*
$\exists 1 \in R \;\; \forall a \in R: \quad 1 \cdot a = a$ *heißt R Ring mit Einselement* [4]

(5) *Wenn zusätzlich gilt*
$\forall a, b \in R: \quad a \cdot b = b \cdot a$ *heißt R kommutativer Ring* [5]

Beispiel 8.1.2. $(\mathbb{Z}, +, \cdot)$ *ist ein kommutativer Ring mit* 1.

Beispiel 8.1.3. $(n\mathbb{Z}, +, \cdot)$ *mit* $n \in \mathbb{N}_{>1}$ *ist ein kommutativer Ring ohne* 1.

Beispiel 8.1.4. $(\mathbb{Q}, +, \cdot)$ *ist ein kommutativer Ring mit* 1.
Das Gleiche gilt auch bei den anderen bekannten Körpern, wie $\mathbb{R}$ oder $\mathbb{C}$. Mehr dazu im Abschnitt über Körper.

Satz 8.1.5. *Es sei R ein Ring. Für alle $a, b \in R$ gilt*

(1) $0 \cdot a = a \cdot 0 = 0$

(2) $(-a) \cdot b = a \cdot (-b) = -(a \cdot b)$

Beweis: (1) $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a \Rightarrow 0 \cdot a = 0$
$a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0 \Rightarrow a \cdot 0 = 0$

(2) $a \cdot 0 = a \cdot (b - b) = a \cdot (b + (-b)) = a \cdot b + a \cdot (-b) \Rightarrow -(a \cdot b) = a \cdot (-b)$ und analog für $(-a) \cdot b$.

□

Definition 8.1.6. *Ein kommutativer Ring K mit $1 \neq 0$ heißt Körper, wenn gilt:*

$$\forall k \in K \setminus \{0\} \;\; \exists \ell \in K: \quad k \cdot \ell = 1$$

- Zusammen mit den Axiomen, die ein kommutativer Ring mit 1 schon erfüllt, heißt das, dass nicht nur $(K, +)$ eine abelsche Gruppe ist, sondern auch $(K \setminus \{0\}, \cdot)$. Somit ist auch das zu jedem $k \in K \setminus \{0\}$ geforderte multiplikative Inverse eindeutig.

Beispiel 8.1.7. $(\mathbb{Z}, +, \cdot)$ *ist kein Körper, denn nur* 1 *und* -1 *sind multiplikativ invertierbar.*

Beispiel 8.1.8. $(\mathbb{Q}, +, \cdot)$ *und* $(\mathbb{R}, +, \cdot)$ *sind Körper.*

- An der Stelle sollte man sich klarmachen, dass ein Körper genau das bietet, was man für den Gauß-Algorithmus benötigt:
  - Mit den beiden Verknüpfungen „+" und „·" sind elementare Umformungen vom Typ (ii) und (iii) erst möglich.
  - Beidseitige Multiplikation einer Gleichung mit einer Zahl $\alpha \in K \setminus \{0\}$ ist eine Äquivalenzumformung, denn dieses α ist invertierbar.
  - Insbesondere kann man eine Zeile mit dem Inversen des Pivotelements multiplizieren (bzw. durch dieses Element teilen), also ein Pivotelement 1 herstellen.

[4] bzgl. $+$ gibt es auf jeden Fall ein neutrales Element, das mit 0 bezeichnet wird
[5] die Addition ist in jedem Fall kommutativ!

8.2 Die komplexen Zahlen

Satz 8.2.1. *Der $\mathbb{R}^2$ mit den beiden folgenden Verknüpfungen ist ein Körper:*

$$+ \quad : \quad \begin{cases} \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2 \\ ((a_1, a_2), (b_1, b_2)) \mapsto (a_1 + b_1, a_2 + b_2) \end{cases}$$
$$\cdot \quad : \quad \begin{cases} \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2 \\ ((a_1, a_2), (b_1, b_2)) \mapsto (a_1 b_1 - a_2 b_2, a_1 b_2 + a_2 b_1) \end{cases}$$

Beweis: Der $\mathbb{R}^2$ ist offensichtlich nicht leer und beide Verknüpfungen sind innere Verknüpfungen weil $\mathbb{R}$ ein Körper ist.
Die Addition geschieht einfach elementweise und ist deshalb klarerweise auch assoziativ. Das neutrale Element der Addition ist $(0,0)$, zu jedem $(a,b) \in \mathbb{R}^2$ gibt es ein inverses Element $-(a,b) = (-a,-b)$
Auch die etwas ungewohnte Multiplikation ist assoziativ, was man mit etwas Schreibarbeit nachweisen kann:

$$\begin{aligned} &(a_1, a_2) \cdot ((b_1, b_2) \cdot (c_1, c_2)) = \\ &((a_1 b_1 c_1 - a_1 b_2 c_2 - a_2 b_1 c_2 - a_2 b_2 c_1, a_1 b_1 c_2 + a_1 b_2 c_1 + a_2 b_1 c_1 - a_2 b_2 c_2)) = \qquad \forall (a_1, a_2), (b_1, b_2), (c_1, c_2) \in \mathbb{R}^2 \\ &((a_1, a_2) \cdot (b_1, b_2)) \cdot (c_1, c_2) \end{aligned}$$

Wegen

$$\begin{aligned} (a_1, a_2) \cdot (b_1, b_2) &= (a_1 b_1 - a_2 b_2, a_1 b_2 + a_2 b_1) \\ (b_1, b_2) \cdot (a_1, a_2) &= (b_1 a_1 - b_2 a_2, b_1 a_2 + b_2 a_1) \end{aligned} \qquad \forall (a_1, a_2), (b_1, b_2) \in \mathbb{R}^2$$

ist die Multiplikation auch kommutativ.
Wegen

$$(a_1, a_2) \cdot (1, 0) = (a_1, a_2) \quad \forall (a_1, a_2) \in \mathbb{R}^2$$

ist $(1,0)$ neutrales Element der Multiplikation.
Wegen

$$(a_1, a_2) \cdot \left(\frac{a_1}{a_1^2 + a_2^2}, \frac{-a_2}{a_1^2 + a_2^2} \right) = (1, 0) \quad \forall (a_1, a_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$$

ist jedes von 0 verschiedene Element $(a_1, a_2) \in \mathbb{R}^2$ invertierbar mit dem inversen Element

$$(a_1, a_2)^{-1} = \left(\frac{a_1}{a_1^2 + a_2^2}, \frac{-a_2}{a_1^2 + a_2^2} \right)$$

Nachdem man auch das Distributivgesetz nachrechnen kann

$$\begin{aligned} &(a_1, a_2) \cdot ((b_1, b_2) + (c_1, c_2)) = \\ &(a_1 \cdot b_1 + a_1 \cdot c_1 - a_2 \cdot b_2 - a_2 \cdot c_2, a_1 \cdot b_2 + a_1 \cdot c_2 + a_2 \cdot b_1 + a_2 \cdot c_1) = \qquad \forall (a_1, a_2), (b_1, b_2), (c_1, c_2) \in \mathbb{R}^2 \\ &(a_1, a_2) \cdot (b_1, b_2) + (a_1, a_2) \cdot (c_1, c_2), \end{aligned}$$

sind somit alle Körpergesetze nachgeweisen. □

- Der im letzten Satz eingeführte Körper hat einige wichtige Eigenschaften und wird deshalb noch etwas genauer untersucht.
- Wegen $\begin{aligned} (a_1, 0) &+ (b_1, 0) = (a_1 + b_1, 0) \\ (a_1, 0) &\cdot (b_1, 0) = (a_1 b_1, 0) \end{aligned}$
  sieht man, dass es sich mit den Elementen $(a, 0)$ mit $a \in \mathbb{R}$ genauso rechnet, wie mit reellen Zahlen.
- Man schreibt deshalb statt $(a, 0)$ einfach a und rechnet wie mit einer reellen Zahl üblich.
- Insbesondere ist bei dieser Schreibweise 0 das neutrale Element der Addition und 1 das neutrale Element der Multiplikation.
- Für das Element $(0,1) \in \mathbb{R}^2$ gilt $(0,1) \cdot (0,1) = (-1,0) = -1$, d.h. $(0,1)$ ist eine Lösung der (in $\mathbb{R}$ nicht lösbaren) Gleichung $x^2 = -1$.
- Dieses Körperelement bezeichnet man meist mit dem Buchstaben i (E-Techniker schreiben j, weil i schon der Strom ist) und nennt es die *imaginäre Einheit*.

- Damit schreibt man statt des Tupels $(a,b) \in \mathbb{R}^2$ auch $a + i \cdot b$ und obige Rechenregeln sehen in dieser Schreibweise so aus:
- $$\begin{aligned} (a_1 + i \cdot a_2) + (b_1 + i \cdot b_2) &= (a_1 + b_1) + i \cdot (a_2 + b_2) \\ (a_1 + i \cdot a_2) \cdot (b_1 + i \cdot b_2) &= (a_1 b_1 - a_2 b_2) + i \cdot (a_1 b_2 + a_2 b_1)\,, \end{aligned}$$
- d.h. man behandelt i wie eine Unbestimmte, nutzt aber die Eigenschaft $i^2 = -1$ aus, wenn mal eine größere Potenz von i benötigt wird.

Definition 8.2.2.

$$\mathbb{C} := \{a + ib \mid a, b \in \mathbb{R}\,,\ i^2 = -1\}$$

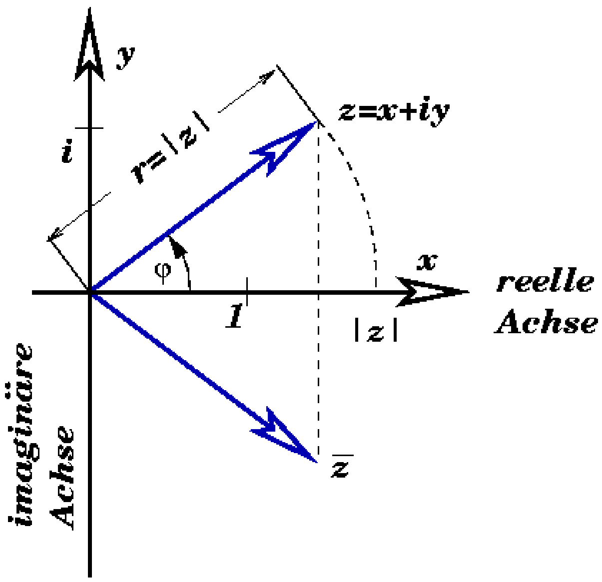
heißt Körper der komplexen Zahlen. Für $z = a+ib \in \mathbb{C}$ heißt $\mathrm{Re}(z) := a$ der Realteil von z und $\mathrm{Im}(z) := b$ der Imaginärteil von z. Skizziert man komplexe Zahlen als Vektoren bzw. Punkte (a,b) in einem kartesischen Koordinatensystem, so nennt man $\mathbb{R}^2$ Gaußsche Zahlenebene. Die Länge eines Vektors in der Gaußschen Zahlenebene heißt Betrag der komplexen Zahl, i.Z.

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

Mit Polarkoordinaten (r, φ) von (a,b) in der Zahlenebene gilt

$$z = a + ib = r(\cos(\varphi) + i \sin(\varphi)) = r\cos(\varphi) + ir\sin(\varphi) \text{ mit } r := |z|\ .$$

Das Argument von z, $\arg(z) := \varphi$, ist der von der positiven reellen Achse entgegen dem Uhrzeigersinn gemessene Winkel zu $z = (a,b)$ $(0 \le \varphi < 2\pi!)$.



Satz 8.2.3. *Die Abbildung*

$$k : \begin{cases} \mathbb{C} \to \mathbb{C} \\ a + ib \mapsto a - ib \end{cases}$$

ist ein Körperautomorphismus, also bijektiv und es gilt für alle $z_1, z_2 \in \mathbb{C}$:

$$k(z_1 + z_2) = k(z_1) + k(z_2) \quad , \quad k(z_1 \cdot z_2) = k(z_1) \cdot k(z_2)$$

Beweis: Einfache Übung. □

Definition 8.2.4. *$k(a + ib) = a - ib$ heißt das konjugiert Komplexe von $a + ib$. Man schreibt üblicherweise $\overline{a + ib} = a - ib$ statt $k(a + ib)$.*

- In der Gaußschen Zahlenebene entspricht die Konjugation einer Spiegelung an der reellen Achse.
- Obige Rechenregeln lauten umgeschrieben und noch erweitert auf die Inversen:

Folgerung 8.2.5.

$$\begin{array}{rclcrclcrcl}
\overline{z_1+z_2} & = & \overline{z_1}+\overline{z_2} & , & \overline{z_1-z_2} & = & \overline{z_1}-\overline{z_2} & , & \overline{\overline{z}} & = & z \\
\overline{z_1\cdot z_2} & = & \overline{z_1}\cdot\overline{z_2} & , & \overline{\left(\frac{z_1}{z_2}\right)} & = & \frac{\overline{z_1}}{\overline{z_2}} & , & z\cdot\overline{z} & = & |z|^2
\end{array}$$

- Besonders nützlich für das Rechnen mit Potenzen komplexer Zahlen ist der folgende Satz, der hier ohne Beweis (kommt sicher irgendwann in der Analysis) mitgeteilt wird:

Satz 8.2.6. *Formel von de Moivre:*

$$\forall n\in\mathbb{N}:\quad (\cos(\varphi)+i\sin(\varphi))^n=\cos(n\varphi)+i\sin(n\varphi)$$

2.11.10

- Interessant ist die geometrische Deutung der Multiplikation:
- Sind z und z' komplexe Zahlen mit den Polarformen

$$\begin{aligned} z &= r\cdot(\cos(\varphi)+i\sin(\varphi)) \\ z' &= r'\cdot(\cos(\varphi')+i\sin(\varphi'))\,, \end{aligned}$$

- so folgt mit Hilfe der Additionstheoreme

$$\begin{aligned} \sin(\varphi+\varphi') &= \cos(\varphi)\sin(\varphi')+\sin(\varphi)\cos(\varphi') \\ \cos(\varphi+\varphi') &= \cos(\varphi)\cos(\varphi')-\sin(\varphi)\sin(\varphi')\,, \end{aligned}$$

- und damit $z\cdot z'=r\cdot r'(\cos(\varphi+\varphi')+i\cdot\sin(\varphi+\varphi'))$
- d.h. eine Multiplikation mit z bewirkt in der Gaußschen Zahlenebene eine *Drehstreckung* (Länge mal r, Winkel plus φ, siehe auch das MatheVital-Applet dazu).

Satz 8.2.7. *Für alle $z,w\in\mathbb{C}$ gilt:*

- $\mathrm{Re}(z)=\frac{z+\overline{z}}{2}$
- $\mathrm{Im}(z)=\frac{z-\overline{z}}{2i}$
- $|w\cdot z|=|w|\cdot|z|$
- $|w+z|\leq|w|+|z|$ *Dreiecksungleichung*

Beweis: Die ersten 3 Punkte sind eine einfache Übung, der letzte Punkt wird bei der Besprechung der Norm später allgemeiner gezeigt. □

Komm. Gr. +	Ass.	$(a+b)+c=a+(b+c)$	**Ring**	**kommutativer Ring**	**kommutativer Ring mit 1**	**Körper**
	Komm.	$a+b=b+a$				
	Neutr.	$\exists\, 0: 0+a=a+0=a$				
	Inv.	$\exists -a: a+(-a)=0$				
	Distr.	$(a+b)\cdot c=a\cdot c+b\cdot c$				
Komm. Gr. ·	Ass.	$(a\cdot b)\cdot c=a\cdot(b\cdot c)$				
	Komm.	$a\cdot b=b\cdot a$				
	Neutr.	$\exists 1: 1\cdot a=a\cdot 1=a$				
	Inv.	$\exists\, a^{-1}: a\cdot a^{-1}=1$				

Abbildung 20: Algebraische Strukturen

9 Matrizenrechnung

9.1 Gleichheit, Addition, Vielfache, Transposition

Definition 9.1.1. *Es sei K ein Körper. Zwei Matrizen $A = (a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} \in K^{m\times n}$ und $B = (b_{ij})_{\substack{1\le i\le p\\1\le j\le q}} \in K^{p\times q}$ heißen gleich, i.Z. $A = B$, wenn gilt $m = p$, $n = q$ und $a_{ij} = b_{ij}$ für $1 \le i \le m$ und $1 \le j \le n$.*

Definition 9.1.2. *Es sei K ein Körper. Für zwei Matrizen*

$$A = (a_{ij})_{\substack{1\le i\le m\\1\le j\le n}}\,,\ B = (b_{ij})_{\substack{1\le i\le m\\1\le j\le n}} \in K^{m\times n}$$

sei deren Summe $A + B := (a_{ij} + b_{ij})_{\substack{1\le i\le m\\1\le j\le n}} \in K^{m\times n}$.

Definition 9.1.3. *Es sei K ein Körper. Für eine Matrix $A = (a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} \in K^{m\times n}$ und einen Skalar $\lambda \in K$ sei*

$$\lambda A = \lambda(a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} := (\lambda a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} \in K^{m\times n}\,.$$

Satz 9.1.4. *Es sei K ein Körper. $(K^{m\times n}, +)$ ist eine abelsche Gruppe.*

Beweis: Die eingeführte Matrizen-Addition ist einfach die elementweise Addition in K, also sicher eine innere, assoziative und abelsche Verknüpfung. Die nur mit dem Nullelement des Körpers gefüllte *Nullmatrix* $\mathbf{0} := (0)_{\substack{1\le i\le m\\1\le j\le n}}$ ist das Nullelement von $(K^{m\times n}, +)$ und zu jedem $A = (a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} \in K^{m\times n}$ gibt es das *inverse Element* $-A = (-a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} \in K^{m\times n}$. □

Definition 9.1.5. *Es sei K ein Körper. Für eine Matrix $A = (a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} \in K^{m\times n}$ ist*

$${}^tA := (a'_{ij})_{\substack{1\le i\le n\\1\le j\le m}} \in K^{n\times m}$$

mit $a'_{ij} = a_{ji}$ für $1 \le i \le n$ und $1 \le j \le m$ die so genannte Transponierte von A.
Oft findet man auch die Schreibweise A^T statt ${}^tA$.
Ausgeschrieben sieht das so aus

$${}^tA = {}^t\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ & & \vdots & \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}$$

Satz 9.1.6. *Es sei K ein Körper. Für $A, B \in K^{m\times n}$ und $\lambda \in K$ gilt*

(1) ${}^t(A + B) = {}^tA + {}^tB$,

(2) ${}^t(\lambda \cdot A) = \lambda \cdot {}^tA$,

(3) ${}^t({}^tA) = A$.

Beweis: Einfache Übung □

Definition 9.1.7. *Es sei K ein Körper. Ein $A \in K^{n\times n}$ mit ${}^tA = A$ heißt symmetrisch.*

9.2 Einige besondere Matrizen

- Die folgenden speziellen Matrizen werden immer wieder gebraucht und teilweise mit eigenen Namen versehen. Dabei sei jeweils K ein Körper.

- Die *Einheitsmatrix*

$$\mathbb{1}_n := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & & & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \in K^{n\times n}$$

lässt sich auch mit Hilfe des *Kronecker-Symbols* schreiben

$$\delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} : \quad \mathbb{1}_n = (\delta_{ij})_{\substack{1\le i\le n \\ 1\le j\le n}}$$

- Die *Einsermatrizen*

$$E_{ij} := \begin{array}{cc} & \begin{array}{ccccc} & & j & & \\ & & \downarrow & & \end{array} \\ i\rightarrow & \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ \vdots & & 1 & & \vdots \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \end{array} \in K^{m\times n}$$

- *Obere Dreiecksmatrizen*

$$\begin{pmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & * \end{pmatrix} \in K^{n\times n}$$

wobei $*$ für einen beliebigen Skalar aus K steht. Formal kann man schreiben $(a_{ij})_{\substack{1\le i\le n \\ 1\le j\le n}} \in K^{n\times n}$ mit $a_{ij} = 0$ für $i > j$.

- *Echte obere Dreiecksmatrizen*

$$\begin{pmatrix} 0 & * & * & \dots & * \\ 0 & 0 & * & \dots & * \\ 0 & 0 & 0 & & * \\ \vdots & & & \ddots & * \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \in K^{n\times n}$$

wobei $*$ für einen beliebigen Skalar aus K steht, oder $(a_{ij})_{\substack{1\le i\le n \\ 1\le j\le n}} \in K^{n\times n}$ mit $a_{ij} = 0$ für $i \geq j$.

- Die *Nullmatrix*

$$\mathbf{O} := \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \in K^{m\times n}$$

- *Untere Dreiecksmatrizen*

$$\begin{pmatrix} * & 0 & 0 & \dots & 0 \\ * & * & 0 & \dots & 0 \\ * & * & * & & 0 \\ \vdots & & & \ddots & 0 \\ * & * & * & \dots & * \end{pmatrix} \in K^{n\times n}$$

wobei $*$ für einen beliebigen Skalar aus K steht, oder $(a_{ij})_{\substack{1\le i\le n \\ 1\le j\le n}} \in K^{n\times n}$ mit $a_{ij} = 0$ für $i < j$.

- *Echte untere Dreiecksmatrizen*
$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ * & 0 & 0 & \dots & 0 \\ * & * & 0 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ * & * & * & \dots & 0 \end{pmatrix} \in K^{n\times n}$$
wobei $*$ für einen beliebigen Skalar aus K steht, oder $(a_{ij})_{\substack{1\le i\le n\\ 1\le j\le n}} \in K^{n\times n}$ mit $a_{ij}=0$ für $i\le j$.

9.3 Multiplikation

Definition 9.3.1. *Es sei K ein Körper. Für $A=(a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} \in K^{m\times n}$, $B=(b_{jk})_{\substack{1\le j\le n\\ 1\le k\le p}} \in K^{n\times p}$ sei*
$$A\cdot B := C = (c_{ik})_{\substack{1\le i\le m\\ 1\le k\le p}} \in K^{m\times p} \text{ mit } c_{ik} := \sum_{j=1}^{n} a_{ij}b_{jk}\,.$$

- Es wird also nach dem Schema „Zeile · Spalte“ gerechnet:
-
$$\begin{array}{cc} & \begin{pmatrix} b_{11} & \dots & b_{1k} & \dots & b_{1p} \\ b_{21} & \dots & b_{2k} & \dots & b_{2p} \\ \vdots & & \dots & & \vdots \\ b_{n1} & \dots & b_{nk} & \dots & b_{np} \end{pmatrix} \\ \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & & \vdots & \\ a_{i1} & a_{i2} & \dots & a_{in} \\ & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} & \begin{pmatrix} & & & & \\ & & \downarrow & & \\ & \rightarrow & c_{ik} & & \\ & & & & \\ & & & & \end{pmatrix} \end{array}$$

- Insbesondere gilt für eine $n\times 1$-Matrix $x = {}^t\begin{pmatrix} x_1 & \dots & x_n\end{pmatrix}$ (bzw. einen Spaltenvektor aus K^n)
$$\begin{aligned} A\cdot x &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{pmatrix} \end{aligned}$$

- Damit ist nachträglich auch die Schreibweise $Ax=b$ für lineare Gleichungssysteme klar.

- Ist $x = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$ eine $1\times n$-Matrix (=Zeilenvektor), so gilt
$$\begin{aligned} x\cdot B &= \begin{pmatrix} x_1b_{11}+\dots+x_nb_{n1} & ,\dots, & x_1b_{1p}+\dots+x_nb_{np} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^{n} x_jb_{j1} & ,\dots, & \sum_{j=1}^{n} x_jb_{jp} \end{pmatrix} \end{aligned}$$

- Es sei speziell $x=e_i=(\delta_{ij})_{1\le j\le n}$ (*Kronecker-Symbol*) einer der so genannten *kanonischen Einheitsvektoren*, d.h. ein Vektor mit 0 in allen Positionen und nur einer 1 an der i-ten Stelle.

- $A\cdot e_i$ liefert dann gerade die i-te Spalte von A.

- Entsprechend liefert ${}^t(e_i) \cdot B$ die i-te Zeile von B.
- Sind $z_1, \ldots, z_m \in K^{1\times n}$ die Zeilen von A, so kann man damit schreiben

$$A \cdot B = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix} \cdot B = \begin{pmatrix} z_1 \cdot B \\ \vdots \\ z_m \cdot B \end{pmatrix} .$$

- Sind $s_1, \ldots, s_p \in K^{n\times 1}$ die Spalten von B, so kann man damit schreiben

$$A \cdot B = A \cdot \begin{pmatrix} s_1 & \ldots & s_p \end{pmatrix} = \begin{pmatrix} A \cdot s_1 & \ldots & A \cdot s_p \end{pmatrix} .$$

Satz 9.3.2. *Für $A, A_1, A_2 \in K^{m\times n}$, $\alpha \in K$, $B, B_1, B_2 \in K^{n\times p}$ und $C \in K^{p\times q}$ gilt:*

(1) $(A_1 + A_2) \cdot B = A_1 \cdot B + A_2 \cdot B$ *(1. Distributivgesetz),*

(2) $A \cdot (B_1 + B_2) = A \cdot B_1 + A \cdot B_2$ *(2. Distributivgesetz),*

(3) $\alpha \cdot (A \cdot B) = (\alpha \cdot A) \cdot B = A \cdot (\alpha \cdot B)$,

(4) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ *(Assoziativgesetz),*

(5) ${}^t(A \cdot B) = {}^tB \cdot {}^tA$.

Beweis: All diese Gesetze werden durch Einsetzen in die Definition der Matrizenmultiplikation und Verwendung der entsprechenden Rechengesetze im zugrunde liegenden Körper K bewiesen. Hier wird nur das 1. Distributivgesetz (1) als Beispiel gezeigt. Dazu seien $A_1 = (a_{ij}^{(1)})$, $A_2 = (a_{ij}^{(2)})$ und $B = (b_{ij})$:

$$\begin{aligned} (A_1 + A_2) \cdot B = (a_{ij}^{(1)} + a_{ij}^{(2)}) \cdot (b_{ij}) &= \left(\sum_{j=1}^n (a_{ij}^{(1)} + a_{ij}^{(2)}) b_{jk} \right) = \\ &= \left(\sum_{j=1}^n \left(a_{ij}^{(1)} b_{jk} + a_{ij}^{(2)} b_{jk} \right) \right) = \\ &= \left(\sum_{j=1}^n a_{ij}^{(1)} b_{jk} \right) + \left(\sum_{j=1}^n a_{ij}^{(2)} b_{jk} \right) = \\ &= (a_{ij}^{(1)}) \cdot (b_{ij}) + (a_{ij}^{(2)}) \cdot (b_{ij}) = \\ &= A_1 \cdot B + A_2 \cdot B \end{aligned}$$

□

- Speziell für $K^{n\times n}$ sind $+$ und $\cdot$ also innere, assoziative Verknüpfungen.
- Die Nullmatrix $\mathbf{0}$ und die Einheitsmatrix $\mathbb{1}_n$ sind neutrale Elemente der Addition und der Multiplikation.
- Nach dem letzten Satz gelten die Distributivgesetze.
- $$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} ,$$

  d.h. die Multiplikation ist *nicht kommutativ.*
- $$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

  d.h. das Matrizenprodukt ist *nicht nullteilerfrei.*

Satz 9.3.3. $(K^{n\times n}, +, \cdot)$ *ist ein nicht-kommutativer Ring mit Einselement.*

9.4 Inverse Matrizen

- Die Komposition von Abbildungen einer Menge X in sich zwar eine innere, assoziative Verknüpfung mit neutralem Element id_X, die Menge $A(X)$ all dieser Abbildungen ist aber üblicherweise keine Gruppe, weil nicht alle Abbildungen invertierbar sind (s. S.28)
- Nun sei $X = K^n$ und F die Menge aller Abbildungen der Gestalt
$$f_A : \begin{cases} K^n \to K^n \\ x \mapsto Ax \end{cases}$$
mit einer festen $(n \times n)$-Matrix $A \in K^{n\times n}$.
- Mit den Rechenregeln für das Matrizenprodukt folgt
$$f_A \circ f_B(x) = f_A(f_B(x)) = f_A(Bx) = A(Bx) = (AB)x\,,$$
d.h. die Abbildung $f_A \circ f_B$ ist wieder in F.
- Die Hintereinanderausführung solcher Abbildungen wird gerade durch das Matrizenprodukt beschrieben. Wegen $\mathrm{id}_{K^n} = f_{\mathbb{1}_n}$ liegt auch das neutrale Element in F.
- Es sind nicht alle Abbildungen dieses Typs invertierbar (s. etwa das vorangegangene Beispiel mit den Nullteilern)
- Es ist also $(F, \circ)$ keine Gruppe.
- Deshalb betrachtet man die Menge aller invertierbaren Abbildungen aus F. Diese bildet mit der Komposition eine Gruppe
$$\mathrm{GL}(n,K) := \{f_A \mid f_A \ \text{ ist invertierbar}\}\,,$$
die so genannte *allgemeine lineare Gruppe* (englisch: **G**eneral **L**inear Group)
- Eine Abbildung $f_A \in F$ ist genau dann invertierbar, also in $\mathrm{GL}(n,K)$, wenn es ein $B \in K^{n\times n}$ bzw. ein $f_B \in F$ mit
$$f_A \circ f_B(x) = A(Bx) = (AB)x = x \quad \forall x \in K^n \iff$$
$$f_A \circ f_B = \mathrm{id}_{K^n} \iff A \cdot B = \mathbb{1}_n$$
gibt. Die Matrix B heißt die zu A *inverse Matrix*, i.Z. $B = A^{-1}$.
- Da f_A und A eineindeutig durcheinander festgelegt sind, kann man auch
$$\mathrm{GL}(n,K) := \{A \in K^{n\times n} \mid A \ \text{ ist invertierbar}\}\,,$$
schreiben.
- Statt invertierbar sagt man auch *regulär*, eine nicht invertierbare Matrix heißt auch *singulär*.

Beispiel 9.4.1. *Es ist*
$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}(2,\mathbb{R}), \quad \textit{denn}$$
$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \mathbb{1}_2 \iff \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

Beispiel 9.4.2. *Es ist*
$$\begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} -i & 1 \\ 1 & -i \end{pmatrix} \quad \textit{also}$$
$$\begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}, \frac{1}{2}\begin{pmatrix} -i & 1 \\ 1 & -i \end{pmatrix} \in \mathrm{GL}(2,\mathbb{C})$$

Satz 9.4.3. *Es sei K ein Körper und $n \in \mathbb{N}$. Für beliebige $A, B \in \mathrm{GL}(n,K)$ und $C, D \in K^{n\times n}$ gilt:*

(1) $AA^{-1} = A^{-1}A = \mathbb{1}_n$,

(2) $AC = \mathbb{1}_n \Rightarrow C = A^{-1}$, $DA = \mathbb{1}_n \Rightarrow D = A^{-1}$,

(3) $\left(A^{-1}\right)^{-1} = A$,

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

(5) $X, Y \in K^{n\times p} \Rightarrow AX = Y$ *hat die eindeutige Lösung* $X = A^{-1}Y$.

Beweis: Da $\mathrm{GL}(n, K)$ eine muliplikative nichtabelsche Gruppe ist, sind die hier aufgelisteten Eigenschaften direkte Folgerungen der Gruppendefinition 7.1.1 und der Sätze 7.1.6 und 7.1.8. Nur der letzte Punkt erweitert 7.1.6(4) etwas, da X, Y nicht zwingend quadratische Matrizen sind. Es ist durch Einsetzen klar, dass $X = A^{-1}Y$ eine Lösung ist. Ist $X' \in K^{n\times p}$ eine weitere Lösung, so ist $0 = AX - AX' = A(X - X')$. Multiplikation mit A^{-1} von links liefert dann $X = X'$. □

Satz 9.4.4. *Es sei K ein Körper und $n \in \mathbb{N}$. Hat $A \in K^{n\times n}$ eine Nullzeile oder -spalte, so ist A nicht invertierbar.*

Beweis: Ist die i-te Spalte von A eine Nullspalte, so ist für jede Matrix $B \in K^{n\times n}$ die i-te Spalte von BA eine Nullspalte, also kann nicht $BA = \mathbb{1}_n$ ein. Entsprechend ist die j-te Zeile von AB eine Nullzeile, wenn die j-te Zeile von A verschwindet. □

- Für eine invertierbare Matrix $A \in \mathrm{GL}(n, K)$ gibt es $A^{-1} \in \mathrm{GL}(n, K)$ mit $A \cdot A^{-1} = \mathbb{1}_n$.
- Fasst man diese Matrizengleichung spaltenweise auf, so sieht man, dass es sich um n LGSe handelt:
$$A \cdot (i\text{-te Spalte von } A^{-1}) = i\text{-te Spalte von } \mathbb{1}_n = e_i \text{ für } i = 1, \dots, n\,.$$
- Wegen $A \in \mathrm{GL}(n, K)$ sind diese n LGSe jeweils eindeutig lösbar.
- Wegen der jeweils gleichen Koeffizientenmatrix dieser n LGSe kann man diese aber simultan lösen:
- *Rechenschema:*
$$\boxed{(A|\mathbb{1}_n) \rightsquigarrow \cdots \rightsquigarrow (\mathbb{1}_n|\underbrace{B}_{=A^{-1}})}$$

10 Vektorräume

10.1 Grundlagen

4.11.10

Definition 10.1.1. *Es seien $V \neq \emptyset$ eine Menge, und K ein Körper. Ferner seien $+ : V \times V \to V$ eine innere Verknüpfung in V und $\cdot : K \times V \to V$ eine äußere Verknüpfung von V mit K. V heißt ein K-Vektorraum (oder linearer Raum), wenn gilt:*

(V_1) $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}\ \forall \mathbf{v}, \mathbf{w} \in V$.

(V_2) $(\mathbf{v} + \mathbf{w}) + \mathbf{u} = \mathbf{v} + (\mathbf{w} + \mathbf{u})\ \forall \mathbf{v}, \mathbf{w}, \mathbf{u} \in V$.

(V_3) $\exists \mathbf{0} \in V$ *mit* $\mathbf{0} + \mathbf{v} = \mathbf{v}\ \forall \mathbf{v} \in V$.

(V_4) $\forall \mathbf{v} \in V\ \exists \bar{\mathbf{v}} \in V$ *mit* $\bar{\mathbf{v}} + \mathbf{v} = \mathbf{0}$ $(\bar{\mathbf{v}} = -\mathbf{v})$.

(V_5) $(\alpha + \beta) \cdot \mathbf{v} = \alpha \cdot \mathbf{v} + \beta \cdot \mathbf{v}\ \forall \alpha, \beta \in K,\ \mathbf{v} \in V$.

(V_6) $(\alpha\beta) \cdot \mathbf{v} = \alpha \cdot (\beta \cdot \mathbf{v})\ \forall \alpha, \beta \in K,\ \mathbf{v} \in V$.

(V_7) $\alpha \cdot (\mathbf{v} + \mathbf{w}) = \alpha \cdot \mathbf{v} + \alpha \cdot \mathbf{w}\ \forall \alpha \in K,\ \mathbf{v}, \mathbf{w} \in V$.

(V_8) $1 \cdot \mathbf{v} = \mathbf{v}\ \forall \mathbf{v} \in V$.

- Die Punkte (V_1)-(V_4) sind die bereits bekannten Axiome einer abelschen Gruppe
- (V_5)-(V_8) regeln, dass die Multiplikation mit Skalaren ordentlich mit der Addition in dieser Gruppe zusammenwirkt.

Folgerung 10.1.2. *Es sei K ein Körper und $V \neq \emptyset$ eine Menge mit einer inneren Verknüpfung $+ : V \times V \to V$ und einer äußeren Verknüpfung $\cdot : K \times V \to V$. V heißt ein K-Vektorraum, wenn gilt:*

(VR_1) $(V, +)$ *ist ein abelsche Gruppe.*

(VR_2) *Die Multiplikation mit Skalaren ist auf folgende Weise mit $(V, +)$ verträglich*

(a) $(\alpha + \beta) \cdot \mathbf{v} = \alpha \cdot \mathbf{v} + \beta \cdot \mathbf{v}\ \forall \alpha, \beta \in K,\ \mathbf{v} \in V$.
(b) $(\alpha\beta) \cdot \mathbf{v} = \alpha \cdot (\beta \cdot \mathbf{v})\ \forall \alpha, \beta \in K,\ \mathbf{v} \in V$.
(c) $\alpha \cdot (\mathbf{v} + \mathbf{w}) = \alpha \cdot \mathbf{v} + \alpha \cdot \mathbf{w}\ \forall \alpha \in K,\ \mathbf{v}, \mathbf{w} \in V$.
(d) $1 \cdot \mathbf{v} = \mathbf{v}\ \forall \mathbf{v} \in V$.

10.2 Wichtige Beispiele

Beispiel 10.2.1. *$V = (\{0\}, +)$ ist mit der Multiplikation mit Skalaren $k \cdot 0 := 0$ für alle $k \in K$ für jeden beliebigen Körper K ein Vektorraum, der so genannte nulldimensionale Vektorraum.*

Beispiel 10.2.2. *Jeder Körper K ist auch ein K-Vektorraum, wenn man für die Multiplikation mit Skalaren einfach die Körpermultiplikation verwendet. Man spricht in diesem Zusammenhang auch von dem eindimensionalen arithmetischen K-Vektorraum.*

Beispiel 10.2.3. *Für jeden Körper K ist das n-fache kartesische Produkt $V = K^n$ mit der Addition*

$$(v_1, \ldots, v_n) + (w_1, \ldots, w_n) := (v_1 + w_1, \ldots, v_n + w_n)$$

und der Multiplikation mit Skalaren

$$k \cdot (v_1, \ldots, v_n) := (kv_1, \ldots, kv_n)$$

ein K-Vektorraum. Man spricht in diesem Zusammenhang auch von dem n-dimensionalen arithmetischen K-Vektorraum. Dies ist eine Verallgemeinerung des vorhergehenden Beispiels.

Beispiel 10.2.4. *Ist K ein Körper, S eine beliebige nichtleere Menge und*

$$V := K^S = \{f \mid S \to K\}\,.$$

Dann ist V mit den Verknüpfungen

$$\begin{aligned} (f+g)(s) :&= f(s) + g(s) \qquad \forall f, g \in V,\ s \in S \\ (kf)(s) :&= k \cdot f(s) \qquad \forall f \in V,\ k \in K,\ s \in S \end{aligned}$$

ein K-Vektorraum.

Beispiel 10.2.5. *Vektorräume des Typs 10.2.4 spielen auch in der Analysis eine wichtige Rolle, etwa mit $K = S = \mathbb{R}$, also $V = \mathbb{R}^{\mathbb{R}}$. Statt ganz V betrachtet man dort aber meist geeignete Teilmengen, etwa*

$$\{f : \mathbb{R} \to \mathbb{R} \mid f \ \ \textit{ist stetig}\}$$

(oft mit $\mathcal{C}(\mathbb{R}, \mathbb{R})$ bezeichnet, stetig=Continuous) oder

$$\{f : \mathbb{R} \to \mathbb{R} \mid f \ \ \textit{ist differenzierbar}\}\,.$$

Beispiel 10.2.6. *Ist K ein Körper und speziell $S = \mathbb{N}_0$, also*

$$V := K^{\mathbb{N}_0} = \{f \mid \mathbb{N}_0 \to K\}\,,$$

so kann man diese Abbildungen auch durch Angabe all ihrer Bilder eindeutig beschreiben, etwa in der Gestalt $(f(0), f(1), f(2), f(3), \dots)$. Meist schreibt man stattdessen $(a_0, a_1, a_2, a_3, \dots)$ mit $f(i) =: a_i \in K$ für $i \in \mathbb{N}_0$. Die Element von $V := K^{\mathbb{N}_0}$ heißen Folgen über K und bilden mit den in 10.2.4 eingeführten Abbildungen einen Vektorraum. In der Folgenschreibweise sehen die Verknüpfungen dann so aus:

$$\begin{aligned} &(a_0, a_1, a_2, \dots) + (b_0, b_1, b_2, \dots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots)\,, \\ &k(a_0, a_1, a_2, a_3, \dots) = (ka_0, ka_1, ka_2, ka_3, \dots) \\ &\forall k \in K, (a_0, a_1, a_2, \dots), (b_0, b_1, b_2, \dots) \in K^{\mathbb{N}_0}\,. \end{aligned}$$

Beispiel 10.2.7. *Ist K ein Körper und speziell $S = \{1, 2, \dots, n\}$ mit $n \in \mathbb{N}$, so kann man diese Abbildungen wie im letzten Beispiel wieder durch Angabe all ihrer Bilder eindeutig beschreiben, etwa in der Gestalt $(f(1), \dots, f(n))$. Schreibt man stattdessen wieder $(a_1, \dots, a_n)$ mit $f(i) =: a_i \in K$ für $i \in \{1, 2, \dots, n\}$, so sieht man*

$$K^{\{1,2,\dots,n\}} = K^n\,,$$

was nochmals zeigt, dass die Schreibweise K^S sehr passend ist.

Satz 10.2.8. *Es sei K ein Körper und V ein K-Vektorraum. Dann gilt:*

(1) $0 \cdot \mathbf{v} = \mathbf{0}\ \forall\, \mathbf{v} \in V$.

(2) $\alpha \cdot \mathbf{0} = \mathbf{0}\ \forall\, \alpha \in K$.

(3) $\alpha \cdot \mathbf{v} = \mathbf{0} \Rightarrow \alpha = 0 \vee \mathbf{v} = \mathbf{0}\ \forall\, \alpha \in K\,,\ \mathbf{v} \in V$.

(4) $(-\alpha) \cdot \mathbf{v} = -\alpha \cdot \mathbf{v} = \alpha \cdot (-\mathbf{v})\ \forall\, \alpha \in K\,,\ \mathbf{v} \in V$.

Beweis: Einfache Übung □

- Auch hier wurden nochmal zur Deutlichkeit die Vektoren fett geschrieben. Die gezeigten Rechenregeln zeigen aber, dass das nicht nötig ist.
- Selbst den Skalar 0 und den Nullvektor muss man nicht durch die Schreibweise unterscheiden, denn es rechnet sich mit beiden ganz ähnlich und aus dem Zusammenhang ist jeweils klar, was gemeint ist.
- Ab jetzt wird deshalb für Vektoren keine besondere Schreibweise mehr verwendet.

Definition 10.2.9. *Es seien K ein Körper, V ein K-Vektorraum und $U \subseteq V$. Die Menge U heißt Untervektorraum (oder einfach Unterraum, Teilraum oder auch linearer Teilraum), falls gilt:*

(U_1) $U \neq \emptyset$

(U_2) $u_1, u_2 \in U \Rightarrow u_1 + u_2 \in U$

(U_3) $u \in U\,,\, \alpha \in K \Rightarrow \alpha u \in U$

Folgerung 10.2.10. *Damit ist U (mit den gleichen Verknüpfungen wie in V) ebenfalls ein Vektorraum, denn (U_3) mit $\alpha = -1$ garantiert $-u \in U$ und mit (U_2) ist das gerade der Untergruppentest, d.h. $(U, +)$ ist eine Untergruppe von $(V, +)$. Wegen (U_3) ist diese Untergruppe auch bzgl. der Multiplikation mit Skalaren abgeschlossen. Da die Regeln (VR_2) in V gelten, gelten sie auch in U.*

Beispiel 10.2.11.

$$U := \{(x, 2x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\} = \mathbb{R}\begin{pmatrix}1\\2\end{pmatrix}$$

ist ein Untervektorraum (UVR) des 2-dimensionalen arithmetischen Vektorraumes $\mathbb{R}^2$, denn

$$\begin{aligned}
&U \neq \emptyset,\\
&\lambda_1\begin{pmatrix}1\\2\end{pmatrix}, \lambda_2\begin{pmatrix}1\\2\end{pmatrix} \in U \Rightarrow \lambda_1\begin{pmatrix}1\\2\end{pmatrix} + \lambda_2\begin{pmatrix}1\\2\end{pmatrix} \in U \quad \textit{und}\\
&\lambda\begin{pmatrix}1\\2\end{pmatrix} \in U \Rightarrow \alpha\lambda\begin{pmatrix}1\\2\end{pmatrix} \in U \quad \forall \alpha \in \mathbb{R}\,.
\end{aligned}$$

Geometrisch betrachtet, ist das eine Gerade durch den Ursprung.

Beispiel 10.2.12. *Die Analysis liefert mit Sätzen wie „Die Summe stetiger Funktionen ist wieder stetig“ die Begründung, dass die Teilmengen $\mathcal{C}(\mathbb{R}, \mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} \mid f \;\text{ ist stetig}\}$ oder $\{f : \mathbb{R} \to \mathbb{R} \mid f \;\text{ ist differenzierbar}\}$ von $V = \mathbb{R}^{\mathbb{R}}$ (und viele ähnliche Gebilde) Untervektorräume von V sind.*

Satz 10.2.13. *Der Durchschnitt beliebig vieler Unterräume eines Vektorraumes ist ein Unterraum.*

Beweis: Es sei $\mathcal{U}$ eine Menge von Untervektorräumen eines Vektorraumes V und

$$D := \bigcap_{U \in \mathcal{U}} U\,.$$

Da 0 in jedem Untervektorraum enthalten ist, ist 0 auch in D, also D nicht leer. Für $a, b \in D$ gilt auch $a, b \in U$ für alle $U \in \mathcal{U}$. Da U ein Teilraum ist, ist auch $a + b \in U$ und $ka \in U$ für jedes $k \in K$, also auch $a + b \in D$ und $ka \in D$ für jedes $k \in K$. $\square$

Beispiel 10.2.14. *Die Vereinigung von Vektorräumen ist im Allgemeinen kein Vektorraum:*

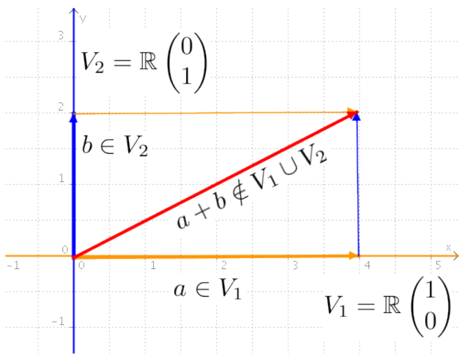


Abbildung 21: Vereinigung von Vektorräumen

10.3 Erzeugendensysteme

Definition 10.3.1. *Es seien K ein Körper, V ein K-Vektorraum, $M \subseteq V$ und*

$$\mathcal{U} := \{U \mid U \;\; \textit{UVR von} \;\; V \;\; \textit{mit} \;\; M \subseteq U\}$$

die Menge der Unterräume von V, die M umfassen. Der von M aufgespannte oder erzeugte Unterraum *(auch lineare Hülle oder Spann genannt) ist definiert als*

$$\operatorname{Lin}(M) := \bigcap_{U \in \mathcal{U}} U \,.$$

Ist $V = \operatorname{Lin}(M)$, so heißt M ein Erzeugendensystem von V. Statt $\operatorname{Lin}(M)$ findet man oft auch die Schreibweisen $\operatorname{span}(M)$ oder $\langle M \rangle$.

- Nach 10.2.13 ist $\operatorname{Lin}(M)$ für jedes $M \subseteq V$ ein Teilraum.
- Direkt aus der Definition folgt, dass $\operatorname{Lin}(M)$ der kleinste M umfassende Teilraum von V ist.
- Die Definition eines Erzeugendensystems hat zwar viele Vorteile, ist aber nicht sonderlich geeignet, wenn man die Elemente aus $\operatorname{Lin}(M)$ direkt angeben möchte. Hierzu ist der folgende Begriff nützlich:

Definition 10.3.2. *Es seien K ein Körper und $v_1, \ldots, v_n \in V$ endlich viele Vektoren eines K-Vektorraumes V und $M \subseteq V$ eine nichtleere Teilmenge von V. Jeder Vektor $v = \sum_{i=1}^{n} c_i v_i$ mit $c_i \in K$ heißt Linearkombination von $v_1, \ldots, v_n$. Ein Vektor v heißt Linearkombination der Menge M, wenn er Linearkombination von endlich vielen Vektoren $v_1, \ldots, v_n \in M$ ist.*

Satz 10.3.3. *Es sei $M \subseteq V$ eine nichtleere Teilmenge von V und M^* die Menge aller Linearkombinationen von M. Dann gilt*

$$\operatorname{Lin}(M) = M^* \,.$$

Beweis: M^* ist ein Unterraum von V, denn

- $M^* \supseteq M \neq \emptyset$
- Mit $v, w \in M^*$ ist auch $v + w \in M^*$, denn v und w sind Linearkombinationen, also ist auch $v + w$ eine.
- Mit $v \in M^*$ und $k \in K$ ist auch $kv \in M^*$.
- Da $M^* \supseteq M$ und $\operatorname{Lin}(M)$ der kleinste M umfassende Teilraum von V ist, folgt $\operatorname{Lin}(M) \subseteq M^*$. Andrerseits enthält $\operatorname{Lin}(M)$ als Vektorraum insbesondere alle Linearkombinationen von Vektoren aus M und somit auch M^*, d.h. $M^* \subseteq \operatorname{Lin}(M)$, d.h. $M^* = \operatorname{Lin}(M)$.

□

Definition 10.3.4. *Es seien K ein Körper, V ein K-Vektorraum und $\mathcal{U}$ eine Menge von Teilräumen von V. Der Summenraum von $\mathcal{U}$ ist*

$$\sum_{U \in \mathcal{U}} U := \operatorname{Lin}\left(\bigcup_{U \in \mathcal{U}} U \right) .$$

Ist $\mathcal{U} = \{U_1, \ldots, U_n\}$, so schreibt man den Summenraum auch in der Form $U_1 + \cdots + U_n$.

Beispiel 10.3.5. *In $\mathbb{C}^3$ seien $u = \begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix}$, $v = \begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}$, $w = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.*
Dann gilt $\operatorname{Lin}(u) + \operatorname{Lin}(v) + \operatorname{Lin}(w) = \operatorname{Lin}(u, v)$, denn $u - iv = w$.

Satz 10.3.6. *Es seien K ein Körper, V ein K-Vektorraum und $\mathcal{U} = \{U_1, \ldots, U_n\}$ eine Menge von (paarweise verschiedenen) Teilräumen von V. Der Summenraum $U_1 + U_2 + \cdots + U_n$ besteht genau aus den Vektoren $v \in V$, die sich in der Form $v = \sum_{i=1}^{n} u_i$ mit $u_i \in U_i$ schreiben lassen.*

Beweis: Da $M := \bigcup_{U \in \mathcal{U}} U \neq \emptyset$, lässt sich jedes $v \in \sum_{U \in \mathcal{U}} U$ als Linearkombination von M darstellen, d.h. es gibt $\tilde{u}_1, \ldots, \tilde{u}_m \in \bigcup_{U \in \mathcal{U}} U$ und $d_1, \ldots, d_m \in K$ mit

$$v = d_1 \tilde{u}_1 + \cdots + d_m \tilde{u}_m \,.$$

Diese Vektoren kann man so nummerieren, dass die ersten in U_1, die nächsten in U_2 usw. und die letzten in U_n sind. Da dies alles Untervektorräume sind, kann man aber diese Grüppchen auch jeweils zusammenfassen zu einem einzigen $u_1 \in U_1$, $u_2 \in U_2$ usw. bis $u_n \in U_n$. Dies zeigt die eine Hälfte der Behauptung, die andere ist
8.11.10 eh klar. □

10.4 Basen

Beispiel 10.4.1. *$v_1 := (3,0)$, $v_2 := (0,2)$, $v_3 := (2,3)$, ist ein Erzeugendensystem von $\mathbb{R}^2$ denn jeder Vektor des $\mathbb{R}^2$ kann als Linearkombination dieser Vektoren dargestellt werden, etwa*

$$(x,y) = \left(\tfrac{x}{3} - 4\right) v_1 + \left(\tfrac{y}{2} - 9\right) v_2 + 6v_3$$

Diese Darstellung ist aber nicht eindeutig:

$$w := (1,0) \Rightarrow w = \tfrac{1}{3}v_1 \quad \textit{oder} \quad w = \tfrac{1}{2}v_3 - \tfrac{3}{4}v_2 \,.$$

Man versteht das Problem besser, wenn man sich klar macht, dass sich der Nullvektor als nichttriviale Linearkombination von v_1, v_2, v_3 darstellen lässt:

$$w = \tfrac{1}{3}v_1 = \tfrac{1}{2}v_3 - \tfrac{3}{4}v_2 \Rightarrow \tfrac{1}{3}v_1 + \tfrac{3}{4}v_2 - \tfrac{1}{2}v_3 = 0$$

Definition 10.4.2. *Es seien K ein Körper und V ein beliebiger K-Vektorraum. Vektoren $v_1, \ldots, v_n \in V$ heißen linear unabhängig, wenn der Nullvektor nur die so genannte triviale Linearkombination*

$$0 = 0 \cdot v_1 + \cdots + 0 \cdot v_n$$

zulässt, d.h. wenn aus $0 = \sum_{i=1}^{n} c_i v_i$ mit $c_i \in K$ zwingend $c_1 = c_2 = \cdots = c_n = 0$ folgt.
Gibt es dagegen auch nichttriviale Linearkombinationen der 0, so heißen $v_1, \ldots, v_n$ linear abhängig.
Eine Teilmenge $M \subseteq V$ heißt linear unabhängig, wenn je endlich viele(!) verschiedene Vektoren aus M linear unabhängig sind, sonst linear abhängig.

Beispiel 10.4.3. *In 10.4.1 wurde gezeigt, dass $\{v_1, v_2, v_3\}$ linear abhängig sind, denn es gibt z.B die nichttriviale Linearkombination*

$$4v_1 + 9v_2 - 6v_3 = 0$$

der 0. Dagegen ist $\{v_1, v_2\}$ linear unabhängig, denn

$$c_1 v_1 + c_2 v_2 = (3c_1, 2c_2) = (0,0) \iff c_1 = c_2 = 0$$

Entsprechend sind $\{v_1, v_3\}$, $\{v_2, v_3\}$, $\{v_1\}$, $\{v_2\}$ und $\{v_3\}$ linear unabhängig.

Beispiel 10.4.4. *Im Sinne der Definition ist auch $\emptyset$ eine linear unabhängige Menge, da mit ihr keine nichttriviale Linearkombination der 0 möglich ist. Die Linearkombination $\sum_{v_i \in \emptyset} c_i v_i$ (leere Summe) setzt man üblicherweise gleich dem Nullvektor. Da jeder Vektorraum den Nullvektor enthält, gilt $\mathrm{Lin}(\emptyset) = \{0\}$, d.h. auch für die leere Menge gilt 10.3.3, das Erzeugnis der leeren Menge ist der nulldimensionalen Vektorraum $\{0\}$.*

Satz 10.4.5. *Es seien K ein Körper und V ein beliebiger K-Vektorraum. Dann gilt:*
(1) Teilmengen linear unabhängiger Mengen in V sind linear unabhängig.
(2) Obermengen linear abhängiger Mengen in V sind linear abhängig.
(3) Ein einzelner Vektor $v \in V \setminus \{0\}$ ist immer linear unabhängig.

Beweis: in der Übung □

Satz 10.4.6. *Es seien K ein Körper und V ein beliebiger K-Vektorraum. Eine aus mindestens zwei Vektoren bestehende Teilmenge $M \subset V$ ist genau dann linear abhängig, wenn ein Vektor $m \in M$ existiert, der sich als Linearkombination von $M \setminus \{m\}$ schreiben lässt, d.h. wenn es endlich viele $c_i \in K$ und $v_i \in M$ gibt mit $v_i \neq m$ und*

$$m = \sum_{i=1}^{n} c_i v_i \,.$$

Beweis: „$\Leftarrow$": Ist $m = \sum_{i=1}^{n} c_i v_i$, so folgt $0 = \sum_{i=1}^{n} c_i v_i - m$. Da dies wegen der -1 bei m eine nichttriviale Linearkombination ist, sind $v_1, \ldots, v_n, m$ und somit M linear abhängig.
„$\Rightarrow$": Ist M linear abhängig, so gibt es endlich viele $d_i \in K$ (nicht alle $= 0$, etwa $d_r \neq 0$) und $v_i \in M$ mit $0 = \sum_{i=1}^{r} d_i v_i$. Somit kann man nach v_r auflösen und erhält $m = v_r = \sum_{i=1}^{r-1} c_i v_i$ mit $c_i = -\frac{d_i}{d_r}$. □

Definition 10.4.7. *Es seien K ein Körper und V ein beliebiger K-Vektorraum. Eine Teilmenge B eines K-Vektorraumes V heißt Basis von V, wenn gilt*

(1) $\mathrm{Lin}\,(B) = V$,

(2) *B ist linear unabhängig.*

Beispiel 10.4.8. *Nach 10.4.4 ist klar, dass $\emptyset$ eine Basis des nulldimensionalen Vektorraumes $V = \{0\}$ ist.*

Beispiel 10.4.9. *Für jede Zahl $k \in K \setminus \{0\}$ ist $\{k\}$ eine Basis des eindimensionalen arithmetischen K-Vektorraumes K, denn jedes $k' \in K$ lässt sich als $\frac{k'}{k}$-faches von k darstellen und $\{k\}$ ist nach 10.4.5(3) linear unabhängig.*

Beispiel 10.4.10.

- *Laut 10.4.3 sind $\{v_1, v_2, v_3\}$ linear abhängig, also keine Basis des $\mathbb{R}^2$. Die Mengen $\{v_1, v_2\}$, $\{v_1, v_3\}$, $\{v_2, v_3\}$, $\{v_1\}$, $\{v_2\}$ und $\{v_3\}$ sind dagegen linear unabhängig.*
- *Die drei zweielementigen Mengen $\{v_1, v_2\}$, $\{v_1, v_3\}$ und $\{v_2, v_3\}$ erzeugen den ganzen $\mathbb{R}^2$, sind also Basen des $\mathbb{R}^2$.*
- *Die einelementigen Mengen $\{v_1\}$, $\{v_2\}$ und $\{v_3\}$ erzeugen dagegen nur echte Teilräume des $\mathbb{R}^2$, sind also nicht Basen des $\mathbb{R}^2$.*

Beispiel 10.4.11. *Im arithmetischen Vektorraum K^n bilden die Vektoren*

$$\begin{aligned} e_1 &:= (1, 0, 0, \ldots, 0, 0, 0) \\ e_2 &:= (0, 1, 0, \ldots, 0, 0, 0) \\ &\;\vdots \\ e_{n-1} &:= (0, 0, 0, \ldots, 0, 1, 0) \\ e_n &:= (0, 0, 0, \ldots, 0, 0, 1) \end{aligned}$$

eine Basis, die so genannte kanonische Basis oder Standardbasis. Die Vektoren $e_1, \ldots, e_n$ heißen kanonische Einheitsvektoren des K^n.

Satz 10.4.12. *Es seien K ein Körper, V ein K-Vektorraum und $B \subseteq V$. Die folgenden Aussagen sind paarweise äquivalent (Punkt (4) nur für Vektorräume ungleich dem Nullraum):*

(1) *B ist eine Basis von V.*

Eine Basis ist ein linear unabhängiges Erzeugendensystem.

(2) $\mathrm{Lin}\,(B) = V$, *aber für jedes $C \subsetneq B$ gilt* $\mathrm{Lin}\,(C) \neq V$.

Eine Basis ist ein minimales Erzeugendensystem.

(3) *B ist linear unabhängig, aber für jedes $C \supsetneq B$ ist C linear abhängig.*

Eine Basis ist eine maximale linear unabhängige Menge.

(4) *Jeder Vektor aus V kann auf genau eine Weise als Linearkombination von B dargestellt werden.*

Basis ist Erzeugendensystem, das eindeutige Darstellung erlaubt.

Beweis: Zuerst der einfache Fall $V = \{0\}$, d.h. es ist nur die Äquivalenz der Punkte (1)-(3) zu zeigen. V hat nur die Teilmengen $\emptyset$ und sich selbst. Für $\emptyset$ sind die Aussagen (1)-(3) wahr, für V falsch. Dies zeigt die Behauptung.
Nun sei also $V \neq \{0\}$. Statt der insgesamt 6 zu zeigenden Äquivalenzen, d.h. 12 Implikationen, kann man sich die Arbeit deutlich erleichtern, wenn man eine geeignete Teilmenge dieser Implikationen zeigt und die Transitivität von „$\Rightarrow$“ ausnutzt, etwa

$$(1) \Rightarrow (4) \Rightarrow (2) \Rightarrow (3) \Rightarrow (1)$$

Trotzdem ist das noch einiges an Arbeit. Ich zeige hier nur exemplarisch die erste Implikation $(1) \Rightarrow (4)$ bzw. die äquivalente Aussage $\neg(4) \Rightarrow \neg(1)$. Die anderen Beweise finden sich zum Nachlesen praktisch in jedem Buch über Lineare Algebra.
Es gibt zwei mögliche Gründe für $\neg(4)$, nämlich
(i) Nicht jeder Vektor $v \in V$ lässt sich als Linearkombination von B darstellen. Das heißt $\mathrm{Lin}\,(B) \neq V$, d.h. B ist keine Basis.
(ii) Es lässt sich zwar jeder Vektor $v \in V$ als Linearkombination von B darstellen, die Darstellung ist aber nicht eindeutig, d.h. es gibt $c_i, d_i \in K$ mit $c_i \neq d_i$ für mindestens ein i zwischen 1 und n und $b_i \in B$ (paarweise verschieden) mit

$$\sum_{i=1}^{n} c_i b_i = v = \sum_{i=1}^{n} d_i b_i \,.$$

Daraus folgt aber sofort $\sum_{i=1}^{n} (c_i - d_i) b_i = 0$, also die lineare Abhängigkeit von B. Dann ist B aber keine Basis. $\square$

10.5 Existenz

- Jeder Vektorraum besitzt ein Erzeugendensystem (z.B. sich selbst).
- Die Frage, ob jeder Vektorraum auch eine Basis besitzt, ist aber keineswegs einfach zu beantworten.
- Die Antwort lautet „ja", zum Nachweis ist aber ein ziemliches mathematische Geschütz nötig, das so genannte *Zornsche Lemma.* [6]
- Die Beweisidee ist dabei aber ziemlich einfach:
  - Man startet „von unten", d.h. von der linear unabhängigen Menge $\emptyset$, und versucht diese durch Hinzunahme von Vektoren so zu erweitern, dass die neue Menge jeweils linear unabhängig bleibt.
  - Bei diesem Prozess kommt man irgendwann bei einer Menge an, die selbst noch linear unabhängig ist, aber durch die Hinzunahme eines jeden beliebigen Vektors linear abhängig werden würde (hierzu braucht man im unendlichen Fall das Zornsche Lemma).
  - Nach 10.4.12(3) ist solch eine maximale linear unabhängige Menge eine Basis des betrachteten Vektorraums.

Definition 10.5.1. *Es seien K ein Körper und V ein K-Vektorraum. Besitzt V eine endliche Teilmenge $E \subseteq V$ mit $\operatorname{Lin}(E) = V$, so heißt V endlichdimensional, sonst unendlichdimensional.*

- Ist ein Vektorraum V endlichdimensional, so kann man die Existenz einer Basis auch wie folgt (ohne das Zornsche Lemma) zeigen:
  - Man startet „von oben", d.h. von dem endlichen Erzeugendensystem E, das es laut Definition des endlichdimensionalen Vektorraumes gibt. Ist E linear unabhängig, hat man bereits eine Basis.
  - Ist E linear abhängig, so wirft man mit Hilfe von 10.4.6 „unnötige" Vektoren aus E raus, bis man bei einem minimalen Erzeugendensystem anlangt.
  - Nach 10.4.12(2) ist solch ein minimales Erzeugendensystem dann eine Basis des betrachteten Vektorraums

Satz 10.5.2 (Basissatz)**.** *Jeder Vektorraum besitzt eine Basis.*

Satz 10.5.3 (Basisergänzungssatz)**.** *Es seien K ein Körper, V ein K-Vektorraum und $A \subseteq V$ linear unabhängig. Dann gibt es eine Basis B von V mit $A \subseteq B$, d.h. A lässt sich zu einer Basis B von V ergänzen.*

Beweis: Man geht prinzipiell wie bei dem Beweis mit dem Zornschen Lemma vor (s. S.49), startet aber nicht mit $\emptyset$, sondern mit der Menge A. □

Satz 10.5.4 (Kleiner Austauschsatz)**.** *Es seien K ein Körper, V ein endlich erzeugter K-Vektorraum mit einer Basis $B = \{b_1, \dots, b_n\}$. Besitzt $b \in V$ die (eindeutige) Darstellung*

$$b = \lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_n b_n$$

mit $\lambda_1, \dots, \lambda_n \in K$ und $\lambda_j \neq 0$ für ein j zwischen 1 und n, so erhält man eine andere Basis B' von V, wenn man b_j gegen dieses b austauscht, d.h. $B' := \{b_1, \dots, b_{j-1}, b, b_{j+1}, \dots, b_n\}$ ist auch eine Basis von V.

Beweis: Der Einfachheit halber werden die Basisvektoren so nummeriert, dass in $b = \lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_n b_n$ gilt $\lambda_1 \neq 0$.
Ist nun $v = \sigma_1 b_1 + \sigma_2 b_2 + \cdots + \sigma_n b_n$ die Darstellung eines Vektors $v \in V$ bezüglich der gegebenen Basis B, so gilt

$$\begin{aligned} &\tfrac{\sigma_1}{\lambda_1} b = \sigma_1 b_1 + \tfrac{\sigma_1 \lambda_2}{\lambda_1} b_2 + \cdots + \tfrac{\sigma_1 \lambda_n}{\lambda_1} b_n \iff \\ &\sigma_1 b_1 = \tfrac{\sigma_1}{\lambda_1} b - \tfrac{\sigma_1 \lambda_2}{\lambda_1} b_2 - \cdots - \tfrac{\sigma_1 \lambda_n}{\lambda_1} b_n = \tfrac{\sigma_1}{\lambda_1} b - \sum_{i=2}^n \tfrac{\sigma_1 \lambda_i}{\lambda_1} b_i \end{aligned}$$

[6] Das Wort „Lemma" kommt aus dem Lateinischen (und das wiederum von dem griechischen Wort „lambanein" - nehmen) und bedeutet „alles, was man nimmt". In der Mathematik wird das meist für einen Hilfssatz bei Beweisführungen benutzt. Im vorliegenden Fall ist das etwas untertrieben, denn es handelt sich um eines der Grundaxiome der Mengenlehre. Zusammen mit der Zermelo-Fraenkel-Mengenlehre kann man zeigen, dass dieses Lemma zu anderen wesentlichen Grundsätzen, wie dem *Auswahlaxiom* oder dem *Wohlordnungssatz* beweisbar äquivalent ist, d.h. man setzt jeweils einen dieser Sätze als Axiom voraus und folgert die beiden anderen.

Setzt man dies statt $\sigma_1 b_1$ in die Basisdarstellung von v ein, so erhält man

$$v = \sigma_1 b_1 + \sigma_2 b_2 + \cdots + \sigma_n b_n = \tfrac{\sigma_1}{\lambda_1} b - \sum_{i=2}^n \tfrac{\sigma_1 \lambda_i}{\lambda_1} b_i + \sum_{i=2}^n \sigma_i b_i =$$
$$= \tfrac{\sigma_1}{\lambda_1} b + \sum_{i=2}^n (\sigma_i - \tfrac{\sigma_1 \lambda_i}{\lambda_1}) b_i$$

d.h. jedes $v \in V$ lässt sich als Linearkombination von $B' = \{b, b_2, \ldots, b_n\}$ darstellen, B' ist Erzeugendensystem von V.
Dieses B' ist auch linear unabhängig, denn es gilt

$$\mu_1 b + \mu_2 b_2 + \cdots + \mu_n b_n = 0 \iff \mu_1 \lambda_1 b_1 + \sum_{i=2}^{n} (\mu_1 \lambda_i + \mu_i) b_i = 0$$

und daraus folgt wegen der linearen Unabhängigkeit von B $\mu_1 \lambda_1 = 0$ und $\mu_1 \lambda_i + \mu_i = 0$ für $i = 2, \ldots, n$.
Wegen $b_1 \neq 0$ und $\lambda_1 \neq 0$ liest man $\mu_1 = 0$ und somit auch $\mu_i = 0$ für $i = 2, \ldots, n$ ab, d.h. B' ist linear unabhängig. □

Satz 10.5.5 (Steinitzscher Austauschsatz). *Es seien K ein Körper und V ein endlich erzeugter K-Vektorraum mit einer Basis $B = \{b_1, \ldots, b_n\}$. Sind $a_1, \ldots, a_k \in V$ linear unabhängig, so gilt $k \leq n$ und es gibt eine Permutation $\pi \in S_n$, so dass*

$$B' := \{a_1, \ldots, a_k, b_{\pi(k+1)}, \ldots, b_{\pi(n)}\}$$

ebenfalls eine Basis von V ist.

Beweis: Induktion nach k: Für $k = 1$ ist die Aussage des Satzes gerade die Aussage des kleinen Austauschsatzes, ist also bereits bewiesen.
Nun sei also $k > 1$ und wir gehen von der Induktionsvoraussetzung $k - 1 \leq n$ und einer Basis $B'' := \{a_1, \ldots, a_{k-1}, b_{\sigma(k)}, \ldots, b_{\sigma(n)}\}$ aus, wobei $\sigma \in S_n$ sei. Wäre $k - 1 = n$, so wäre schon $\{a_1, \ldots, a_{k-1}\}$ eine Basis von V, was aber nicht möglich ist, denn nach Voraussetzung ist $\{a_1, \ldots, a_k\}$ linear unabhängig. Es ist also $k - 1 < n \iff k \leq n$.
Bezüglich der Basis B'' von V besitzt $a_k \in V$ die eindeutige Darstellung

$$a_k = \sum_{i=1}^{k-1} \lambda_i a_i + \sum_{i=k}^{n} \lambda_i b_{\sigma(i)} \text{ mit } \lambda_1, \ldots, \lambda_n \in K\,.$$

In dieser Darstellung muss mindestens einer der Koeffizienten $\lambda_k, \ldots, \lambda_n$ von Null verschieden sein, denn sonst ergäbe sich ein Widerspruch zur linearen Unabhängigkeit von $\{a_1, \ldots, a_k\}$.
Ist nun etwa $k \leq \ell \leq n$ und $\lambda_\ell \neq 0$, so kann man nach dem kleinen Austauschsatz den zugehörigen Basisvektor $b_{\sigma(\ell)}$ der Basis B'' gegen a_k austauschen und erhält so eine neue Basis.

$$\{a_1, \ldots, a_{k-1}, b_{\sigma(k)}, \ldots, b_{\sigma(\ell-1)}, a_k, b_{\sigma(\ell+1)}, \ldots, b_{\sigma(n)}\}$$

Um genau die gewünschte Form von B' zu bekommen, vertauscht man noch die Position ℓ und die Position k, d.h. man wendet noch die entsprechende Transposition ρ an. Dann ist $\pi = \rho \circ \sigma$. □

- Der Steinitzsche Austauschsatz gibt jetzt direkt Auskunft über die Größe von Basen.
- Sind nämlich $B = \{b_1, \ldots, b_n\}$ und $B' = \{a_1, \ldots, a_k\}$ zwei Basen eines endlich erzeugten K-Vektorraumes V, so folgt
  - mit der Basis B und der linear unabhängige Menge B': $k \leq n$,
  - mit der Basis B' und der linear unabhängige Menge B: $n \leq k$.
- d.h. es gilt $k = n$; je zwei Basen eines endlich erzeugten K-Vektorraumes V sind gleichmächtig!
- Der Austauschsatz lässt sich (bei entsprechender Umformulierung) auf den Fall unendlicher Basen ausdehnen.
- Damit kann man dann auch zeigen, dass je zwei unendliche Basen die gleiche Kardinalität besitzen.
- Dies ermöglicht die folgende Definition

Definition 10.5.6. *Es seien K ein Körper und V ein endlich erzeugter K-Vektorraum mit einer Basis $B = \{b_1, \dots, b_n\}$. Dann heißt $\dim(V) := |B| = n$ die Dimension von V. Ist V nicht endlich erzeugt, so wurde bereits in 10.5.1 definiert, dass er dann unendlichdimensional heißt. In diesem Fall schreibt man $\dim(V) = \infty$.*

Beispiel 10.5.7. *Es ist $\dim(\{0\}) = 0$, denn die Basis $\emptyset$ von $\{0\}$ hat 0 Elemente.*

Beispiel 10.5.8. *Es ist $\dim(K^n) = n$, denn die kanonische Basis $\{e_1, \dots, e_n\}$ hat n Elemente.*

Beispiel 10.5.9. *Es ist $\dim(\mathcal{C}(\mathbb{R}, \mathbb{R})) = \infty$; man kann nicht einmal eine Basis explizit angeben.* 11.11.10

Satz 10.5.10. *Es seien K ein Körper und V ein endlich erzeugter K-Vektorraum, also $\dim(V) < \infty$. Für jeden Unterraum U von V gilt dann*

(1) $\dim(U) \leq \dim(V)$,

(2) Ist $\dim(U) = \dim(V)$, so gilt $U = V$.

Beweis: (1) Ist $\dim(U) = n$, so besitzt U eine Basis $\{b_1, \dots, b_n\}$, die sich nach dem Basisergänzungssatz 10.5.3 zu einer Basis von V erweitern lässt. Somit ist die Kardinalität der Basis (=Dimension) von V größer oder gleich der von U.
(2) Wieder geht man von einer Basis $\{b_1, \dots, b_n\}$ von U aus, die sich nach dem Basisergänzungssatz 10.5.3 zu einer Basis von V erweitern lässt. Ist nun aber auch $\dim(V) = n$, so ist da nichts mehr zu tun, d.h. $\{b_1, \dots, b_n\}$ ist auch schon eine Basis von V. Wenn U, V eine gleiche Basis besitzen, sind sie natürlich gleich. □

- Nach 10.4.12(4) ist die Darstellung eines jeden Vektors $x \in V$ bezüglich einer Basis B von V eindeutig.
- Ist insbesondere V endlichdimensional und $B = \{b_1, \dots, b_n\}$ eine Basis von V, so kann man jedes $v \in V$ darstellen in der Form
$$v = \sum_{i=1}^{n} \nu_i b_i \text{ mit } \nu_1, \dots, \nu_n \in K\,.$$
- Für eine fest vereinbarte Reihenfolge der Basisvektoren ist also jedes $v \in V$ eindeutig festgelegt durch das zugehörige n-Tupel $(\nu_1, \dots, \nu_n) \in K^n$.
- Um die Reihenfolge der Basisvektoren deutlich zu machen, schreibt man auch $B = (b_1, \dots, b_n)$ und spricht von einer *geordneten Basis* von V.

Definition 10.5.11. *Es seien K ein Körper, V ein endlichdimensionaler K-Vektorraum und $B = (b_1, \dots, b_n)$ eine geordnete Basis von V. Ist*
$$v = \sum_{i=1}^{n} \nu_i b_i \textit{ mit } \nu_1, \dots, \nu_n \in K$$
die eindeutige Darstellung eines $v \in V$ bezüglich B, so heißt
$$v_{/B} := \begin{pmatrix} \nu_1 \\ \vdots \\ \nu_n \end{pmatrix} \in K^n$$
die Koordinatendarstellung von v bezüglich B.

- Wählt man sich in einem n-dimensionalen K-Vektorraum eine geordnete Basis B und rechnet man dann mit Koordinatenvektoren bzgl. dieser Basis, so rechnet es sich damit wie im K^n!
- Mit dem Dimensionsbegriff ist man nun in der Lage, den bisher schwammig eingeführten Rang einer Matrix sauber zu definieren.
- Dazu betrachtet man eine beliebige Matrix
$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in K^{m \times n}\,,$$

- Die *Zeilenvektoren*

$$z_1 := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{pmatrix}, \dots, z_m := \begin{pmatrix} a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

  von A erzeugen den so genannten *Zeilenraum* von A

$$\mathrm{Z}(A) := \mathrm{Lin}\,(z_1, \dots, z_m) \subseteq K^n\,.$$

- Die *Spaltenvektoren*

$$s_1 := \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, s_m := \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

  von A erzeugen den so genannten *Spaltenraum* von A

$$\mathrm{S}(A) := \mathrm{Lin}\,(s_1, \dots, s_n) \subseteq K^m\,.$$

- Führt man nun eine der bekannten elementaren Zeilenumformungen

  (i) Vertauschen zweier Zeilen

  (ii) Multiplikation einer Zeile mit einer Konstanten $\alpha \in K \setminus \{0\}$

  (iii) Addition des α-fachen einer Zeile, $\alpha \in K$ beliebig, zu einer anderen.

  an der Matrix aus, so bleiben die Zeilen der Matrix ein Erzeugendensystem des Zeilenraumes $\mathrm{Z}(A)$ (bei Typ (i) und (ii) ist das ohne Beweis klar, bei Typ (iii) eine leichte Übungsaufgabe).

- Somit kann man das Gauß-Verfahren als Übergang von dem ursprünglich gegebenen Erzeugendensystem des Zeilenraumes zu einer Basis des Zeilenraumes interpretieren, denn die Zeilenstufenform am Ende des Umformungsprozesses ist ein minimales Erzeugendensystem und somit eine Basis. Die lineare Unabhängigkeit der verbleibenden Zeilen sieht man auch direkt an den Pivotspalten.

- Bisher war der Rang der Matrix A die Anzahl der Zeilen in einer Zeilenstufenform der Matrix A gewesen, wobei eben nicht klar war, ob diese Zahl überhaupt eindeutig bestimmt ist.

- Mit obiger Überlegung kann man jetzt den Rang einer Matrix sauber definieren:

Definition 10.5.12. *Es seien K ein Körper und $A \in K^{m\times n}$. Der Rang der Matrix A, i.Z.* $\mathrm{Rang}(A)$*, ist*

$$\mathrm{Rang}(A) := \dim(\mathrm{Z}(A))\,.$$

- Die gleichen Überlegungen kann man natürlich auch für den Spaltenraum anstellen.

- Die Anzahl der Spalten in der resultierenden Spaltenstufenform ist dann die Dimension des Spaltenraumes.

- Im Gegensatz zu dem gerade eingeführten *Zeilenrang* der Matrix A ist diese Dimension der so genannte *Spaltenrang* der Matrix A.

- Damit stellt sich die Frage, wie sich Zeilen- und Spaltenrang zueinander verhalten. Diese wird durch den folgenden Satz auf sehr einfache Weise beantwortet, denn die beiden Ränge sind immer gleich.

Satz 10.5.13. *Es seien K ein Körper und $A \in K^{m\times n}$. Dann gilt*

$$\dim(\mathrm{Z}(A)) = \dim(\mathrm{S}(A))\,.$$

Beweis: Wird nachgetragen im Abschnitt über lineare Abbildungen. □

10.6 Affine Teilräume

- Ist V ein K-Vektorraum und $W \subseteq V$ ein Untervektorraum von V, prüft man leicht nach, dass
$$v \sim w :\iff v - w \in W$$
eine Äquivalenzrelation auf V definiert.
- Die Äquivalenzklasse $[v]_\sim$ von v ist $[v]_\sim = v + W$ und wird auch *Nebenklasse* des Untervektorraumes W genannt.
- Im geometrischen Zusammenhang nennt man die Nebenklasse $[v]_\sim$ eines Untervektorraumes W von V auch einen affiner Raum:

Definition 10.6.1. *Eine Teilmenge X eines K-Vektorraumes V heißt affiner Unterraum oder Teilraum, falls es ein $v \in V$ und einen Untervektorraum $W \subseteq V$ gibt, so dass*
$$X = v + W := \{u \in V \mid \exists w \in W : u = v + w\}.$$

Satz 10.6.2. *Es sei $X = v + W$ affiner Teilraum eines K-Vektorraumes V. Dann gilt:*

(1) Für jedes beliebige $v' \in X$ gilt: $X = v' + W$.

(2) Ist $v' \in V$ und $W' \subseteq V$ ein Untervektorraum von V mit $X = v + W = v' + W'$, so gilt $W' = W$ und $v' - v \in W \iff v' \in v + W$.

Beweis: (1) Wegen $v' \in X = v + W$ gibt es ein $w' \in W$ mit $v' = v + w' \iff v = v' - w'$.
$X \subseteq v' + W$: Es sei $x \in X = v + W$. Dann gibt es ein $w \in W$ mit $x = v + w$. Einsetzen von $v = v' - w'$ liefert $x = v' - w' + w = v' + \underbrace{(w - w')}_{\in W} \in v' + W$.
$X \supseteq v' + W$: Es sei $x \in v' + W$. Dann gibt es eine $w \in W$ mit $x = v' + w$. Einsetzen von $v' = v + w'$ liefert $x = v + \underbrace{w' + w}_{\in W} \in v + W$.
(2) $X - X \subseteq W$: Man betrachtet die Differenz $x - x'$ von zwei beliebigen Elementen $x, x' \in X = v + W$. Dann gibt es $w, w' \in W$ mit $x = v + w$ und $x' = v + w'$, d.h. es ist $x - x' = (v + w) - (v + w') = w - w' \in W$. Definiert man
$$X - X := \{x - x' \mid x, x' \in X\},$$
so ist damit gezeigt, dass $X - X \subseteq W$ ist.
$X - X \supseteq W$: Nimmt man anders herum ein $w \in W$, so sieht man sofort, dass $w \in X - X$ ist: man nehme etwa $x = v + w \in X = v + W$ und $x' = v \in X = v + W$. Dies zeigt die andere Inklusion, also
$$W = X - X := \{x - x' \mid x, x' \in X\}.$$
$W = W'$: Macht man die gleiche Überlegung nochmal mit $X = v' + W'$, so erhält man völlig analog $W' = X - X$, also zusammen $W = W'$.
Wegen $W = W'$ gilt jetzt also $X = v + W = v' + W$, insbesondere gilt $v' \in v + W$, also gibt es ein $w \in W$ mit $v' = v + w \iff v' - v = w \in W$. □ 15.11.10

Definition 10.6.3. *Es sei X ein affiner Unterraum eines K-Vektorraumes V. Der eindeutig bestimmte Untervektorraum $\mathrm{D}(X)$ von V mit $X = v + \mathrm{D}(X)$ und $v \in V$ heißt der Differenzraum, Richtungsraum oder Tangentialraum von X. Man setzt*
$$\dim(X) := \dim(\mathrm{D}(X)).$$

Definition 10.6.4. *Es sei K ein Körper. Ein affiner Teilraum X der Dimension 1 eines K-Vektorraumes V heißt Gerade.*

Definition 10.6.5. *Es sei K ein Körper. Ein affiner Teilraum X der Dimension 2 eines K-Vektorraumes V heißt Ebene.*

Definition 10.6.6. *Es seien K ein Körper, V ein K-Vektorraum. Ein affiner Teilraum X der Dimension $n-1$ von V heißt Hyperebene, ein Teilraum der Dimension 0 Punkt*[7].

[7] Diese Punkte sind also identisch mit den Vektoren!

- Aus der Theorie der linearen Gleichungssysteme folgt, dass man eine Hyperebene in einem n-dimensionalen K-Vektorraum V als Lösung einer inhomogenen linearen Gleichung ${}^t a x = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b$ mit $b \in K$ und $a = (a_1, a_2, \ldots, a_n) \in K^n \setminus \{0\}$ beschreiben kann.

- Man beachte, dass diese *Gleichung der Hyperebene* immer einen $n-1$-dimensionalen affinen Teilraum beschreibt; eine Gerade im $\mathbb{R}^3$ bekommt man so z.B. nicht!

Definition 10.6.7. *Es seien K ein Körper und X, X' zwei affine Unterräume eines K-Vektorraumes V. Dann heißen X und X' parallel, in Zeichen $X \parallel X'$, wenn* $\mathrm{D}(X) \subseteq \mathrm{D}(X') \vee \mathrm{D}(X') \subseteq \mathrm{D}(X)$ *ist.*

- Diese Definition ist mit Vorsicht zu genießen: falls die Vektorräume W und W' verschiedene Dimension haben geht nämlich die Transitivität verloren und somit ist die Parallelität keine Äquivalenzrelation mehr.

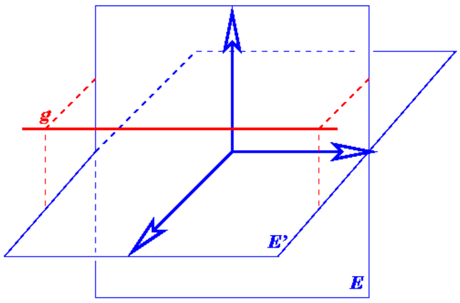


Abbildung 22: Parallele affine Teilräume

11 Lineare und affine Abbildungen

11.1 Homomorphismen

- Aus dem Kapitel über Gruppen sind strukturerhaltenden Abbildungen unter dem Namen „Gruppen-Homomorphismen“ bekannt.
- Da ein Vektorraum insbesondere eine abelsche Gruppe ist, kann man all das zu Gruppenhomomorphismen Gesagte auch auf Vektorräume übertragen (lesen Sie nochmal nach!).
- Da Vektorräume aber wegen des zusätzlichen skalaren Vielfachen noch eine Verknüpfung mehr als Gruppen haben, möchte man Abbildungen untersuchen, die auch diese Struktur erhalten:

Definition 11.1.1. *Es seien K ein Körper und V, W zwei K-Vektorräume. Eine Abbildung $f : V \to W$ heißt linear oder Vektorraum-Homomorphismus, wenn gilt*

(*HOM*$_1$) $f(v_1 + v_2) = f(v_1) + f(v_2)$ *für alle* $v_1, v_2 \in V$, [8]

(*HOM*$_2$) $f(kv) = kf(v)$ *für alle* $v \in V$ *und* $k \in K$.

Die Menge aller linearen Abbildungen von V nach W wird mit $\mathrm{Hom}(V, W)$ *bezeichnet.*

- Die folgenden Beispiele zeigen einige Standard-Homomorphismen und damit insbesondere auch die Existenz linearer Abbildungen.

Beispiel 11.1.2. *Es seien K ein Körper und V, W zwei K-Vektorräume. Die so genannte Nullabbildung ist linear:*

$$f : \begin{cases} V \to W \\ v \mapsto 0 \end{cases} .$$

Beispiel 11.1.3. *Es sei $V = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ der Vektorraum der unendlich oft differenzierbaren Abbildungen f von $\mathbb{R}$ in $\mathbb{R}$. Dann ist*

$$\varphi : \begin{cases} V \to V \\ f \mapsto \frac{d}{dx} f \end{cases}$$

eine lineare Abbildung.

Beispiel 11.1.4. *Es seien S und T Mengen mit $\emptyset \neq S \subset T$. Weiterhin sei V der Vektorraum aller Abbildungen von T in einen Körper K, also $V = K^T$ und analog $W = K^S$. Weiterhin seien $f \in V$, also $f : T \to K$, und*

$$f|_S : \begin{cases} S \to K \\ x \mapsto f(x) \end{cases}$$

die Restriktion von f auf S. Dann ist

$$\varphi : \begin{cases} V \to W \\ f \mapsto f|_S \end{cases}$$

eine lineare Abbildung.

- Da es jetzt wieder um Abbildungen geht, spielen auch wieder die Begriffe injektiv, surjektiv und bijektiv eine Rolle (wiederholen!):

Definition 11.1.5. *Es seien K ein Körper und V, W zwei K-Vektorräume. Ein surjektiver Homomorphismus $f : V \to W$ heißt Epimorphismus.*
Ein injektiver Homomorphismus $f : V \to W$ heißt Monomorphismus.
Ein bijektiver Homomorphismus $f : V \to W$ heißt Isomorphismus.
Ein Homomorphismus $f : V \to V$ heißt Endomorphismus.
Ein Isomorphismus $f : V \to V$ heißt Automorphismus.
Statt $\mathrm{Hom}(V, V)$ *schreibt man kurz* $\mathrm{End}(V)$.

- Wie schon bei Gruppen, ist auch bei Vektorräumen der Kern eines Homomorphismus das Urbild des neutralen Elements:

[8] Das ist gerade die Definition eines Gruppenhomomorphismus!

Definition 11.1.6. $\mathrm{Kern}(f) := f^{-1}(\{0\})$

- Bereits im Abschnitt über Gruppen ist gezeigt worden, dass der Kern eine Untergruppe des Urbildes ist:

Satz 11.1.7. *Es seien K ein Körper, V, W zwei K-Vektorräume und $f \in \mathrm{Hom}(V, W)$. Dann ist $\mathrm{Kern}(f)$ ein Untervektorraum von V.*

Beweis: Leichte Übung (kleine Abwandlung des entsprechenden Beweises 7.4.5 für Gruppen, jetzt mit dem Untervektorraum-Test 10.2.9 statt mit dem Untergruppen-Test 7.4.2). □

Satz 11.1.8. *Es seien K ein Körper, V, W zwei K-Vektorräume und $f \in \mathrm{Hom}(V, W)$. Dann gilt:*
f ist injektiv $\iff$ $\mathrm{Kern}(f) = \{0\}$.

Beweis: Wegen $f(0) = 0$ für jeden Homomorphismus bzw. wegen der Untervektorraum-Eigenschaft von $\mathrm{Kern}(f)$ ist $0 \in \mathrm{Kern}(f)$.

„$\Rightarrow$“ Ist ein $v \neq 0$ in $\mathrm{Kern}(f)$, so ist $f(0) = 0 = f(v)$, was aber ein Widerspruch zur Injektivität ist.

„$\Leftarrow$“ Ist dagegen f nicht injektiv, so gibt es zwei Vektoren $v \neq w \in V$ mit $f(v) = f(w)$. Dann gilt aber wegen der Linearität von f: $f(v) - f(w) = f(v - w) = 0$, d.h. $v - w$ ist ein von 0 veschiedener Vektor im Kern von f, also $\mathrm{Kern}(f) \neq \{0\}$. □

Beispiel 11.1.9. *Es sei $V = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ (vgl. 11.1.3) und*

$$\varphi : \begin{cases} V \to V \\ f \mapsto \frac{d}{dx} f \end{cases} \in \mathrm{End}(V)\,.$$

Wegen $\mathrm{Kern}(f) = \mathbb{R} \neq \{0\}$ ist f nicht injektiv (es sind genau die Konstanten, die die Ableitung 0 haben).
Da man jedes Element von $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ integrieren kann ($\to$ Analysis), besitzt jedes Element ein Urbild bzgl. φ ($g(x) = \int_0^x f(y)\,dy$ ist ein Urbild von $\frac{d}{dx} g(x) = f(x)$), d.h. φ ist surjektiv.
Dies kann nur in unendlich großen Mengen passieren. Es wird noch gezeigt werden, dass sich die Aussage von endlichen Mengen auf endlichdimensionale Vektorräume erweitern lässt.

Satz 11.1.10. *Es seien K ein Körper, $m, n \in \mathbb{N}$ und $A \in K^{m \times n}$. Dann ist*

$$f_A : \begin{cases} K^n \to K^m \\ v \mapsto Av \end{cases} \in \mathrm{Hom}(K^n, K^m)\,.$$

Beweis: Die Behauptung folgt direkt aus den Rechengesetzen für Matrizen:

$$\begin{aligned} f_A(k \cdot v) &= A(k \cdot v) = k \cdot Av = k \cdot f_A(v) \\ f_A(v + w) &= A(v + w) = Av + Aw = f_A(v) + f_A(w) \end{aligned}$$ □

- Mit diesem Satz kennt man eigentlich schon alle linearen Abbildungen zwischen endlichdimensionalen Vektorräumen. Dies wird noch ausführlich hergeleitet werden.

Beispiel 11.1.11. *Eine lineare Abbildung f_A des arithmetischen Vektorraumes $\mathbb{R}^2$ in sich ist etwa durch die Abbildungsmatrix*

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

gegeben.
Ein Vektor $v = \binom{v_1}{v_2}$ wird durch diese Abbildung auf $\binom{-v_2}{v_1}$ abgebildet.
Aus der Schulmathematik weiß man vielleicht noch, dass dieser Vektor die gleiche Länge wie v hat und auf v orthogonal steht.
Deshalb handelt es sich um eine Drehung um $\frac{\pi}{2}$. Die Drehrichtung erkennt man, wenn man etwa das Bild $f_A(e_1) = e_2$ betrachtet, d.h. es geht entgegen dem Uhrzeigersinn.
Wendet man f_A zweimal hintereinander auf einen Vektor v an (also Drehung um π bzw. Spiegelung am Ursprung), so ist dies wieder eine lineare Abbildung mit der Abbildungsmatrix A^2, denn

$$f_A(f_A(v)) = f_A(Av) = A(Av) = A^2 v = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} v = -\mathbb{1}_2 v = -v\,,$$

weshalb man oft auch f_A^2 statt $f_A \circ f_A$ schreibt.
Entsprechend ist f_A^4 die Identität (Drehung um 2π), bzw. $A^4 = \mathbb{1}_2$. [9]

[9] Man nennt solche eine Abbildung f, zu der es ein $k \in \mathbb{N}$ mit $f^k = id$ gibt, *idempotent*.

Satz 11.1.12. *Es seien K ein Körper, V, W zwei K-Vektorräume und $\varphi \in \mathrm{Hom}(V, W)$. Dann gilt*

(1) $M \subseteq V \Rightarrow \varphi(\mathrm{Lin}\,(M)) = \mathrm{Lin}\,(\varphi(M))$

(2) *Ist $U \subseteq V$ ein Untervektorraum von V, so ist $\varphi(U) \subseteq W$ ein Untervektorraum von W mit* $\dim(\varphi(U)) \leq \dim(U)$ *für ein endlichdimensionales U.*

Beweis: (1) Ist $M = \emptyset$, so ist $\mathrm{Lin}\,(M) = \{0\}$, also $\varphi(\mathrm{Lin}\,(M)) = \{0\}$ und $\varphi(M) = \emptyset$, also auch $\mathrm{Lin}\,(\varphi(M)) = \{0\}$. Nun seien also $M \neq \emptyset$ und $m_1, \dots, m_n \in M$. Dann gilt für eine beliebige Linearkombination $\sum_{i=1}^n \lambda_i m_i$ von M:

$$\varphi\left(\sum_{i=1}^n \lambda_i m_i\right) = \sum_{i=1}^n \lambda_i \varphi(m_i)\,,$$

d.h. das Bild einer Linearkombination unter einer linearen Abbildung ist eine Linearkombination der Bilder. Liest man die Gleichung dagegen von rechts nach links, so sieht man dass auch jede Linearkombination von Bildern aus M ein Bild einer Linearkombination aus M ist. Da bekanntlich die Menge aller Linearkombinationen einer Menge identisch mit dem Erzeugnis dieser Menge ist, folgt der erste Teil der Behauptung.
(2) Da U ein Untervektorraum von V ist, gilt $\mathrm{Lin}\,(U) = U$ und somit folgt mit Hilfe des bereits bewiesenen Teils des Satzes, dass

$$\varphi(U) = \varphi(\mathrm{Lin}\,(U)) = \mathrm{Lin}\,(\varphi(U))$$

ein Untervektorraum von W ist. Zur Bestimmung der Dimension betrachtet man eine Basis B von U. Es ist $\dim(U) = |B|$ und

$$\varphi(U) = \varphi(\mathrm{Lin}\,(B)) = \mathrm{Lin}\,(\varphi(B))\ .$$

Somit ist $\dim(\varphi(U)) \leq |\varphi(B)|$, denn $\varphi(B)$ enthält eine Basis von $\varphi(U)$. □

11.2 Rang und Defekt

Definition 11.2.1. *Es seien K ein Körper, V, W zwei K-Vektorräume und $f \in \mathrm{Hom}(V, W)$. Das Bild der linearen Abbildung f ist*

$$\mathrm{Bild}(f) := f(V) \subseteq W$$

und ist nach dem vorhergehenden Satz ein Untervektorraum von W.[10]
Der Rang der linearen Abbildung f ist die Dimension ihres Bildraumes, in Zeichen

$$\mathrm{Rang}(f) := \dim(\mathrm{Bild}(f))\,.$$

- Betrachtet man die lineare Abbildung
$$f_A : \begin{cases} K^n \to K^m \\ v \mapsto Av \end{cases}$$
mit einem $A \in K^{m \times n}$, so folgt aus der Definition des Matrizenprodukts, dass der Bildraum von f_A gerade aus allen Linearkombinationen der Spalten von A besteht.

- Es gilt also
$$\mathrm{Rang}(f_A) = \dim(\mathrm{Bild}(f_A)) = \dim(\mathrm{S}(A)) = \mathrm{Rang}(A)\,.$$

- Der Rang einer Matrix A und der Rang der von ihr vermittelten linearen Abbildung f_A stimmen also überein.

- Da sich noch herausstellen wird, dass sich jede lineare Abbildung zwischen endlichdimensionalen Vektorräumen wie f_A mit Hilfe eines Matrizenprodukts schreiben lässt, ist diese Erkenntnis von großer Tragweite.

- Für eine lineare Abbildung f ist bekanntlich $\mathrm{Kern}(f) := f^{-1}(\{0\})$ ein Untervektorraum von V. Es gilt sogar allgemeiner:

Satz 11.2.2. *Es seien K ein Körper, V, W zwei K-Vektorräume, $f \in \mathrm{Hom}(V, W)$ und $Y \subseteq W$ ein Untervektorraum. Dann ist $f^{-1}(Y) \subseteq V$ ein Untervektorraum von V.*

[10] englisch: $\mathrm{Im}(f)$ vom ‘image’ - nicht mit dem Imaginärteil verwechseln!

Beweis: Es seien $X := f^{-1}(Y)$ und $x, x_1, x_2 \in X$, $k \in K$. Dann gilt $f(x_1), f(x_2) \in Y$ und somit wegen der Linearität von f auch

$$\begin{aligned} f(x_1 + x_2) = f(x_1) + f(x_2) \in Y \Rightarrow x_1 + x_2 \in X \\ f(kx) = kf(x) \in Y \Rightarrow kx \in X\,, \end{aligned}$$

d.h. Y besteht den Untervektorraumtest (damit ist auch der Beweis, dass der Kern ein Untervektorraum ist, nochmal allgemeiner wiederholt). □

Definition 11.2.3. *Es seien K ein Körper, V, W zwei K-Vektorräume und $f \in \operatorname{Hom}(V, W)$. Der Defekt von f ist die Dimension des Kerns, in Zeichen*

$$\operatorname{Def}(f) := \dim(\operatorname{Kern}(f))\,.$$

Satz 11.2.4. *Es seien K ein Körper, V, W zwei K-Vektorräume, $f \in \operatorname{Hom}(V, W)$ und $\dim(V) < \infty$. Dann gilt*

$$\operatorname{Def}(f) + \operatorname{Rang}(f) = \dim(V)\,.$$

Beweis: Nach dem Basisergänzungssatz kann man eine Basis $B' = \{b'_1, \ldots, b'_n\}$ von $\operatorname{Kern}(f)$ zu einer Basis $B = B' \cup B''$ von V mit $B'' = \{b''_1, \ldots, b''_m\}$ ergänzen.
Es gilt $f(B') = \{0\}$ und

$$\operatorname{Lin}(f(B'')) = \operatorname{Lin}(f(B)) = f(\operatorname{Lin}(B)) = f(V)\,.$$

Das dabei verwendete $f(B'')$ ist sogar linear unabhängig, denn

$$0 = \sum_{i=1}^{m} \lambda_i f(b''_i) = f\left(\sum_{i=1}^{m} \lambda_i b''_i\right)$$

heißt $\sum_{i=1}^m \lambda_i b''_i \in \operatorname{Kern}(f)$ und damit folgt

$$\sum_{i=1}^{m} \lambda_i b''_i = \sum_{i=1}^{n} \sigma_i b'_i \iff \sum_{i=1}^{m} \lambda_i b''_i - \sum_{i=1}^{n} \sigma_i b'_i = 0$$

Da $B = B' \cup B''$ eine Basis von V ist, folgt damit $\lambda_i = 0$ für $i = 1, \ldots, m$ (und $\sigma_i = 0$ für $i = 1, \ldots, n$).
Da $f(B'')$ linear unabhängig ist, und wegen $\operatorname{Lin}(f(B'')) = f(V)$, ist $f(B'')$ eine Basis von $f(V) = \operatorname{Bild}(f)$. Die Behauptung kann man jetzt ablesen aus

$$\begin{aligned} \operatorname{Rang}(f) &= \dim(\operatorname{Bild}(f)) = |B''| = m\,, \\ \operatorname{Def}(f) &= \dim(\operatorname{Kern}(f)) = |B'| = n\,, \\ \dim(V) &= |B| = |B' \cup B''| = m + n\,. \end{aligned}$$

□

- Mit diesem Satz ist man nun in der Lage Satz 10.5.13 relativ kurz zu beweisen:

Beweis: Ist $f := f_A : x \mapsto Ax$ die durch A gegebene lineare Abbildung, so ist $\text{Spaltenrang}(A) = \dim(\mathrm{S}(A)) = \dim(\operatorname{Bild}(f))$.
Andererseits ist $\dim(\operatorname{Kern}(f)) = n - \textit{Zeilenrang}(A)$
Mit dem Dimensionssatz $\dim(\operatorname{Kern}(f)) + \dim(\operatorname{Bild}(f)) = n$ folgt nun aber $\dim(\operatorname{Kern}(f)) = n - \dim(\operatorname{Bild}(f))$ und somit $\dim(\operatorname{Bild}(f)) = \text{Zeilenrang}(A) = \text{Spaltenrang}(A)$. □

18.11.10

Satz 11.2.5. *Es seien K ein Körper, V und W zwei K-Vektorräume und B eine Basis von V. Die Abbildung $f \in \operatorname{Hom}(V, W)$ ist*

(1) surjektiv $\iff \operatorname{Lin}(f(B)) = W \overset{\substack{\textit{falls}\\ \dim(W)<\infty}}{\iff} \operatorname{Rang}(f) = \dim(W)$.

(2) injektiv $\iff \operatorname{Kern}(f) = \{0\} \iff \operatorname{Def}(f) = 0 \iff$ *$f|_B$ ist injektiv und $f(B)$ linear unabhängig* $\overset{\substack{\textit{falls}\\ \dim(V)<\infty}}{\iff} \operatorname{Rang}(f) = \dim(V)$.

(3) bijektiv $\iff$ *$f|_B$ ist injektiv und $f(B) \subset W$ ist eine Basis von W* $\overset{\substack{\textit{falls}\\ \dim(V),\dim(W)<\infty}}{\iff} \operatorname{Rang}(f) = \dim(V) = \dim(W)$.

Beweis: (1) $B = \{b_i \mid i \in N\} \Rightarrow f(B) = \{f(b_i) \mid i \in N\}$ [11] und somit

$$\begin{aligned}\mathrm{Lin}\,(f(B)) = &\left\{ \sum_{i\in M} \lambda_i f(b_i) \mid \lambda_i \in K\,, M \subseteq N, M \text{ endlich} \right\} = \\ = &f\left(\left\{ \sum_{i\in M} \lambda_i b_i \mid \lambda_i \in K\,, M \subseteq N, M \text{ endlich} \right\} \right) = \\ = &f(V) \overset{f \text{ surj.}}{=} W\,.\end{aligned}$$

Wegen $\mathrm{Rang}(f) := \dim f(V) \overset{f \text{ surj.}}{=} \dim W$ folgt $\mathrm{Rang}(f) = \dim W$ aus der Surjektivität.
Aus $\dim f(V) = \dim W$ folgt nur im endlichdimensionalen Fall $f(V) = W$ und damit die Surjektivität von f (siehe 11.5.22); im ∞-dimensionalen Fall stimmt das i.A. nicht! (2) wurde teilweise schon in 11.1.8 erledigt, weitere Teile werden in den Übungen gezeigt. $\square$

- Der im letzten Punkt festgestellte Sachverhalt, dass Vektorräume genau dann isomorph sind, wenn sie gleiche Dimension haben, gilt sogar im ∞-dimensionalen Fall (es gibt eine Bijektion zwischen den Basen).

- Ist $V = W$ endlichdimensional, so ist für ein injektives f nach 11.2.5(2) $\mathrm{Kern}(f) = \{0\}$, also nach 11.2.4 $\mathrm{Rang}(f) = \dim(W)$ und somit nach 11.2.5(1) f surjektiv.

- Entsprechend folgt aus f surjektiv, dass für $V = W$ endlichdimensional f auch injektiv ist.

- Es gilt also:

Folgerung 11.2.6. *Es sei f ein Endomorphismus eines endlichdimensionalen Vektorraumes V. Dann gilt: f ist injektiv $\iff$ f ist surjektiv.*

11.3 Matrixdarstellung

- Ein K-Vektorraum V mit $m := \dim(V) < \infty$ lässt sich wie folgt mit dem arithmetischen K-Vektorraum gleicher Dimension in Verbindung bringen:

  – Es sei $B = (b_1, \ldots, b_m)$ eine geordnete Basis von V.

  – Es sei
  $$\varphi_B : \begin{cases} K^m \to V \\ (x_1, \ldots, x_m) \mapsto \sum_{i=1}^{m} x_i b_i \end{cases}$$
  (also insbesondere $\varphi_B(e_j) = b_j$).

  – Das so definierte φ_B ist ein Isomorphismus, d.h. V und K^m sind isomorph. ($\to$ Übung 5.7).

- Die folgenden Generalvoraussetzungen gelten für den gesamten Abschnitt:

  – Es sei V ein endlichdimensionaler K-Vektorraum mit $m := \dim(V)$ und der geordneten Basis $B = (b_1, \ldots, b_m)$.

  – Es sei W ein endlichdimensionaler K-Vektorraum mit $n := \dim(W)$ und der geordneten Basis $C = (c_1, \ldots, c_n)$.

  – Es sei $f \in \mathrm{Hom}(V, W)$.

- Es sei
$$\psi := \varphi_C^{-1} \circ f \circ \varphi_B \quad \text{bzw.} \quad f = \varphi_C \circ \psi \circ \varphi_B^{-1}$$

- Da φ_C (und damit auch φ_C^{-1} $\to$ Übung),f und φ_B linear sind, ist ψ ebenfalls linear ($\to$ Übung).

[11] Ist B und somit die *Indexmenge* N unendlich (möglicherweise sogar überabzählbar), so wählt man ein geeignetes $g \in V^N$ und vereinbart damit $b_i := g(i)$. $(b_i)_{i\in N}$ nennt man eine *Familie* in V und b_i ein Element dieser Familie zum Index i.

- Damit haben wir die folgende Situation:

$$\begin{tikzcd} V \arrow[r, "f"] & W \arrow[d, "\varphi_C^{-1}"] \\ K^m \arrow[u, "\varphi_B"] \arrow[r, "\psi"'] & K^n \end{tikzcd}$$

- Als lineare Abbildung ist ψ durch die Bilder aller Basisvektoren einer geordneten Basis eindeutig festgelegt.
- Nimmt man etwa die kanonische Basis $E^{(m)} = (e_1^{(m)}, \dots, e_m^{(m)})$ von K^m und einen Vektor ${}^t\begin{pmatrix} v_1 & v_2 \dots & v_m \end{pmatrix} = \sum_{\ell=1}^m v_\ell e_\ell^{(m)} \in K^m$, so gilt:
- $\psi({}^t\begin{pmatrix} v_1 & v_2 \dots & v_m \end{pmatrix}) = \psi(\sum_{\ell=1}^{m} v_\ell e_\ell^{(m)}) = \sum_{\ell=1}^{m} v_\ell \psi(e_\ell^{(m)})\,,$
- Dies kann man auch in Matrixschreibweise notieren. Ist

$$A := (\psi(e_1^{(m)}), \psi(e_2^{(m)}), \dots, \psi(e_m^{(m)}))$$

  die Matrix, deren Spalten die Bilder der Basisvektoren sind (geschrieben als Koordinatenvektoren bezüglich der kanonischen Basis $E^{(n)}$ des K^n), so gilt:
- $\psi(\begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}) = A \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$.
- Wegen $f = \varphi_C \circ \psi \circ \varphi_B^{-1}$ bedeutet das also für die ursprünglich betrachtete lineare Abbildung $f \in \operatorname{Hom}(V, W)$ und ein $v \in V$

$$f(v) = \varphi_C \circ \psi \circ \varphi_B^{-1}(v)\,.$$

- Ist nun $v = \sum_{\ell=1}^m v_\ell b_\ell$ die eindeutige Darstellung von v bezüglich der Basis $B = (b_1, \dots, b_m)$ von V, so ist

$$\varphi_B^{-1}(v) = \varphi_B^{-1}(\sum_{\ell=1}^{m} v_\ell b_\ell) = \sum_{\ell=1}^{m} v_\ell e_\ell^{(m)} = {}^t\begin{pmatrix} v_1 & v_2 & \dots & v_m \end{pmatrix} \in K^m\,.$$

- Somit ist nach dem oben Gesagten

$$\psi \circ \varphi_B^{-1}(v) = \psi(\begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}) = A \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} =: \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}\,.$$

- Damit ergibt sich für f insgesamt

$$f(v) = \varphi_C(\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}) = \varphi_C(\sum_{\ell=1}^{n} w_\ell e_\ell^{(n)}) = \sum_{\ell=1}^{n} w_\ell c_\ell =: w$$

- Für $\varphi_B^{-1}(v)$ kann man auch $v_{/B}$ schreiben, denn das ist die bereits früher eingeführte Koordinatendarstellung des Vektors v bezüglich der Basis B.
- Damit sieht die ganze Rechnung so aus:

$$w_{/C} = A \cdot v_{/B}$$

- Die Matrix A sieht dabei mit der neuen Schreibweise so aus

$$A = (f(b_1)_{/C}, f(b_2)_{/C}, \dots, f(b_m)_{/C})\,, \tag{11.1}$$

  d.h. ihre Spalten sind die Bilder der Basisvektoren von V, dargestellt als Koordinatenvektoren bezüglich der Basis von W.

Definition 11.3.1. *Die Matrix $A \in K^{n\times m}$ zu der linearen Abbildung f ist (für festgelegte Basen B, C von V bzw. W) eindeutig bestimmt. Man nennt sie deshalb die Matrixdarstellung von $f \in \mathrm{Hom}(V,W)$ bzgl. der Basen B und C und schreibt $A =: {}_C[f]_B$.*

- Diese Schreibweise ist in der Literatur nicht einheitlich:
- Einige andere Schreibweisen mit der gleichen Bedeutung sind etwa $M(f)^B_C$ oder ${}_CM(f)_B$.
- Die Schreibweise ist sehr einprägsam, was auf der folgenden Seite farblich unterstrichen wird.

V endlichdim. K-Vektorraum	$f \in \mathrm{Hom}(V,W)$	W endlichdim. K-Vektorraum
$B = (b_1,\dots,b_m)$ geordnete Basis	$K^{n\times m} \ni {}_C[f]_B = (f(b_1)_{/C},\dots,(f(b_m)_{/C})$	$C = (c_1,\dots,c_n)$ geordnete Basis
$v = \sum_{i=1}^m v_i b_i \in V$		$w = \sum_{j=1}^n w_j c_j \in W$

$$f(v) = w \iff w_{/C} = {}_C[f]_B \cdot v_{/B}$$

Beispiel 11.3.2. *Es sei*

$$h : \begin{cases} \mathbb{R}^{2\times 2} \to \mathbb{R} \\ \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \mapsto x_{11} + x_{22} \end{cases} \qquad (\to \text{ Übung}).$$

Eine mögliche Basis des $m = 4$-dimensionalen $\mathbb{R}$-Vektorraumes $\mathbb{R}^{2\times 2}$ ist etwa

$$B := \left(\underbrace{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)}_{=b_1}, \underbrace{\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)}_{=b_2}, \underbrace{\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)}_{=b_3}, \underbrace{\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)}_{=b_4} \right)$$

Eine mögliche Basis des $n = 1$-dimensionalen $\mathbb{R}$-Vektorraumes $\mathbb{R}$ ist

$$C := (\underbrace{1}_{=c_1})$$

Zum Aufstellen von ${}_C[h]_B \in \mathbb{R}^{1\times 4}$ werden die Bilder der Basisvektoren gebraucht:

$$h(b_1) = 1\,,\ h(b_2) = 0\,,\ h(b_3) = 0\,,\ h(b_4) = 1\,.$$

Bezüglich der Basis C sind das die Koordinaten-Vektoren

$$h(b_1)_{/C} = (1)\,,\ h(b_2)_{/C} = (0)\,,\ h(b_3)_{/C} = (0)\,,\ h(b_4)_{/C} = (1)\,.$$

Mit diesen Vektoren als Spalten ergibt sich

$${}_C[h]_B = \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix}$$

Möchte man nun also z.B. $h(X)$ für $X = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ mit Hilfe der Matrixdarstellung berechnen, so muss man X zuerst bezüglich der Basis B darstellen.
Wegen $X = b_1 + b_2 + b_3 + b_4$ gilt

$$X_{/B} = \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \\ 1 \end{smallmatrix}\right)$$

und somit

$$y_{/C} = {}_C[h]_B \cdot X_{/B} = \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix} \cdot \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \\ 1 \end{smallmatrix}\right) = (2)$$

Damit ist

$$h(X) = \sum_{j=1}^{1} y_j c_j = y_1 \cdot c_1 = 2 \cdot 1 = 2 = y\,,$$

was man auch direkt aus der Abbildungsvorschrift abliest. 22.11.10

Beispiel 11.3.3. *Es sei*

$$f : \begin{cases} \mathbb{R}[z]_2 \to \mathbb{R}[z]_2 \\ p(z) \mapsto 2\frac{d}{dz}p(z) - z\frac{d^2}{dz^2}p(z) \end{cases}$$

Da die Ableitung linear ist, ist f linear.
Schreibt man $p \in \mathbb{R}[z]_2$ aus als $p = az^2 + bz + c$ mit $a, b, c \in \mathbb{R}$, so ist $f(p(z)) = 2az + 2b$.
Da es sich bei f um einen Endomorphismus handelt, nimmt man üblicherweise die gleiche Basis für Definitions- und Bildbereich. Eine mögliche Basis des $m = 3$-dimensionalen $\mathbb{R}$-Vektorraumes $\mathbb{R}[z]_2$ ist etwa

$$B := \left(\underbrace{z^2}_{=b_1}, \underbrace{z}_{=b_2}, \underbrace{1}_{=b_3}\right)$$

Zum Aufstellen von ${}_B[f]_B \in \mathbb{R}^{3\times 3}$ werden die Bilder der Basisvektoren gebraucht:

$$f(b_1) = 2z\,,\; f(b_2) = 2\,,\; f(b_3) = 0\,,$$

Bezüglich der Basis B sind das die Koordinaten-Vektoren

$$f(b_1)_{/B} = \left(\begin{smallmatrix}0\\2\\0\end{smallmatrix}\right)\,,\; f(b_2)_{/B} = \left(\begin{smallmatrix}0\\0\\2\end{smallmatrix}\right)\,,\; f(b_3)_{/B} = \left(\begin{smallmatrix}0\\0\\0\end{smallmatrix}\right)\,.$$

Mit diesen Vektoren als Spalten ergibt sich

$${}_B[f]_B = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

Möchte man nun also z.B. $f(q)$ für $q(z) = z^2 - 3z + 2$ mit Hilfe der Matrixdarstellung berechnen, so muss man q zuerst bezüglich der Basis B darstellen.
Wegen $q = b_1 - 3b_2 + 2b_3$ gilt

$$q_{/B} = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$$

und somit

$${}_B[f]_B \cdot q_{/B} = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ -6 \end{pmatrix}$$

Damit ist

$$f(q) = \sum_{j=1}^{3} q_j b_j = 2z - 6\,.$$

Aus der Darstellungsmatrix ${}_B[f]_B$ von f kann man nun auch weitere Informationen über f ablesen:
Da offensichtlich $\mathrm{Rang}({}_B[f]_B) = 2$ ist, folgt $\mathrm{Rang}(f) = \dim(\mathrm{Bild}(f)) = 2$. Da die Bildpolynome von f einen Grad ≤ 1 haben, bedeutet das $\mathrm{Bild}(f) = \mathbb{R}[z]_1$.
Dies folgt auch aus

$$\begin{aligned} \mathrm{Bild}(f) = &\mathrm{S}({}_B[f]_B) = \mathrm{Lin}\left(\begin{pmatrix}0\\2\\0\end{pmatrix}, \begin{pmatrix}0\\0\\2\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}\right) = \\ &= \mathrm{Lin}\left(\begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right) = \mathrm{Lin}\,(z, 1) = \mathbb{R}[z]_1 \end{aligned}$$

Ebenfalls wegen $\mathrm{Rang}({}_B[f]_B) = \mathrm{Rang}(f) = 2$ folgt mit dem Dimensionssatz $\mathrm{Def}(f) = \dim(\mathrm{Kern}(f)) = 1$, d.h. f ist weder injektiv noch surjektiv.
Der $\mathrm{Kern}(f)$ ist die Lösung des homogenen LGS

$${}_B[f]_B \cdot x = 0 \iff \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} \cdot x = 0$$

hier also $\mathrm{Kern}(f) = \mathrm{Lin}\,(e_3) = \mathrm{Lin}\,(1) = \mathbb{R}$.
Ändert man die Definition ab auf

$$\tilde{f} : \begin{cases} \mathbb{R}[z]_2 \to \mathbb{R}[z]_1 \\ p(z) \mapsto 2\frac{d}{dz}p(z) - z\frac{d^2}{dz^2}p(z) \end{cases},$$

so erhält man entsprechend

$$_C\left[\tilde{f}\right]_B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

(wobei nun $C = (z, 1)$ *eine Basis von* $\mathbb{R}[z]_1$ *sei) und liest an dem vollen Rang* 2 *der Matrix ab, dass* $\tilde{f}$ *surjektiv ist.*

11.4 Basiswechsel

- Ist $V = W$ und betrachtet man einen Endomorphismus $f \in \mathrm{Hom}(V, V) = \mathrm{End}(V)$, so wird man meist auch nur eine Basis B von V wählen und schreibt dann einfach $[f]_B$ statt $_B[f]_B$.
- Betrachtet man nun einen Endomorphismus $f \in \mathrm{Hom}(V, V) = \mathrm{End}(V)$, aber doch zwei verschiedene Basen $B \neq \tilde{B}$, so ist insbesondere
$$T := {}_{\tilde{B}}[\mathrm{id}]_B$$
interessant.
- Diese quadratische Transformationsmatrix T beschreibt die Umrechnung eines Koordinatenvektors bezüglich der Basis B in einen Koordinatenvektor bezüglich der Basis $\tilde{B}$, d.h. für jedes $v \in V$ gilt
$$v_{/\tilde{B}} = {}_{\tilde{B}}[\mathrm{id}]_B \cdot v_{/B} = T \cdot v_{/B}\,. \tag{11.2}$$
- Betrachtet man nun also einen Endomorphismus $f \in \mathrm{End}(V)$ einmal bezüglich der Basis B von V und einmal bezüglich der Basis $\tilde{B}$ von V und die beiden zugehörigen Abbildungsmatrizen $[f]_B$ und $[f]_{\tilde{B}}$, so kann man zwischen diesen mit Hilfe geeigneter Transformationsmatrizen umrechnen:
- Aus
$$w_{/\tilde{B}} = [f]_{\tilde{B}} \cdot v_{/\tilde{B}} = {}_{\tilde{B}}[f]_{\tilde{B}} \cdot v_{/\tilde{B}}$$
- wird durch Einsetzen von (11.2)
$$w_{/\tilde{B}} = {}_{\tilde{B}}[f]_{\tilde{B}} \cdot {}_{\tilde{B}}[\mathrm{id}]_B \cdot v_{/B}$$
- und durch nochmalige Transformation
$$w_{/B} = {}_B[\mathrm{id}]_{\tilde{B}} \cdot w_{/\tilde{B}} = \underbrace{{}_B[\mathrm{id}]_{\tilde{B}} \cdot {}_{\tilde{B}}[f]_{\tilde{B}} \cdot {}_{\tilde{B}}[\mathrm{id}]_B}_{={}_B[f]_B} \cdot v_{/B} \tag{11.3}$$
- Wegen
$$v_{/\tilde{B}} = {}_{\tilde{B}}[\mathrm{id}]_B \cdot v_{/B} \quad \text{bzw.} \quad v_{/B} = {}_B[\mathrm{id}]_{\tilde{B}} \cdot v_{/\tilde{B}} \qquad \forall v \in V$$
ist $T = {}_{\tilde{B}}[\mathrm{id}]_B$ invertierbar mit $T^{-1} = {}_B[\mathrm{id}]_{\tilde{B}}$.
- Daraus liest man ab
$${}_B[f]_B = {}_B[\mathrm{id}]_{\tilde{B}} \cdot {}_{\tilde{B}}[f]_{\tilde{B}} \cdot {}_{\tilde{B}}[\mathrm{id}]_B = T^{-1} \cdot {}_{\tilde{B}}[f]_{\tilde{B}} \cdot T \tag{11.4}$$

Definition 11.4.1. *Es seien* K *ein Körper und* $n \in \mathbb{N}$. *Matrizen* A, $A' \in K^{n\times n}$ *heißen ähnlich, wenn es ein invertierbares* $T \in K^{n\times n}$ *mit* $A' = TAT^{-1} \iff A = T^{-1}A'T$ *gibt.*

- Die *Einsteinsche Summenkonvention* ist eine Konvention zur Notation mathematischer Ausdrücke.
- Sie wird in der Tensorrechnung und insbesondere in der theoretischen Physik verwendet.
- Mit der *Summenkonvention* werden Summenzeichen zur Verbesserung der Übersicht einfach weggelassen und stattdessen wird über doppelt auftretende Indizes summiert:

- Im einfachsten Fall der Summenkonvention gilt: *Über doppelt auftretende Indizes innerhalb eines Produkts wird summiert.*
- So ist z.B. für $A = (a_{ij}) \in K^{m\times n}$ die i-te Zeile des Matrizenproduktes Ax bekanntlich $(Ax)_i = \sum_{j=1}^n a_{ij}x_j$. Mit der einfachen Summenkonvention könnte man kurz schreiben $(Ax)_i = a_{ij}x_j$.
- Das ist natürlich fehleranfällig, da man wissen muss, von wo bis wo zu summieren ist. Bei $\alpha_{ki}\beta_{\ell k}\gamma_{ij}$ könnte z.B. $\sum_{i=1}^m \sum_{k=1}^n \alpha_{ki}\beta_{\ell k}\gamma_{ij}$ gemeint sein.
- In der Relativitätstheorie gilt als zusätzliche Regel: *Summiert wird nur, wenn der Index sowohl als oberer und als unterer Index auftritt.*
- In diesem Zusammenhang schreibt man dann etwa $(Ax)_i = a^i_j x^j$.
- Ein Vektor x eines endlichdimensionalen K-Vektorraumes V bzgl. einer Basis $B = (b_1, \ldots, b_n)$ von V sieht dann kurz so aus: $x = x^j b_j$ mit $x^j \in K$ für $j = 1, \ldots, n$ bzw. $x_{/B} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$ statt bisher $x = \sum_{j=1}^n x_j b_j$.
- Die Spalten von $T = (t_{ij})_{1\le i,j\le n} = {}_{\tilde{B}}[\mathrm{id}]_B$ sind laut (11.1) gerade die Vektoren $b_{j/\tilde{B}}$, d.h. man kann für $j = 1, \ldots, n$ schreiben
$$b_j = \sum_{i=1}^n t_{ij}\tilde{b}_i \quad \text{oder in Summennotation} \quad \boxed{b_j = t^i_j \tilde{b}_i}.$$
- Betrachtet man dagegen $x_{/\tilde{B}} = {}_{\tilde{B}}[\mathrm{id}]_B \cdot x_{/B} \iff x_{/\tilde{B}} = T \cdot x_{/B}$ so folgt für die Koordinaten x_i von $x_{/B}$ und $\tilde{x}_i$ von $x_{/\tilde{B}}$
$$\tilde{x}_i = \sum_{j=1}^n t_{ij}x_j \quad \text{oder in Summennotation} \quad \boxed{\tilde{x}^i = t^i_j x^j}.$$
- Betrachtet man die umrandeten Gleichungen, so sieht man, dass man beide Male die gleichen t^i_j hat, einmal aber für die Umrechnung $B \to \tilde{B}$, das andere Mal für $\tilde{B} \to B$, d.h. die x^i transformieren sich genau entgegengesetzt=*kontravariant* zu den b_j.
- Noch allgemeiner kann man einen Basiswechsel beschreiben, wenn man wieder von einer linearen Abbildung $f : V \to W$ ausgeht, jetzt aber zwischen zwei Basen B und $\tilde{B}$ von V und C und $\tilde{C}$ von W wechselt.
- Dann gilt
$${}_{\tilde{C}}[f]_{\tilde{B}} = {}_{\tilde{C}}[\mathrm{id}_W]_C \cdot {}_C[f]_B \cdot {}_B[\mathrm{id}_V]_{\tilde{B}}$$
- bzw. mit den Transformationsmatrizen $T := {}_{\tilde{C}}[\mathrm{id}_W]_C$ von C auf $\tilde{C}$ und $S := {}_{\tilde{B}}[\mathrm{id}_V]_B$ von B auf $\tilde{B}$
$${}_{\tilde{C}}[f]_{\tilde{B}} = T \cdot {}_C[f]_B \cdot S^{-1}.$$

Definition 11.4.2. *Es seien K ein Körper und $m, n \in \mathbb{N}$. Matrizen A, $A' \in K^{n\times m}$ heißen äquivalent, wenn es invertierbare Matrizen $T \in K^{n\times n}$, $S \in K^{m\times m}$ mit $A' = TAS^{-1}$ gibt.*

Beispiel 11.4.3. *Wie in 11.3.2 sei*
$$h : \begin{cases} \mathbb{R}^{2\times 2} \to \mathbb{R} \\ \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \mapsto x_{11} + x_{22} \end{cases}$$
Die Abbildungsmatrix bezüglich der Basen B und C ist bereits aus 11.3.2 bekannt. Jetzt soll auf die beiden neuen Basen
$$\tilde{B} := \Big(\underbrace{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)}_{=\tilde{b}_1}, \underbrace{\left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)}_{=\tilde{b}_2}, \underbrace{\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)}_{=\tilde{b}_3}, \underbrace{\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)}_{=\tilde{b}_4} \Big) \quad \text{von} \quad \mathbb{R}^{2\times 2}$$
und
$$\tilde{C} := \big(\underbrace{2}_{=\tilde{c}_1} \big) \quad \text{von} \quad \mathbb{R}$$
übergegangen werden.

Wegen $\tilde{b}_1 = b_1$, $\tilde{b}_2 = b_1 + b_2$, $\tilde{b}_3 = b_1 + b_2 + b_3$ *und* $\tilde{b}_4 = b_1 + b_2 + b_3 + b_4$ *gilt*

$${}_B[\mathrm{id}_{\mathbb{R}^{2\times 2}}]_{\tilde{B}} = S^{-1} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Wegen $\tilde{c}_1 = 2 = 2c_1$ *gilt*

$${}_C[\mathrm{id}_{\mathbb{R}}]_{\tilde{C}} = (2) \quad \textit{bzw.} \quad T = {}_{\tilde{C}}[\mathrm{id}_{\mathbb{R}}]_C = ({}_C[\mathrm{id}_{\mathbb{R}}]_{\tilde{C}})^{-1} = (\frac{1}{2})$$

Damit hat man zusammen

$$\begin{aligned} {}_{\tilde{C}}[h]_{\tilde{B}} =& {}_{\tilde{C}}[\mathrm{id}_W]_C \cdot {}_C[h]_B \cdot {}_B[\mathrm{id}_V]_{\tilde{B}} = T \cdot {}_C[h]_B \cdot S^{-1} = \\ =& (\frac{1}{2})\begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \end{aligned}$$

Damit kann man jetzt das Bild der Matrix $X = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ *nochmal bezüglich der neuen Basen berechnen. Wegen* $X = \tilde{b}_4$ *ist*

$$X_{/\tilde{B}} = \left(\begin{smallmatrix} 0 \\ 0 \\ 0 \\ 1 \end{smallmatrix}\right)$$

und somit

$$y_{/\tilde{C}} = {}_{\tilde{C}}[h]_{\tilde{B}} \cdot X_{/\tilde{B}} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \cdot \left(\begin{smallmatrix} 0 \\ 0 \\ 0 \\ 1 \end{smallmatrix}\right) = (1)\,.$$

Damit ist

$$h(X) = \sum_{j=1}^{1} y_j \tilde{c}_j = y_1 \cdot \tilde{c}_1 = 1 \cdot 2 = 2 = y$$

Beispiel 11.4.4. *Wie in 11.3.3 sei*

$$f : \begin{cases} \mathbb{R}[z]_2 \longrightarrow \mathbb{R}[z]_2 \\ p(z) \mapsto 2\frac{d}{dz}p(z) - z\frac{d^2}{dz^2}p(z) \end{cases}$$

Die Abbildung f *soll nun allerdings bezüglich der neuen Basis*

$$\tilde{B} := \left(\underbrace{4}_{=\tilde{b}_1}, \underbrace{2z}_{=\tilde{b}_2}, \underbrace{z^2}_{=\tilde{b}_3} \right)$$

beschrieben werden.
Wegen $\tilde{b}_1 = 4b_3$, $\tilde{b}_2 = 2b_2$, $\tilde{b}_3 = b_1$ *gilt (vgl.* (11.4)*)*

$${}_B\left[\mathrm{id}_{\mathbb{R}[z]_2}\right]_{\tilde{B}} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 4 & 0 & 0 \end{pmatrix} = T^{-1}\,.$$

Damit hat man zusammen

$$\begin{aligned} [f]_{\tilde{B}} =& {}_{\tilde{B}}[f]_{\tilde{B}} = {}_{\tilde{B}}\left[\mathrm{id}_{\mathbb{R}[z]_2}\right]_B \cdot {}_B[f]_B \cdot {}_B\left[\mathrm{id}_{\mathbb{R}[z]_2}\right]_{\tilde{B}} = \\ =& {}_B\left[\mathrm{id}_{\mathbb{R}[z]_2}\right]_{\tilde{B}}^{-1} \cdot {}_B[f]_B \cdot {}_B\left[\mathrm{id}_{\mathbb{R}[z]_2}\right]_{\tilde{B}} = T \cdot [f]_B \cdot T^{-1} = \\ =& \begin{pmatrix} 0 & 0 & \frac{1}{4} \\ 0 & \frac{1}{2} & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 4 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

11.5 $\mathrm{Hom}(V, W)$ und der Dualraum

Satz 11.5.1. *Es seien K ein Körper, V und W zwei K-Vektorräume.* $\mathrm{Hom}(V, W)$ *ist mit der üblichen Summe und dem skalaren Vielfachen von Abbildungen selbst ein K-Vektorraum. Sind V, W endlichdimensional, so gilt*

$$\dim(\mathrm{Hom}(V, W)) = \dim(V) \cdot \dim(W)\,.$$

Beweis: Sind $f, g \in \mathrm{Hom}(V, W)$ und $v, v_1, v_2 \in V$, $k, k', c \in K$, so setzt man, wie bei Abbildungen üblich

$$\begin{aligned}(f + g)(v) &:= f(v) + g(v)\,,\\ (kf)(v) &:= kf(v)\,.\end{aligned}$$

Diese Abbildungen sind wieder lineare Abbildungen von V nach W, also Elemente von $\mathrm{Hom}(V, W)$. Dies ist viel Schreibarbeit, aber nicht sonderlich schwer und wird in den Übungen erledigt ($\Rightarrow$5.6).
Für den Nachweis der Dimensionsformel im endlichdimensionalen Fall wählt man zwei Basen $B = (b_1, \ldots, b_m)$ und $C = (c_1, \ldots, c_n)$ von V und W und betrachtet die linearen Abbildungen f_{ij} mit $i = 1, \ldots, m$ und $j = 1, \ldots, n$, die wie folgt durch die Bilder der Basisvektoren eindeutig bestimmt sind:

$$f_{ij}(b_i) := c_j\,,\ f_{ij}(b_\ell) := 0 \ \text{ für } \ \ell \neq i\,.$$

Diese Abbildungen sind linear unabhängig:

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n} \lambda_{ij} f_{ij} = 0 &\iff \sum_{j=1}^{n}\sum_{i=1}^{m} \lambda_{ij} f_{ij}(b_k) = 0 \quad \forall 1 \leq k \leq m\\ &\iff \sum_{j=1}^{n} \lambda_{kj} c_j = 0 \quad \forall 1 \leq k \leq m\\ &\Rightarrow \lambda_{kj} = 0 \ \text{ für } \ j = 1, \ldots, n\,,\ k = 1, \ldots, m\,.\end{aligned}$$

Die f_{ij} erzeugen den ganzen Vektorraum: Ist nämlich $f \in \mathrm{Hom}(V, W)$, so ist dieses f bekanntlich durch die Bilder der Basisvektoren eindeutig festgelegt.
Für geeignete Koeffizienten $\mu_{kj} \in K$ kann man die Bilder der Basisvektoren schreiben als $f(b_k) = \sum_{j=1}^{n} \mu_{kj} c_j$. Nutzt man nun wie oben die Definition der f_{ij} aus, so ergibt sich:

$$\begin{aligned}f(b_k) &= \sum_{j=1}^{n} \mu_{kj} c_j = \sum_{j=1}^{n}\sum_{i=1}^{m} \mu_{ij} f_{ij}(b_k) \quad \forall k = 1, \ldots, m\\ \Rightarrow f &= \sum_{i=1}^{m}\sum_{j=1}^{n} \mu_{ij} f_{ij}\,,\end{aligned}$$

d.h. f ist eine Linearkombination der f_{ij}.
Da es $n \cdot m$ verschiedene Basisvektoren f_{ij} sind, ist also $\dim(\mathrm{Hom}(V, W)) = n \cdot m$. □

25.11.10

Definition 11.5.2. *Es seien K ein Körper und V ein K-Vektorraum. $V^* :=$ $\mathrm{Hom}(V, K)$ heißt der Dualraum von V. Die Elemente von V^* heißen Linearformen oder Lineare Funktionale.*

- Eine Linearform wurde bereits ausführlich in 11.3.2 untersucht.
- Ist V endlichdimensional, so betrachtet man nochmal die Basis aus dem Beweis von 11.5.1.
- Zu zwei Basen $B = (b_1, \ldots, b_m)$ von V und $C = (1)$ von K werden die Linearformen f_i mit $i = 1, \ldots, m$ durch die Bilder der Basisvektoren eindeutig bestimmt:

$$f_i(b_i) := 1\,,\ f_i(b_j) := 0 \ \text{ für } \ i \neq j\,.$$

- Wegen dieser Eigenschaft schreibt man b^i statt f_i:

Definition 11.5.3. *Es seien K ein Körper, V ein endlichdimensionaler K-Vektorraum und $B = (b_1, \dots, b_n)$ eine Basis von V. Für $i = 1, \dots, n$ seien $b^i \in V^*$ mit* [12]

$$b^i(b_j) := \delta^i_j = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{sonst} \end{cases} .$$

Satz 11.5.4. *Es seien K ein Körper und V ein endlichdimensionaler K-Vektorraum mit einer Basis $B = (b_1, \dots, b_n)$. Dann ist $B^* := \left(b^1, \dots, b^n\right)$ eine Basis von V^*, die so genannte Dualbasis von B. Insbesondere gilt $\dim V = \dim V^*$, also $V \cong V^*$.*

Beweis: Der Beweis wurde bereits in 11.5.1 erbracht. □

- Da die Dualbasis $B^* := \{b^1, \dots, b^n\}$ mit oberen Indizes geschrieben wurde, verwendet man für die Koordinaten nun untere Indizes, also für $y \in V^*$ etwa $y = \sum_{i=1}^n y_i b^i$ oder kurz $y = y_i b^i$ mit Summenkonvention.
- Mit linearer Fortsetzung von 11.5.3 folgt nun für $x = x^j b_j$:

$$b^i(x) = b^i(x^j b_j) = x^j b^i(b_j) = x^j \delta^i_j = x^i$$

  oder ohne Summenkonvention für $x = \sum_{j=1}^n x^j b_j$:

$$b^i(x) = b^i(\sum_{j=1}^n x^j b_j) = \sum_{j=1}^n x^j b^i(b_j) = \sum_{j=1}^n x^j \delta^i_j = x^i$$

- *Vorsicht*: Für $\dim V = \infty$ bilden die b^i keine Basis von V^*: Die durch $f : b_i \mapsto 1$ für alle $b_i \in B$ gegebene Linearform lässt sich z.B. nicht als Linearkombination von B^* darstellen (die Summe aller b^* ist nicht definiert und erst recht keine Linearkombination)!
- Ist $y \in V^*$ und $B = (b_1, \dots, b_n)$ Basis von V, so ist ${}_{(1)}[y]_B = \begin{pmatrix} \eta_1 & \dots & \eta_n \end{pmatrix} \in K^{1\times n}$.
- Damit gilt für $x \in V$ mit dem Koordinatenvektor $x_{/B} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$, dass $y(x)_{/(1)} = \sum_{j=1}^n \eta_j x^j$ oder kurz $y(x) = \eta_j x^j$.
- Wechselt man auf die Basis $\tilde{B} = \left(\tilde{b}_1, \dots, \tilde{b}_n\right)$ von V, so gilt entsprechend mit $x_{/\tilde{B}} = \begin{pmatrix} \tilde{x}^1 \\ \vdots \\ \tilde{x}^n \end{pmatrix}$ und ${}_{(1)}[y]_{\tilde{B}} = \begin{pmatrix} \tilde{\eta}_1 & \dots & \tilde{\eta}_n \end{pmatrix} \in K^{1\times n}$, dass $y(x)_{/(1)} = \sum_{i=1}^n \tilde{\eta}_i \tilde{x}^i$ oder kurz $y(x) = \tilde{\eta}_i \tilde{x}^i$.
- Zusammen folgt also $\eta_j x^j = \tilde{\eta}_i \tilde{x}^i$, also mit der Transformationsgleichung $\tilde{x}^i = t^i_j x^j$: $\eta_j x^j = \tilde{\eta}_i t^i_j x^j$.
- Da dies für alle $x \in V$ richtig ist, folgt $\eta_j = t^i_j \tilde{\eta}_i$, d.h. die η_j transformieren sich genauso wie die Basis $b_j = t^i_j \tilde{b}_i$, also *kovariant*.

Definition 11.5.5. *Es sei V ein K-Vektorraum mit $n := \dim(V) < \infty$.*
Ein Element von K heißt auch Tensor 0.Stufe.
Ein Tensor 1.Stufe ist eine Größe, deren einfach indizierte Komponenten v^i oder v_i bei einem durch $b_j = t^i_j \tilde{b}_i$ beschriebenen Wechsel zwischen den Basen $B = (b_1, \dots, b_n)$ und $\tilde{B} = \left(\tilde{b}_1, \dots, \tilde{b}_n\right)$ von V ein bestimmtes Transformationsverhalten zeigen, nämlich im kontravarianten Fall $\tilde{v}^i = t^i_j v^j$, im kovarianten Fall $v_j = t^i_j \tilde{v}_i$.

- Vektoren sind *kontravariante Tensoren 1.Stufe.*
- Linearformen sind *kovariante Tensoren 1.Stufe.*
- Im Fall arithmetischer Vektorräume $V = K^m$ und $W = K^n$ mit den kanonischen Basen $E^{(m)}$ und $E^{(n)}$ gilt gerade

$${}_{E^{(n)}}[f_{ij}]_{E^{(m)}} = E_{ji} \in K^{n\times m} .$$

[12] δ^i_j ist das bekannte *Kroneckersymbol* das hier nur für die spätere Verwendung der Summenkonvention etwas anders geschrieben wird.

- Dies zeigt
$$\operatorname{Hom}(K^m, K^n) \cong K^{n\times m}\,.$$
- Der zugehörige Isomorphismus ist die Abbildung einer linearen Abbildung auf ihre darstellende Matrix.
- Dieser Zusammenhang verdeutlicht auch nochmals die hergeleitete Dimensionsformel.
- Ist $V = K^m$ und $W = K$ mit den kanonischen Basen $E^{(m)}$ und (1) gilt gerade
$$_{(1)}\left[b^j\right]_{E^{(m)}} = {}^t(e_j) \in K^{1\times m}\,,$$
- also
$$\operatorname{Hom}(K^m, K) \cong K^{1\times m} \cong K^m\,.$$
- Betrachtet man den Spezialfall von Endomorphismen, also Elementen von $\operatorname{End}(V) = \operatorname{Hom}(V, V)$, so kann man zusätzlich zu „+“ und dem skalaren Vielfachen, mit denen das ein Vektorraum ist, auch noch die Komposition dieser Abbildungen betrachten.
- In dem zu $\operatorname{End}(V)$ isomorphen Vektorraum der quadratischen Matrizen aus $K^{m\times m}$ entspricht die Komposition der Matrizenmultiplikation, d.h. für $f, g \in \operatorname{End}(V)$ gilt (bezüglich einer Basis B von V)
$$[f \circ g]_B = [f]_B \cdot [g]_B\,.$$
(Man beachte: weder $\circ$ noch das Matrizenprodukt sind i.Allg. kommutativ!)
- Mit dieser Multiplikation wird $\operatorname{End}(V)$ bzw. $K^{m\times m}$ sogar zum *Ring* (folgt direkt aus den Rechengesetzen für Matrizen).
- Man spricht vom *Endomorphismenring* des Vektorraumes V, in Zeichen: $(\operatorname{End}(V), +, \circ)$ (Einselement id_V bzw. Einheitsmatrix $\mathbb{1}_m$).

11.6 Translationen

Definition 11.6.1. *Es seien K ein Körper, V ein K-Vektorraum, $a \in V$ und*
$$t_a : \begin{cases} V \to V \\ x \mapsto x + a \end{cases}$$
Dieses t_a heißt Translation oder Verschiebung von V.

Satz 11.6.2. *Es seien K ein Körper, V ein K-Vektorraum. Dann gilt:*
(1) t_a ist eine Permutation von $V \quad \forall a \in V$,
(2) $t_a \circ t_b = t_{a+b} = t_b \circ t_a \quad \forall a, b \in V$,
(3) $(t_a)^{-1} = t_{-a} \quad \forall a \in V$, $\operatorname{id}_V = t_0$,
(4) Ist $T := \{t_a \mid a \in V\}$, so ist $f : V \to T$ mit $f(a) = t_a$ eine bijektive Abbildung mit $f(a + b) = t_{a+b} = t_a \circ t_b$.

Beweis: (1) $t_a(x) = t_a(y) \iff a + x = a + y \iff x = y$, d.h. t_a ist injektiv,
$t_a(x - a) = (x - a) + a = x$, d.h. ein beliebiges $x \in V$ hat das Urbild $x - a$, t_a ist also surjektiv.
Damit ist t_a eine Bijektion von V in sich, also eine Permutation von V.
(2)
$$\begin{aligned} t_a \circ t_b(x) &= t_a(t_b(x)) = t_a(x + b) = (x + b) + a = x + (b + a) = \\ &= \underbrace{x + (a + b)}_{=t_{a+b}(x)} = (x + a) + b = t_b(x + a) = t_b(t_a(x)) = \\ &= t_b \circ t_a(x)\,. \end{aligned}$$
(3) Nach dem vorhergehenden Punkt gilt $(t_{-a} \circ t_a)(x) = t_0(x) = x \; \forall x \in V \Rightarrow t_{-a} = t_a^{-1}$, denn $t_0(x) = x \quad \forall x \in V \iff t_0 = \operatorname{id}_V$.
(4) Es ist $f(a) = f(b) \iff t_a = t_b \iff t_a(x) = t_b(x) \quad \forall x \in V \iff x + b = x + a \quad \forall x \in V \iff a = b$, also f injektiv. Da die Surjektivität direkt aus der Definition klar ist, ist f also bijektiv. Die Eigenschaft $f(a + b) = t_a \circ t_b$ folgt direkt aus der Definition zusammen mit dem oben gezeigten $t_{a+b} = t_a \circ t_b$. □

- Da die Komposition von Abbildungen bekanntlich assoziativ ist, ist wegen 11.6.2(2)-(3) $(T, \circ)$ eine abelsche Gruppe.

- 11.6.2(4) zeigt dann, dass f eine strukturerhaltende Abbildung von $(V,+)$ nach $(T,\circ)$, also ein Gruppenhomomorphismus, ist.
- Da dieser Homomorphismus sogar bijektiv ist, handelt es sich bei f um einen Isomorphismus, d.h. die beiden betrachteten Gruppen sind isomorph: $(V,+)\cong(T,\circ)$.

11.7 Affine Abbildungen

Definition 11.7.1. *Es seien K ein Körper, V und W zwei K-Vektorräume. Eine Abbildung $g:V\to W$ heißt affine Abbildung, falls es eine Translation $t:W\to W$ und eine lineare Abbildung $f:V\to W$ gibt mit $g=t\circ f$.*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "g=t\circ f"'] & W \arrow[d, "t"] \\ & W \end{tikzcd}$$

- Nimmt man $t=t_0=\mathrm{id}_W$, so hat man $g=f$, d.h. alle linearen Abbildungen sind auch affin.
- Ist $V=W$ und nimmt man $f=\mathrm{id}_V$, so sieht man, dass alle Translationen affine Abbildungen von V in sich sind.
- Die Umkehrung dieser beiden Aussagen gilt nicht.

Folgerung 11.7.2. *Aus den Eigenschaften von Homomorphismen folgt sofort, dass die Bilder affiner Teilräume unter affinen Abbildungen wieder affine Teilräume sind!*

Beispiel 11.7.3. *Es sei $g:\mathbb{R}^2\to\mathbb{R}^2$ gegeben durch*

$$g(x)=\begin{pmatrix}3&0\\0&3\end{pmatrix}x+\begin{pmatrix}-4\\6\end{pmatrix}.$$

Dies ist eine affine Abbildung mit

$$t=t_a\,,\ a=\begin{pmatrix}-4\\6\end{pmatrix}\quad\text{und}\quad f=f_A\,,\ A=\begin{pmatrix}3&0\\0&3\end{pmatrix}.$$

Einige Funktionswerte dieser Abbildung sind z.B.

$$g(\begin{pmatrix}2\\3\end{pmatrix})=\begin{pmatrix}2\\15\end{pmatrix}\,,\ g(\begin{pmatrix}2\\-3\end{pmatrix})=\underbrace{\begin{pmatrix}2\\-3\end{pmatrix}}_{:=w}\,,\ g(\begin{pmatrix}4\\2\end{pmatrix})=\begin{pmatrix}8\\12\end{pmatrix}.$$

Dabei fällt auf, dass w von g festgehalten wird.
Zusammen mit $A=3\mathbb{1}_2$ kann man sich so verdeutlichen, was g geometrisch macht:

$$g(x)=Ax+a=A(w+x-w)+a=\underbrace{Aw+a}_{=g(w)=w}+A(x-w)=$$
$$=w+3\mathbb{1}_2(x-w)=w+3(x-w)=W+3\overrightarrow{WX}$$

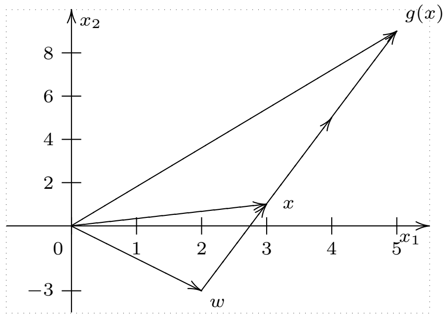


Abbildung 23: Zentrische Streckung

Satz 11.7.4. *Es seien K ein Körper, V und W zwei K-Vektorräume und $g_1, g_2 : V \to W$ zwei affine Abbildungen mit $g_1 = t_a \circ f_1$, $g_2 = t_b \circ f_2$ und $a, b \in W$, $f_1, f_2 \in \mathrm{Hom}(V, W)$. Dann gilt: Ist $g_1(x) = g_2(x)\ \forall x \in V$, so ist $f_1 = f_2$ und $a = b$.*

Beweis: Ist $g_1(x) = g_2(x)\ \forall x \in V$, so ist $(t_a \circ f_1)(x) = (t_b \circ f_2)(x)$. Da t_a invertierbar ist, folgt

$$f_1(x) = (t_a^{-1} \circ t_b \circ f_2)(x) \quad \forall x \in V\,,$$

also insbesondere

$$f_1(0) = t_a^{-1} \circ t_b \circ f_2(0)\,.$$

Da für die linearen Abbildungen f_1 und f_2 gilt
$f_1(0) = 0 = f_2(0)$, folgt daraus

$$0 = t_a^{-1} \circ t_b(0) = t_{-a}(b) = b - a \iff b = a\,,$$

d.h. die Translationsanteile der beiden affinen Abbildungen sind identisch.
Setzt man dies in die ursprüngliche Gleichung ein, so ergibt sich

$$\begin{aligned} t_a \circ f_1(x) = t_b \circ f_2(x) &\iff t_a \circ f_1(x) = t_a \circ f_2(x) \quad \forall x \in V \\ \iff f_1(x) = f_2(x) \quad \forall x \in V &\iff f_1 = f_2\,, \end{aligned}$$

d.h. auch die linearen Anteile sind identisch. □

- Schaltet man zwei affine Abbildungen von V in sich hintereinander, so ergibt sich
- $$\begin{aligned} g_1 \circ g_2(x) &= (t_a \circ f_1) \circ (t_b \circ f_2)(x) = f_1(f_2(x) + b) + a = \\ &= f_1(f_2(x)) + f_1(b) + a = f_1 \circ f_2(x) + (f_1(b) + a) = \\ &= t_c \circ g(x) \end{aligned}$$

  mit $c = f_1(b) + a$ und $g = f_1 \circ f_2$.
- Da mit f_1 und f_2 auch $f_1 \circ f_2$ linear ist, ist $g_1 \circ g_2$ auch wieder eine affine Abbildung.
- Da nicht alle affinen Abbildungen invertierbar sind, bildet die Menge aller affinen Abbildungen von V in sich trotz der gezeigten Eigenschaften sicher keine Gruppe.
- Betrachtet man aber nur die invertierbaren affinen Abbildungen von V in sich, so bildet diese Menge zusammen mit der Komposition von Abbildungen eine Gruppe:
  - Komposition ist immer assoziativ,
  - Hintereinanderausführung bijektiver Abbildungen ist bijektiv,
  - Identität ist eine affine Abbildung

Definition 11.7.5. *Es seien K ein Körper und V ein K-Vektorraum. Eine bijektive affine Abbildung von V in sich heißt Affinität. Die Gruppe aller Affinitäten mit der Komposition wird mit $\mathrm{Aff}(V)$ bzw. $(\mathrm{Aff}(V), \circ)$ bezeichnet (auch $\mathrm{GA}(V)$ für affine Gruppe).*

- Da Translationen grundsätzlich invertierbar sind, ist eine affine Abbildung von V in sich genau dann eine Affinität, wenn ihr linearer Anteil invertierbar ist.
- Die affine Abbildung (zentrische Streckung) aus 11.7.3 ist ein Element von $\mathrm{Aff}(\mathbb{R}^2)$, denn der lineare Anteil ist das 3-fache der Identität.

Satz 11.7.6. *Es seien V und W zwei K-Vektorräume und $g : V \to W$ eine affine Abbildung. Sind X und X' parallele affine Teilräume von V, so sind $g(X)$ und $g(X')$ parallele Teilräume von W.*

Beweis: Zu den affinen Teilräume X und X' von V gibt es $a, a' \in V$ und Untervektorräume U und U' von V mit $X = a + U$ und $X' = a' + U'$. $X \| X'$ bedeutet $U \subseteq U'$ oder $U' \subseteq U$. Der Einfachheit halber seien die Räume so bezeichnet, dass $U \subseteq U'$ gilt.
Die Bilder der beiden affinen Räume unter der affinen Abbildung $g = t_b \circ f$ sind $g(X) = (f(a) + b) + f(U)$ und $g(X') = (f(a') + b) + f(U')$. Dies sind affine Teilräume von W, denn $f(a) + b, f(a') + b \in W$ und $f(U)$ und $f(U')$ sind nach 11.1.12 Untervektorräume von W.
Da U ein Untervektorraum von U' ist, ist auch $f(U)$ ein Untervektorraum von $f(U')$, d.h. auch die affinen Räume $f(X)$ und $f(X')$ sind parallel. □

- Vorsicht: Machen Sie keine falschen Umkehrungen dieses Satzes! Nichtparallele affine Räume können durch
29.11.10 eine affine Abbildung parallel werden.

12 Determinante und Spur

12.1 Motivation

- In den Übungen (→2.4 & 7.5) wurde die Abbildung

$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \end{cases} \qquad a, b, c, d \in \mathbb{R}$$

  untersucht.
- Dies ist eine allgemeine lineare Abbildung $f_A \in \operatorname{End}(\mathbb{R}^2)$ mit $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$.
- Es wird gezeigt, dass f_A (und damit A) genau dann invertierbar ist, wenn $ad - bc \neq 0$ ist.
- Schreibt man für die spätere Verallgemeinerung um auf $A = \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right)$, so heißt dies:

$$A \text{ invertierbar} \iff a_{11} \cdot a_{22} - a_{21} \cdot a_{12} \neq 0\,.$$

- Betrachtet man nun also die Abbildung

$$D : A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \to a_{11} \cdot a_{22} - a_{21} \cdot a_{12}\,,$$

  so weiß man demnach, dass $D(A) = 0$ genau dann der Fall ist, wenn A singulär[13] ist.
- Da A genau dann singulär ist, wenn der Rang der Matrix A nicht voll ist, d.h. wenn die Zeilen oder Spalten von A linear abhängig sind, so kann man die Abbildung D auch interpretieren als Abbildung $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ der Spalten von A (oder analog für die Zeilen von A), die je nach linearer Abhängigkeit oder Unabhängigkeit Werte $= 0$ oder $\neq 0$ liefert, also etwa
- $$D : \Big(\underbrace{\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}}_{=b}, \underbrace{\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}}_{=d} \Big) \to a_{11} \cdot a_{22} - a_{21} \cdot a_{12} = b_1 \cdot d_2 - b_2 \cdot d_1\,.$$
- Die Abbildung D ist linear in allen Komponenten, d.h. in allen Spalten (oder analog Zeilen) von A:
- Die Linearität in der 1. Spalte sieht man etwa, indem man die 2. Spalte festhält, und in der 1. Spalte den üblichen Linearitätstest macht:
- $$\begin{aligned} D(b + c, d) &= (b_1 + c_1) \cdot d_2 - (b_2 + c_2) \cdot d_1 = \\ &= b_1 \cdot d_2 - b_2 \cdot d_1 + c_1 \cdot d_2 - c_2 \cdot d_1 = D(b, d) + D(c, d) \\ D(\lambda b, d) &= \lambda b_1 \cdot d_2 - \lambda b_2 \cdot d_1 = \lambda(b_1 \cdot d_2 - b_2 \cdot d_1) = \lambda D(b, d) \end{aligned}$$

  für alle $b, c, d \in \mathbb{R}^2$, $\lambda \in \mathbb{R}$.
- Analog zeigt man für die 2.Komponente von D (also 2.Spalte von A)

$$\begin{aligned} D(b, d + e) =& D(b, d) + D(b, e) \\ D(b, \sigma d) =& \sigma D(b, d) \end{aligned}$$

  für alle $b, d, e \in \mathbb{R}^2$, $\sigma \in \mathbb{R}$.

Definition 12.1.1. *Es seien K ein Körper, V ein K-Vektorraum und $n \in \mathbb{N}$. Eine Abbildung $V^n \to K$, die in jeder Komponente linear ist, heißt n-Linearform (Bilinearform für $n = 2$) oder allgemeiner Multilinearform.*

Beispiel 12.1.2. *Die oben diskutierte Abbildung $D : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, ist eine Bilinearform.*

- Vertauscht man die Komponenten von $D(b, d) = b_1 d_2 - b_2 d_1$, so erhält man $D(d, b) = b_2 d_1 - b_1 d_2 = -D(b, d)$. Man spricht deshalb von einer *alternierenden Bilinearform.*

[13] singulär=nicht invertierbar

- Das Kreuzprodukt (→später)
$$\begin{pmatrix} a_{11} \\ a_{21} \\ 0 \end{pmatrix} \times \begin{pmatrix} a_{12} \\ a_{22} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ a_{11}a_{22}-a_{21}a_{12} \end{pmatrix},$$
im $\mathbb{R}^3$ ist ein Vektor der Länge (→später)
$$a_{11}a_{22} - a_{21}a_{12} = D\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right).$$
- Dies ist gerade die Fläche des von
$$\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \text{ und } \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$
aufgespannten Parallelogramms.
- D heißt *Determinante* von A und obige Beispiele haben schon mal einige wesentlich Eigenschaften der Determinante angedeutet.

12.2 Determinante

Definition 12.2.1. *Es seien R ein kommutativer Ring mit 1, $n \in \mathbb{N}$ und $A = (a_{ik})_{1 \leq i,k \leq n} \in R^{n \times n}$ eine quadratische Matrix. Dann heißt*
$$\det(A) := \sum_{\pi \in S_n} \left(\operatorname{sgn}(\pi) \prod_{i=1}^{n} (a_{i\pi(i)}) \right) \in R$$
die Determinante von A. Diese Formel ist auch unter dem Namen Leibniz-Regel bekannt (nach Gottfried Wilhelm von Leibniz (1646 - 1716)). Oft schreibt man $|A|$ für $\det(A)$; bei ausgeschriebenen Matrizen werden die runden Klammern durch senkrechte Striche ersetzt.

Beispiel 12.2.2.

$n = 1$
$$\det(A) = \det(a_{11}) = a_{11}.$$

$n = 2$

Nr.	π	*#Fehlst.*	$\operatorname{sgn}(\pi)$	$\prod_{i=1}^{n} a_{i,\pi(i)}$
1	$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$	0	$+$	$a_{11}a_{22}$
2	$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$	1	$-$	$a_{12}a_{21}$

$$\Rightarrow \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \qquad (\text{Sarrus } 2 \times 2)$$

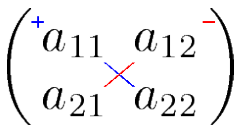


Abbildung 24: Merkregel Sarrus 2×2

$n = 3$

Nr.	π	*#Fehlst.*	$\operatorname{sgn}(\pi)$	$\prod_{i=1}^{n} a_{i,\pi(i)}$
1	$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$	0	$+$	$a_{11}a_{22}a_{33}$
2	$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$	1	$-$	$a_{11}a_{23}a_{32}$
3	$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$	1	$-$	$a_{12}a_{21}a_{33}$
4	$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$	2	$+$	$a_{12}a_{23}a_{31}$
5	$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$	2	$+$	$a_{13}a_{21}a_{32}$
6	$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$	3	$-$	$a_{13}a_{22}a_{31}$

$$\Rightarrow \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = + a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}$$

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

Abbildung 25: Merkregel Sarrus 3×3

$n \geq 4$
Im Fall $n = 4$ ist $|S_n| = n! = 24$, es gibt aber nur 8 Diagonalen in der Matrix A („Diagonalen“ wie oben im Fall $n = 2$ oder $n = 3$ gemeint). Schon aus dieser Überlegung folgt, dass eine Sarrus-Regel wie für $n = 2$ oder $n = 3$ im Fall $n = 4$ nicht existiert. Auch für alle größeren $n \in \mathbb{N}$ ist $n! \neq 2n$ und somit eine Berechnung der Determinante à la Sarrus nicht möglich.
Die Berechnung größerer Determinanten mit der Leibniz-Regel wäre auch wegen der $n!$ Summanden sehr mühsam. Dies zeigt, dass man insbesondere für $n > 3$ ein besseres Verfahren braucht.

12.3 Einfache Rechenregeln für die Determinante

Satz 12.3.1. *Es seien R ein kommutativer Ring mit 1, $n \in \mathbb{N}$ und $A = (a_{ij})_{1 \leq i,j \leq n} \in R^{n \times n}$. Dann gilt*

$$\det {}^tA = \det A\,.$$

Beweis: Laut Definition gilt

$$\begin{aligned} \det {}^tA &= \sum_{\pi \in S_n} \left(\operatorname{sgn}(\pi) \prod_{i=1}^{n} (a_{\pi(i)i}) \right) = \sum_{\pi \in S_n} \left(\operatorname{sgn}(\pi) \prod_{j=1}^{n} (a_{j\pi^{-1}(j)}) \right) = \\ &= \sum_{\tau \in S_n} \left(\operatorname{sgn}(\tau^{-1}) \prod_{j=1}^{n} (a_{j\tau(j)}) \right) = \sum_{\tau \in S_n} \left(\operatorname{sgn}(\tau) \prod_{j=1}^{n} (a_{j\tau(j)}) \right) = \\ &= \det A \end{aligned}$$

$\square$

Satz 12.3.2. *Es seien R ein kommutativer Ring mit 1, $n \in \mathbb{N}$, $A = (a_{ij})_{1 \leq i,j \leq n} \in R^{n \times n}$, $\sigma \in S_n$ und $b_{ij} := a_{i\sigma(j)}$ für $1 \leq i,j \leq n$, also $B = (b_{ij})_{1 \leq i,j \leq n} \in R^{n \times n}$ die quadratische Matrix, die aus A durch die Spaltenpermutation σ entsteht. Dann gilt*

$$\det B = \operatorname{sgn}(\sigma) \det A\,.$$

Beweis:

$$\begin{aligned} \det B &= \sum_{\tau \in S_n} \left(\operatorname{sgn}(\tau) \prod_{i=1}^{n} (b_{i\tau(i)}) \right) = \sum_{\tau \in S_n} \left(\operatorname{sgn}(\tau) \prod_{i=1}^{n} (a_{i\sigma(\tau(i))}) \right) = \\ &= \sum_{\rho \in S_n} \left(\operatorname{sgn}(\sigma^{-1} \circ \rho) \prod_{i=1}^{n} (a_{i\rho(i)}) \right) = \qquad (\rho := \sigma \circ \tau) \\ &= \operatorname{sgn}(\sigma^{-1}) \det A = \operatorname{sgn}(\sigma) \det A\,. \end{aligned}$$

$\square$

Satz 12.3.3. *Es seien R ein kommutativer Ring mit 1, $n \in \mathbb{N}$, $\ell \in \mathbb{N}$ mit $1 \leq \ell \leq n$ und v'_ℓ und $v_1, \ldots, v_\ell, \ldots, v_n \in R^n$ Spaltenvektoren. Dann gilt*

$$\det(v_1, ..., v_\ell + v'_\ell, ..., v_n) = \det(v_1, ..., v_\ell, ..., v_n) + \det(v_1, ..., v'_\ell, ..., v_n)\,.$$

Beweis: Schreibt man $v_i = {}^t\begin{pmatrix} a_{1i} & a_{2i} & \ldots & a_{ni} \end{pmatrix}$ für $i = 1, \ldots, n$ und $v'_\ell = {}^t\begin{pmatrix} a'_{1\ell} & a'_{2\ell} & \ldots & a'_{n\ell} \end{pmatrix}$ so gilt

$$\begin{aligned} \det(v_1, ..., v_\ell + v'_\ell, ..., v_n) &= \sum_{\pi \in S_n} \left(\operatorname{sgn}(\pi)(a_{\ell\pi(\ell)} + a'_{\ell\pi(\ell)}) \prod_{\substack{i=1 \\ i \neq \ell}}^{n} a_{i\pi(i)} \right) = \\ &= \sum_{\pi \in S_n} \left(\operatorname{sgn}(\pi) a_{\ell\pi(\ell)} \prod_{\substack{i=1 \\ i \neq \ell}}^{n} a_{i\pi(i)} \right) + \sum_{\pi \in S_n} \left(\operatorname{sgn}(\pi) a'_{\ell\pi(\ell)} \prod_{\substack{i=1 \\ i \neq \ell}}^{n} a_{i\pi(i)} \right) = \\ &= \det(v_1, ..., v_\ell, ..., v_n) + \det(v_1, ..., v'_\ell, ..., v_n)\,. \end{aligned}$$

$\square$

Satz 12.3.4. *Es seien R ein kommutativer Ring mit 1, $n \in \mathbb{N}$, $\ell \in \mathbb{N}$ mit $1 \le \ell \le n$, $r \in R$ und $v_1, \ldots, v_\ell, \ldots, v_n \in R^n$ Spaltenvektoren. Dann gilt*

$$\det(v_1, ..., r \cdot v_\ell, ..., v_n) = r \cdot \det(v_1, ..., v_\ell, ..., v_n)\,.$$

Beweis: Schreibt man $v_i = {}^t\!\begin{pmatrix} a_{1i} & a_{2i} & \ldots & a_{ni} \end{pmatrix}$ für $i = 1, \ldots, n$, so gilt

$$\begin{aligned}\det(v_1, ..., r \cdot v_\ell, ..., v_n) = \sum_{\pi \in S_n} \left(\operatorname{sgn}(\pi) \cdot r \cdot a_{\ell\pi(\ell)} \prod_{\substack{i=1 \\ i \neq \ell}}^{n} a_{i\pi(i)} \right) = \\ = r \cdot \sum_{\pi \in S_n} \left(\operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i\pi(i)} \right) = r \cdot \det(v_1, ..., v_\ell, ..., v_n)\,.\end{aligned}$$

□

- Arbeitet man über einem Grundkörper K (statt über dem Ring R), so zeigen die Sätze 12.3.3 und 12.3.4, dass die Determinante als Abbildung $K^n \times \cdots \times K^n \to K$ in jeder Komponente linear ist, d.h. es handelt sich um eine Multilinearform.
- Mit 12.3.1 folgt, dass diese Aussagen jeweils auch zeilenweise gelten.
- Multipliziert man jede Spalte einer quadratischen Matrix $A \in R^{n \times n}$ mit einem Faktor $r \in R$, so folgt mit 12.3.4: $\det(rA) = r^n \det(A)$.
- Aus der Multilinearität folgt, dass die Determinante einer Matrix mit einer Nullzeile oder -spalte verschwindet.

Satz 12.3.5. *Es seien R ein kommutativer Ring mit 1, $n \in \mathbb{N}$, $A = (a_{ij})_{1 \le i,j \le n} \in R^{n \times n}$. Hat A zwei identische Spalten (oder Zeilen), so ist* $\det A = 0$.

Beweis: Angenommen, die Spalten i und j von A sind gleich. Vertauscht man diese beiden Spalten, so bleibt die Matrix, und damit auch ihre Determinante, gleich. Andererseits gilt nach 12.3.2, dass sich bei dieser Vertauschung das Vorzeichen der Determinante um $\operatorname{sgn}\left(\begin{smallmatrix} 1 & \ldots & i & \ldots & j & \ldots & n \\ 1 & \ldots & j & \ldots & i & \ldots & n \end{smallmatrix}\right) = -1$ ändert.
Vergleich der beiden Aussagen liefert $\det A = -\det A$, also $2 \det A = 0$. Ist R ein Körper mit $2 \neq 0$ (also etwa $\mathbb{Q}$ oder $\mathbb{R}$), so liest man daraus $\det A = 0$ ab. In Körpern mit $2 = 0$, etwa in $\mathbb{Z}_2$, oder in Ringen mit Nullteilern, etwa $\mathbb{Z}_6$, kann man so nicht argumentieren. Auch dort gilt der Satz aber, man muss nur etwas mehr Mathematik zum Beweis verwenden [siehe z.B. Fischer, 2000, , S.194]. Die Aussage für Zeilen folgt daraus mit 12.3.1. □

Satz 12.3.6. *Es seien R ein kommutativer Ring mit 1, $n \in \mathbb{N}$, $\ell, m \in \mathbb{N}$ mit $1 \le \ell \neq m \le n$, $r \in R$ und $v_1, \ldots, v_\ell, \ldots, v_m, \ldots, v_n \in R^n$ Spaltenvektoren. Dann gilt*

$$\det(v_1, \ldots, v_\ell, \ldots, v_m, \ldots, v_n) = \det(v_1, \ldots, v_\ell + r v_m, \ldots, v_m, \ldots, v_n)\,.$$

Beweis: Wegen 12.3.3 und 12.3.4 (also der Multilinearität, falls R ein Körper ist) gilt

$$\begin{aligned}&\det(v_1, \ldots, v_\ell + r v_m, \ldots, v_m, \ldots, v_n) = \\ &\det(v_1, \ldots, v_\ell, \ldots, v_m, \ldots, v_n) + r \det(v_1, \ldots, v_m, \ldots, v_m, \ldots, v_n)\,.\end{aligned}$$

Die letzte Determinante verschwindet aber nach 12.3.5. □

- Mit 12.3.2, 12.3.4, und 12.3.6 kann man nun zusammenfassend die Wirkung elementarer Umformungen auf Determinanten auflisten:

	Elementare Umformung	*Effekt auf Determinante*
Typ I:	Vertauschen zweier Zeilen (oder Spalten)	Ändert das Vorzeichen
Typ II:	Multiplikation einer Zeile (oder Spalte) mit einer Konstanten $\alpha \in R$	Wert der Determinante mal α
Typ III:	Addition des α-fachen einer Zeile (Spalte), $\alpha \in R$ beliebig, zu einer anderen	Wert der Determinante bleibt gleich

Beispiel 12.3.7. Vorsicht: *Kombinationen dieser elementaren Zeilenumformungen, die man insbesondere gerne benutzt um Bruchrechnungen zu vermeiden, führen bei Determinanten oft zu Fehlern:*

$$\begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} \overset{Z_2 \leftarrow 2Z_2 - 3Z_1}{\rightsquigarrow} \begin{pmatrix} 2 & 4 \\ 0 & -10 \end{pmatrix}$$

ist eine korrekte Kombination aus den beiden elementaren Zeilenumformungen $Z_2 \leftarrow 2Z_2$ und $Z_2 \leftarrow Z_2 - 3Z_1$. Wendet man dies auf eine Determinante an, so übersieht man leicht den Faktor 2 bei der ersten Umformung, der den Wert der Determinante verändert:

$$\begin{vmatrix} 2 & 4 \\ 3 & 1 \end{vmatrix} \underset{12.3.4}{\overset{Z_2 \leftarrow 2Z_2}{=}} \frac{1}{2} \begin{vmatrix} 2 & 4 \\ 6 & 2 \end{vmatrix} \underset{12.3.6}{\overset{Z_2 \leftarrow Z_2 - 3Z_1}{=}} \frac{1}{2} \begin{vmatrix} 2 & 4 \\ 0 & -10 \end{vmatrix}$$

Beispiel 12.3.8. *Ein weiterer Stolperstein ist die verschiedene Verwendung skalarer Vielfacher bei Matrizen und Determinanten. Es gilt:*

$$2 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix},$$

aber

$$2 \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 4 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 6 & 8 \end{vmatrix} = \begin{vmatrix} 2 & 2 \\ 6 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 4 \\ 3 & 8 \end{vmatrix}$$

bzw.

$$4 \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 4 \\ 6 & 8 \end{vmatrix}.$$

- Was hat man nun davon, wenn man eine Determinante durch elementare Zeilenumformungen auf Zeilenstufenform gebracht hat (Spalten analog, das wird im Folgenden nicht jeweils extra dazugesagt)?
- Entsteht eine Nullzeile, so ist die Determinante$= 0$, denn die Zeilen sind damit linear abhängig.
- Entsteht keine Nullzeile oder -spalte, so ist die Zeilen- oder Spaltenstufenform eine Dreiecksmatrix.
- Die Determinante solch einer Dreiecksmatrix ist einfach das Produkt der Diagonalelemente. Dies sieht man an der Leibniz-Formel:
$$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i,\pi(i)} \tag{Leibniz}$$
- Ist nämlich $\pi \neq \mathrm{id}$, so gibt es mindestens ein i mit $1 \leq i \leq n$ und $\pi(i) < i$.
- Da das zugehörige $a_{i,\pi(i)}$ bei einer oberen Dreiecksmatrix $= 0$ ist, ist der gesamte Summand $\operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i,\pi(i)} = 0$.
- Von der gesamten Formel bleibt also nur der Summand, der zu $\pi = \mathrm{id}$ gehört:
$$\det(A) = \prod_{i=1}^{n} a_{i,i}.$$

Beispiel 12.3.9.

$$\begin{vmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 10 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{vmatrix} = 3$$

- Wir haben bereits gesehen, dass die Determinante einer Matrix mit einer Nullzeile oder -spalte $= 0$ ist.
- Weiterhin wurde bereits gezeigt, dass die Determinante $= 0$ ist, wenn die Matrix zwei gleiche Zeilen oder Spalten enthält.
- Mit der nun bekannten Wirkung elementarer Umformungen an quadratischen Matrizen auf deren Determinanten, kann man diese Aussagen nun noch deutlich verallgemeinern:

Satz 12.3.10. *Es seien K ein Körper, $n \in \mathbb{N}$ und $A \in K^{n \times n}$. Dann gilt:*

$$\det(A) = 0 \iff A \text{ ist nicht invertierbar (=singulär).}$$

Beispiel 12.3.11. *A ist genau dann singulär, wenn* $\operatorname{Rang} A < n$ *ist.*
Dies ist wiederum genau dann der Fall, wenn der von den Zeilen (oder Spalten) der Matrix A aufgespannte Raum eine Dimension $< n$ hat.
D.h. A ist genau dann singulär, wenn die Zeilen (oder Spalten) von A linear abhängig sind.
Ist dies der Fall, so lässt sich mit Hilfe elementarer Zeilenumformungen eine Nullzeile (bzw. mit Hilfe elementarer Spaltenumformungen eine Nullspalte) herstellen.
Ist A dagegen regulär, [14] *so ist* $\operatorname{Rang} A = n$*, d.h. A lässt sich mit Hilfe elementarer Umformungen auf Dreiecksgestalt mit Diagonaleinträgen $\neq 0$ bringen.*
Die Determinante der resultierenden Matrix ist also $\neq 0$. Da die elementaren Umformungen nur das Vorzeichen ändern oder mit einem Faktor $\neq 0$ multiplizieren, heißt das, dass in diesem Fall auch $\det A \neq 0$ *ist.*

12.4 Der Laplacesche Entwicklungsssatz

Definition 12.4.1. *Es seien R ein kommutativer Ring mit 1, $n \in \mathbb{N}$ und $A \in R^{n\times n}$.*
Für $i, j \in \mathbb{N}$ mit $1 \leq i, j \leq n$ sei $A^{(ij)} \in R^{(n-1)\times(n-1)}$ diejenige Matrix, die durch Streichen der i-ten Zeile und j-ten Spalte aus A entsteht.
Der Eintrag in der k-ten Zeile und ℓ-ten Spalte dieser Matrix wird mit $(A^{(ij)})_{k,\ell}$ bezeichnet. $A^{(ij)}$ heißt Streichungsmatrix zum Eintrag $a_{ij} = (A)_{ij}$ von A, $\det(A^{(ij)})$ *Minor zu a_{ij}.*
Der Faktor $(-1)^{i+j}\det(A^{(ij)})$ *heißt der Kofaktor oder das algebraische Komplement von a_{ij}.*

Beispiel 12.4.2. *Für* $A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in \mathbb{Z}^{3\times 3}$ *ist* $A^{(21)} = \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} \in \mathbb{Z}^{2\times 2}$

- Es sei $\rho_\ell \in S_n$ für $1 \leq \ell \leq n$ diejenige Permutation, die durch
$$\rho_\ell(k) = \begin{cases} k & \text{für } k < \ell \\ k+1 & \text{für } \ell \leq k < n \\ \ell & \text{für } k = n \end{cases}$$
definiert ist.
- In der üblichen Permutationsschreibweise ist das
$$\rho_\ell = \begin{pmatrix} 1 & 2 & \dots & \ell-1 & \ell & \ell+1 & \dots & n-1 & n \\ 1 & 2 & \dots & \ell-1 & \ell+1 & \ell+2 & \dots & n & \ell \end{pmatrix}$$
- ρ_ℓ besitzt die Fehlstände $(\ell, n), (\ell+1, n), \dots, (n-1, n)$.
- Da das $n - \ell$ Stück sind, gilt $\operatorname{sgn}(\rho_\ell) = (-1)^{n-\ell}$.
- Mit solchen Permutationen lässt sich der Übergang von A zu der Streichungsmatrix zu einem Eintrag beschreiben. Es gilt:
$$(A^{(ij)})_{k,\ell} = (A)_{\rho_i(k),\rho_j(\ell)} = a_{\rho_i(k),\rho_j(\ell)} \quad \text{für} \quad 1 \leq k, \ell \leq n-1\,.$$

6.12.10

Satz 12.4.3. *Es sei $A \in K^{n\times n}$ und $1 \leq i \leq n$. Dann gilt:*
$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A^{(ij)})\,.$$

Beweis: Ist $\pi \in S_n$ eine Permutation mit $\pi(i) = j$, so folgt
$$\rho_j^{-1} \circ \pi(i) = \rho_j^{-1}(j) = n \quad \text{und somit} \quad \rho_j^{-1} \circ \pi \circ \rho_i(n) = n$$
Da sgn bekanntlich ein Gruppenhomomorphismus ist, folgt
$$\begin{aligned} \operatorname{sgn}(\rho_j^{-1} \circ \pi \circ \rho_i) &= \operatorname{sgn}(\rho_j^{-1}) \cdot \operatorname{sgn}(\pi) \cdot \operatorname{sgn}(\rho_i) = \operatorname{sgn}(\rho_j) \cdot \operatorname{sgn}(\pi) \cdot \operatorname{sgn}(\rho_i) \\ &= (-1)^{n-j} \cdot \operatorname{sgn}(\pi) \cdot (-1)^{n-i} = (-1)^{i+j} \cdot \operatorname{sgn}(\pi)\,. \end{aligned}$$

[14] regulär=invertierbar

Für eine Permutation $\pi \in S_n$ mit $\pi(i) = j$ sei nun also

$$\tau := \rho_j^{-1} \circ \pi \circ \rho_i \,,$$

d.h. man kann solch ein π umschreiben zu $\pi = \rho_j \circ \tau \circ \rho_i^{-1}$ und es gilt $\operatorname{sgn}(\tau) = (-1)^{i+j} \operatorname{sgn}(\pi)$. Nun kann man den Entwicklungssatz beweisen, indem man in der Leibniz-Formel jeweils alle $\pi \in S_n$ mit $\pi(i) = j$ bündelt und dann zu diesem τ übergeht:

$$\begin{aligned}
\det(A) &= \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{\ell=1}^{n} a_{\ell,\pi(\ell)} = \sum_{j=1}^{n} \sum_{\substack{\pi \in S_n \\ \pi(i)=j}} \operatorname{sgn}(\pi) \prod_{\ell=1}^{n} a_{\ell,\pi(\ell)} = \sum_{j=1}^{n} \sum_{\substack{\tau \in S_n \\ \tau(n)=n}} (-1)^{i+j} \operatorname{sgn}(\tau) \prod_{\ell=1}^{n} a_{\ell,\rho_j \circ \tau \circ \rho_i^{-1}(\ell)} = \\
&= \sum_{j=1}^{n} (-1)^{i+j} \sum_{\substack{\tau \in S_n \\ \tau(n)=n}} \operatorname{sgn}(\tau) \prod_{k=1}^{n} a_{\rho_i(k),\rho_j \circ \tau(k)} = \sum_{j=1}^{n} (-1)^{i+j} \sum_{\substack{\tau \in S_n \\ \tau(n)=n}} \operatorname{sgn}(\tau) a_{ij} \prod_{k=1}^{n-1} a_{\rho_i(k),\rho_j \circ \tau(k)} = \\
&= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \sum_{\tilde\tau \in S_{n-1}} \operatorname{sgn}(\tilde\tau) \prod_{k=1}^{n-1} (A^{(ij)})_{k,\tilde\tau(k)} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A^{(ij)})\,.
\end{aligned}$$

□

- In der vorliegenden Version beschreibt der Laplacesche Entwicklungssatz die *Entwicklung einer Matrix nach der i-ten Zeile.*
- Wegen $\det(A) = \det({}^tA)$ kann man aber genauso *nach einer Spalte* entwickeln:
- $\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A^{(ij)})$. für ein j mit $1 \le j \le n$.

Beispiel 12.4.4.

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1 \cdot (-1)^{1+1} \cdot \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} + 2 \cdot (-1)^{1+2} \cdot \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} + 3 \cdot (-1)^{1+3} \cdot \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} =$$

$$= 2 \cdot (-1)^{2+1} \cdot \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} + 5 \cdot (-1)^{2+2} \cdot \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} + 8 \cdot (-1)^{2+3} \cdot \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}\,.$$

- Durch Entwicklung nach einer Zeile wird aus einer $n \times n$-Determinante eine Summe von n $(n-1) \times (n-1)$-Determinanten (Spalte analog).
- Entwickelt man diese nochmal, so werden das $n \cdot (n-1)$ $(n-2) \times (n-2)$-Determinanten usw., d.h. der Aufwand wird erheblich!
- Deshalb sollte man nicht einfach drauflos entwickeln, sondern erst die entsprechende Zeile mit geeigneten elementaren Umformungen vereinfachen:

Beispiel 12.4.5.

$$\begin{vmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 10 \end{vmatrix} = \begin{vmatrix} 4 & -3 & -6 \\ 1 & 0 & 0 \\ 7 & -6 & -11 \end{vmatrix} = - \begin{vmatrix} -3 & -6 \\ -6 & -11 \end{vmatrix} = 3 \begin{vmatrix} 1 & 2 \\ -6 & -11 \end{vmatrix} = 3\,.$$

12.5 Der Produktsatz und seine Folgen

Satz 12.5.1. *Es seien R ein kommutativer Ring mit 1, $n \in \mathbb{N}$, $A, B \in R^{n \times n}$. Dann gilt*

$$\det(A \cdot B) = \det A \cdot \det B\,.$$

Beweis: Mit $A = (a_{ij})_{1\le i,j\le n}$, $B = (b_{ij})_{1\le i,j\le n}$ und $C = (c_{ij})_{1\le i,j\le n}$, wobei $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$ für $1 \le i,j \le n$ ist, gilt

$$\begin{aligned}
\det(A\cdot B) &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} c_{i,\pi(i)} = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n}\sum_{k=1}^{n} a_{ik}b_{k,\pi(i)} \overset{\text{Ausmult.}}{=} \\
&= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \sum_{k_1,k_2,\dots,k_n=1}^{n} \prod_{i=1}^{n} a_{i,k_i}b_{k_i,\pi(i)} = \sum_{k_1,k_2,\dots,k_n=1}^{n} \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i,k_i} \prod_{i=1}^{n} b_{k_i,\pi(i)} = \\
&= \sum_{k_1,k_2,\dots,k_n=1}^{n} \prod_{i=1}^{n} a_{i,k_i} \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{j=1}^{n} b_{k_j,\pi(j)} = \sum_{k_1,k_2,\dots,k_n=1}^{n} \prod_{i=1}^{n} a_{i,k_i} \underbrace{\det\left((b_{k_j,\ell})_{1\le j,\ell\le n}\right)}_{=0,\ \text{falls nicht alle } k_j \text{ verschieden}} \overset{k_j=\tau(j)}{=} \\
&= \sum_{\tau\in S_n} \prod_{i=1}^{n} a_{i,\tau(i)} \det\left((b_{\tau(j),\ell})_{1\le j,\ell\le n}\right) \overset{12.3.2}{=} \sum_{\tau\in S_n} \operatorname{sgn}(\tau) \prod_{i=1}^{n} a_{i,\tau(i)} \det B = \det A \cdot \det B\,.
\end{aligned}$$

□

Satz 12.5.2. *Es seien K ein Körper, $n \in \mathbb{N}$ und $A \in \mathrm{GL}(n,K)$. Dann gilt:*

$$\det(A^{-1}) = \det(A)^{-1}.$$

Beweis:

$$\begin{aligned}
1 = \det(\mathbb{1}_n) &= \det(A^{-1}A) \overset{12.5.1}{=} \det(A^{-1})\det(A) \Rightarrow \\
\det(A^{-1}) &= \det(A)^{-1}\,.
\end{aligned}$$

□

Satz 12.5.3. *Es seien K ein Körper, $n \in \mathbb{N}$ und $A, B \in K^{n\times n}$ ähnlich. Dann ist $\det(A) = \det(B)$.*

Beweis:

$$\begin{aligned}
A \text{ ähnlich zu } B \Rightarrow \exists S \in \mathrm{GL}(n,K) : A = S^{-1}BS \Rightarrow& \\
\det A = \det(S^{-1}BS) &\overset{12.5.1}{=} \det(S^{-1})\det B \det S \overset{12.5.2}{=} \\
&= \det(S)^{-1}\det S \det B = \det B\,.
\end{aligned}$$

□

- Ist f ein Endomorphismus eines K-Vektorraumes V mit $\dim V = n$, so besitzt er bezüglich einer Basis C die Abbildungsmatrix $[f]_C$.
- Beim Basiswechsel zu einer anderen Basis D von V ändert sich zwar die Matrix zu $[f]_D$, diese ist aber ähnlich zu $[f]_C$.
- Nach 12.5.3 folgt somit $\det([f]_C) = \det([f]_D)$.
- Die Determinante der Abbildungsmatrix hängt also nur von dem Endomorphismus f ab, nicht von der speziell gewählten Basis von V.
- Man nennt die Determinante deshalb eine *Invariante* des Endomorphismus f und definiert:

Definition 12.5.4. *Es seien K ein Körper, V ein K-Vektorraum mit $\dim V = n \in \mathbb{N}$ und $f \in \mathrm{End}(V)$.*

$$\det(f) := \det([f]_B)$$

für eine beliebige Basis B von V heißt Determinante des Endomorphismus.

12.6 Die Cramersche Regel

Satz 12.6.1. *Es seien K ein Körper, $n \in \mathbb{N}$, $A \in \mathrm{GL}(n,K)$ und $b \in K^n$. Sind $s_1, s_2, \dots, s_n$ die Spalten von A, so gilt für die eindeutige Lösung des LGS $Ax = b$*

$$x_i = \frac{\det(s_1,\dots,s_{i-1},b,s_{i+1},\dots,s_n)}{\det(A)}\,,\ i = 1,\dots,n$$

Beweis: Zum Nachweis wird die angegebene Lösung in Ax eingesetzt und dann wird die Determinante im Zähler von x_i nach Laplace entwickelt (nach der i-ten Spalte). Für die k-te Komponente gilt:

$$(Ax)_k = \sum_{i=1}^{n} a_{ki} x_i = \sum_{i=1}^{n} a_{ki} \frac{\det(s_1, \ldots, s_{i-1}, b, s_{i+1}, \ldots, s_n)}{\det(A)} = \frac{1}{\det(A)} \sum_{i=1}^{n} a_{ki} \sum_{\ell=1}^{n} (-1)^{\ell+i} b_\ell \det(A^{(\ell i)}) =$$

$$= \frac{1}{\det(A)} \sum_{\ell=1}^{n} b_\ell \sum_{i=1}^{n} a_{ki} (-1)^{\ell+i} \det(A^{(\ell i)}) = \frac{1}{\det(A)} \sum_{\ell=1}^{n} b_\ell \delta_{k\ell} \det(A) = b_k$$

denn der Ausdruck in der inneren Summe ist gerade die Entwicklung der Matrix, die aus A durch Ersetzen der ℓ-ten Zeile durch die k-te Zeile entsteht, nach der ℓ-ten Zeile. Ist $k = \ell$, so ist das also gerade $\det(A)$, ist dagegen $k \neq \ell$, so betrachtet man eine Matrix mit 2 gleichen Zeilen, d.h. deren Determinante ist 0. □

- Die Bedeutung der Cramerschen Regel ist relativ gering (trotzdem wird sie leider oft in der Schule dem Gauß-Verfahren vorgezogen):
  - sie kann wegen des Einsatzes der Determinante nur für Systeme mit quadratischer Koeffizientenmatrix verwendet werden,
  - sie kann wegen der Determinante im Nenner nur im Fall $\det A \neq 0$ angewendet werden (also bei einer eindeutigen Lösung),
  - ihre Anwendung ist in den meisten Fällen mit deutlich mehr Aufwand verbunden als z.B. das normale Gauß-Verfahren; so muss man bereits für ein 3×3-LGS vier 3×3-Determinanten berechnen!
- In der Übung wird ein Beispiel für die sinnvolle Verwendung der Cramerschen Regel besprochen.
- Für den Nachweis der linearen Unabhängigkeit von Vektoren oder das Lösen linearer Gleichungssysteme spielen Determinanten eher eine untergeordnete Rolle, weil das Gauß-Verfahren viel einfacher ist.
- Ihre volle Bedeutung wird sich erst im nächsten Kapitel über Eigenwerte und Eigenvektoren herausstellen.

12.7 Die Spur

Definition 12.7.1. *Es seien R ein kommutativer Ring mit 1, $n \in \mathbb{N}$ und $A = (a_{ij})_{1 \leq i,j \leq n} \in R^{n \times n}$. Die Spur der Matrix A (englisch trace) ist*

$$\operatorname{Spur}(A) := \sum_{i=1}^{n} a_{ii} \,.$$

- In den Übungen (→6.1 & 6.4) wurde nachgerechnet, dass sich die Spur einer Abbildungsmatrix beim Basiswechsel in Vektorräumen nicht ändert.
- Somit ist die Spur eine Invariante von Endomorphismen.

Definition 12.7.2. *Es seien K ein Körper, V ein K-Vektorraum mit $\dim V = n \in \mathbb{N}$ und $f \in \operatorname{End}(V)$. Für eine beliebige Basis B von V sei die Spur des Endomorphismus*

$$\operatorname{Spur}(f) := \operatorname{Spur}([f]_B) \,.$$

13 Eigenwerte und Eigenvektoren

13.1 Motivation

Beispiel 13.1.1. *Zwei oder mehr gekoppelte Pendel werden vorteilhaft durch (zwei oder mehr) Differentialgleichungen (DGL) beschrieben.*
Auch, wenn Differentialgleichungen eigentlich in der Analysis behandelt werden, sieht man hier auch eine der wesentlichen Anwendungen der linearen Algebra.
Für dieses Beispiel betrachten wir 2 Massen an 2 Federn:

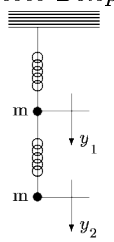


Abbildung 26: gekoppelte Pendel

Die Ruhelagen der Federn seien $y_1 = 0$ und $y_2 = 0$. Beide Federn haben die gleiche Federkonstante k. Beide Massen seien m. Dann sind die Bewegungsgleichungen

$$m\ddot{y}_1 = -ky_1 + k(y_2 - y_1) = k(y_2 - 2y_1)\,,$$
$$m\ddot{y}_2 = -k(y_2 - y_1)\,.$$

Ein Punkt über den Ortskoordinaten bedeutet dabei die Ableitung $\frac{d}{dt}$ nach der Zeit, d.h. $\dot{y}$ ist die Geschwindigkeit, $\ddot{y}$ die Beschleunigung der Masse mit der Koordinate y.

Das Problem bei diesen Gleichungen ist nicht nur, dass es sich um Differentialgleichungen handelt (d.h. y, $\dot{y}$, $\ddot{y}$ etc. werden in einer Gleichung verknüpft und man möchte y berechnen), sondern hier liegt ein gekoppeltes Differentialgleichungssystem vor, d.h. mehrere Funktionen und ihre Ableitungen werden miteinander verknüpft.
Mit $m = k = 1$ und Matrixschreibweise sieht das Beispiel so aus:

$$\begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{pmatrix} = \underbrace{\begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix}}_{B:=} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Deshalb spricht man von einem linearen Differentialgleichungssystem.
Es wäre einem schon sehr geholfen, wenn man solche Systeme entkoppeln könnte, d.h. wenn man aus einem System mit n Variablen n Gleichungen mit nur jeweils einer Variablen machen könnte.
Die Idee dazu stammt aus der linearen Algebra.
Entkopplung von Gleichungen mit n Variablen bedeutet, dass man Funktionen $x_1, \ldots, x_n$ sucht, so dass die lineare (invertierbare!) Transformation

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

auf ein entkoppeltes System

$$\begin{pmatrix} \ddot{x}_1 \\ \vdots \\ \ddot{x}_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

führt.
*Das heißt $\ddot{x}_1 = \lambda_1 x_1\,, \ddot{x}_2 = \lambda_2 x_2\,, \ldots, \ddot{x}_n = \lambda_n x_n\,,$ was sehr leicht lösbar ist (*exp *oder Winkelfunktionen* sin, cos*).*
Wendet man solch eine lineare Transformation im letzten Beispiel an, so ergibt sich

$$A \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Rightarrow A \cdot \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} \Rightarrow A \cdot \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = \begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{pmatrix}$$

und somit

$$A \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \iff$$
$$\begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = \underbrace{A^{-1} \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} A}_{\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \overset{!}{=}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

- Die hieraus resultierende Fragestellung kam bereits im Abschnitt über lineare Abbildungen vor.

- Gesucht ist eine Ähnlichkeitstransformation mit einer invertierbaren Matrix A, so dass die transformierte Matrix Diagonalgestalt hat.
- Bekanntlich beschreibt diese Ähnlichkeitstransformation einen Basiswechsel der linearen Abbildung $f : x \mapsto Bx$.
- Das Problem lässt sich deshalb auch so beschreiben: gesucht wird eine Basis, bzgl. der die Abbildungsmatrix der linearen Abbildung f Diagonalgestalt hat.

Beispiel 13.1.2. *Zu Beispiel 13.1.1 gibt es solch eine Matrix. Mit*

$$A = \begin{pmatrix} -\frac{1}{2} + \frac{1}{2}\sqrt{5} & -\frac{1}{2} - \frac{1}{2}\sqrt{5} \\ 1 & 1 \end{pmatrix}$$

gilt

$$A^{-1} \cdot B \cdot A = \begin{pmatrix} -\frac{3}{2} + \frac{1}{2}\sqrt{5} & 0 \\ 0 & -\frac{3}{2} - \frac{1}{2}\sqrt{5} \end{pmatrix}$$

d.h. die durch die Matrix B bzgl. der kanonischen Basis gegebene lineare Abbildung hat bezüglich der Basis

$$\left(\begin{pmatrix} -\frac{1}{2} + \frac{1}{2}\sqrt{5} \\ 1 \end{pmatrix}, \begin{pmatrix} -\frac{1}{2} - \frac{1}{2}\sqrt{5} \\ 1 \end{pmatrix} \right)$$

Diagonalgestalt. *Den Rest des Beispiels erledigt nun die Analysis. Für die Anfangswerte $y_1(0) = 0$, $y_2(0) = 1$, $\dot{y}_1(0) = 0$, $\dot{y}_2(0) = 0$ ergibt sich z.B. die Lösung:*

$$\begin{aligned} y_1(t) &= \tfrac{1}{5}\sqrt{5}\cos\left(\tfrac{1}{2}t\sqrt{5} - \tfrac{1}{2}t\right) - \tfrac{1}{5}\sqrt{5}\cos\left(\tfrac{1}{2}t + \tfrac{1}{2}t\sqrt{5}\right) \\ y_2(t) &= \left(\tfrac{1}{10}\sqrt{5} + \tfrac{1}{2}\right)\cos\left(\tfrac{1}{2}t\sqrt{5} - \tfrac{1}{2}t\right) + \left(\tfrac{1}{2} - \tfrac{1}{10}\sqrt{5}\right)\cos\left(\tfrac{1}{2}t + \tfrac{1}{2}t\sqrt{5}\right) \end{aligned}$$

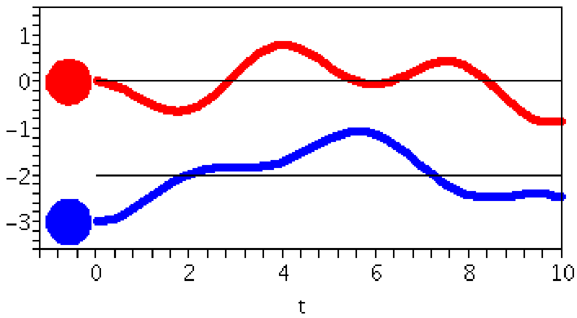


Abbildung 27: Gekoppelte Pendel

13.2 Grundlegende Definitionen

Definition 13.2.1. *Es seien K ein Körper, V ein K-Vektorraum und $f \in \mathrm{End}(V)$. Ein Vektor $v \in V \setminus \{0\}$ heißt Eigenvektor des Endomorphismus f, wenn es ein $\lambda \in K$ gibt, so dass*

$$f(v) = \lambda \cdot v\,.$$

Das λ heißt der zu diesem Eigenvektor gehörige Eigenwert.
Betrachtet man für $\dim V < \infty$ die Abbildungsmatrix $A := [f]_B$ des Endomorphismus bezüglich einer geordneten Basis B von V, so gilt

$$A \cdot v_{/B} = \lambda \cdot v_{/B}$$

und man nennt v Eigenvektor und λ den zugehörigen Eigenwert von A.

- Der Nullvektor hat für jedes $f \in \mathrm{End}(V)$ und jedes $\lambda \in K$ die Eigenschaft $f(0) = \lambda \cdot 0$, ist aber nie ein Eigenvektor!

Beispiel 13.2.2. *Man betrachte Vektoren bei einer Spiegelung an einer Ebene durch den Ursprung im $\mathbb{R}^3$:*

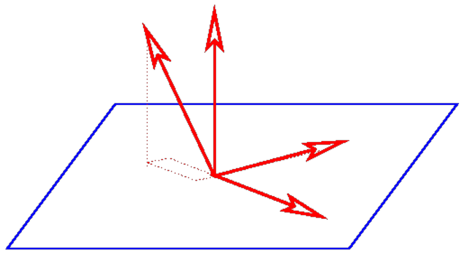


Abbildung 28: EVn bei Spiegelung

Aus der Geometrie der Abbildung klar, dass die Vektoren in der Ebene (außer 0*) Eigenvektoren zum Eigenwert* 1 *sind und die Vektoren, die orthogonal auf der Ebene stehen (außer* 0*), Eigenvektoren zum Eigenwert* -1*.*

- Bisher ist noch nicht geklärt, ob es immer Eigenwerte und -vektoren gibt, und, falls das der Fall ist, wie man diese berechnet.
- Man kann aber schon mal untersuchen, was für die wünschenswerte Diagonalgestalt alles nötig wäre:

Definition 13.2.3. *Es seien* K *ein Körper,* V *ein endlichdimensionaler* K*-Vektorraum und* $f \in \mathrm{End}(V)$*. Der Endomorphismus* f *heißt diagonalisierbar, wenn es eine Basis* B *von* V *gibt, so dass* $[f]_B$ *Diagonalgestalt hat.*

Beispiel 13.2.4. $\begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix}$ *aus 13.1.1 ist diagonalisierbar mit Diagonalgestalt* $\begin{pmatrix} -\frac{3}{2}+\frac{1}{2}\sqrt{5} & 0 \\ 0 & -\frac{3}{2}-\frac{1}{2}\sqrt{5} \end{pmatrix}$

Satz 13.2.5. *Es seien* K *ein Körper,* V *ein endlichdimensionaler* K*-Vektorraum und* $f \in \mathrm{End}(V)$*. Der Endomorphismus* f *ist genau dann diagonalisierbar, wenn* V *eine Basis aus Eigenvektoren von* f *besitzt.*

Beweis: Es sei $B = (b_1, \ldots, b_n)$ eine Basis von V, bezüglich der $[f]_B$ Diagonalgestalt hat, also

$$[f]_B = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & 0 & \ldots & 0 & \lambda_n \end{pmatrix} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)\,.$$

Für die Vektoren in B heißt das

$$f(b_1) = \lambda_1 \cdot b_1\,,\ f(b_2) = \lambda_2 \cdot b_2, \ldots, f(b_n) = \lambda_n \cdot b_n\,,$$

d.h. das sind Eigenvektoren zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$.
Sind umgekehrt die Vektoren in B allesamt Eigenvektoren so folgt

$$f(b_1) = \lambda_1 \cdot b_1\,,\ f(b_2) = \lambda_2 \cdot b_2, \ldots, f(b_n) = \lambda_n \cdot b_n\,,$$

und somit $[f]_B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. □

Satz 13.2.6. *Es seien* K *ein Körper,* V *ein endlichdimensionaler* K*-Vektorraum und* $f \in \mathrm{End}(V)$*. Es ist* $\lambda \in K$ *genau dann ein Eigenwert von* f*, wenn gilt*

$$\det(f - \lambda \cdot \mathrm{id}_V) = 0\,.$$

Beweis: Laut Definition gilt für einen EV v und den zugehörigen EW λ:

$$f(v) = \lambda \cdot v = \lambda \cdot \mathrm{id}_V(v) \iff f(v) - \lambda \cdot \mathrm{id}_V(v) = 0 \iff (f - \lambda \cdot \mathrm{id}_V)(v) = 0$$

Mit f und id_V ist auch $f - \lambda \cdot \mathrm{id}_V$ ein Endomorphismus von V.
Für festes $\lambda \in K$ ist $(f - \lambda \cdot \mathrm{id}_V)(v) = 0$ ein homogenes LGS und hat somit immer die triviale Lösung $v = 0$.
Die triviale Lösung ist aber uninteressant, denn *definitionsgemäß ist ein Eigenvektor von Null verschieden.*

Genau dann sind v und λ also Eigenvektor und zugehöriger Eigenwert, wenn v eine *nichttriviale ($\neq 0$)* Lösungen des LGS $(f - \lambda \cdot \mathrm{id}_V)(v) = 0$ ist.
Nichttriviale Lösungen eines homogenen LGS existieren genau dann, wenn der Rang der Koeffizientenmatrix (bzw. des zugehörigen Homomorphismus) nicht voll ist, hier also genau dann, wenn $f - \lambda \cdot \mathrm{id}_V \notin \mathrm{GL}(V)$ ist.
Dafür gibt es aber bereits ein Kriterium mit Hilfe von Determinanten aus 12.3.10:

$$f - \lambda \cdot \mathrm{id}_V \notin \mathrm{GL}(V) \iff \det(f - \lambda \cdot \mathrm{id}_V) = 0\,. \qquad \square$$

9.12.10

- Ist A die Abbildungsmatrix von f bezüglich der gerade gewählten Basis, also $A = [f]_B$, so ist $C := A - \lambda \mathbb{1}_n$ die Matrix zu $f - \lambda \cdot \mathrm{id}_V$.
- Nach der Leibniz-Regel 12.2.1 ist die Determinante von $f - \lambda \cdot \mathrm{id}_V$ bzw. von C eine Summe von Produkten aus je n Einträgen von C.
- Dies bedeutet, dass $\det(f - \lambda \cdot \mathrm{id}_V)$ ein Polynom in λ vom Grad n ist.
- Die größte λ-Potenz ensteht bei dem Produkt über die Hauptdiagonale

  $$\prod_{i=1}^{n} c_{ii} = \prod_{i=1}^{n} (a_{ii} - \lambda) = (a_{11} - \lambda)(a_{22} - \lambda) \cdot \cdots \cdot (a_{nn} - \lambda)\,,$$

  d.h. der Leitterm des Polynoms ist $(-\lambda)^n = (-1)^n \lambda^n$.
- Auch die zweitgrößte λ-Potenz λ^{n-1} ensteht einzig aus diesem Produkt.
- Der Koeffizient dieser Potenz ist

  $$(-1)^{n-1} \sum_{i=1}^{n} a_{ii} = (-1)^{n-1}\,\mathrm{Spur}(A) = (-1)^{n-1}\,\mathrm{Spur}(f)$$

- Den konstanten Term eines Polynoms bekommt man, indem man für die Variable 0 einsetzt.
- Der konstante Term von $\det(f - \lambda \cdot \mathrm{id}_V)$ ist also $\det(f) = \det(A)$.
- Zusammen hat man somit

$$\det(f - \lambda \mathrm{id}_V) = (-1)^n \lambda^n + (-1)^{n-1}\,\mathrm{Spur}(f) \lambda^{n-1} + \cdots + \det(f), \text{ bzw.} \tag{13.1}$$
$$\det(A - \lambda \mathbb{1}_n) = (-1)^n \lambda^n + (-1)^{n-1}\,\mathrm{Spur}(A) \lambda^{n-1} + \cdots + \det(A) \tag{13.2}$$

Definition 13.2.7. *Es seien K ein Körper, V ein endlichdimensionaler K-Vektorraum, $f \in \mathrm{End}(V)$ und $A \in K^{n \times n}$. Dann heißt*

$$\chi_f(\lambda) := \det(f - \lambda \cdot \mathrm{id}_V) \in K[\lambda]$$

das charakteristische Polynom des Endomorphismus f bzw.

$$\chi_A(\lambda) := \det(A - \lambda \mathbb{1}_n)$$

das charakteristische Polynom der Matrix A. $\chi_f(\lambda) = 0$ bzw. $\chi_A(\lambda) = 0$ heißt die charakteristische Gleichung.

- Die Lösungen der charakteristischen Gleichung, bzw. die Nullstellen (*Wurzeln* genannt) des charakteristischen Polynoms (falls vorhanden - hängt vom Körper K ab), sind die Eigenwerte von f bzw. A.

Beispiel 13.2.8. *In dem einleitenden Beispiel 13.1.1 ging es um die Matrix*

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

bzw. lineare Abbildung $f = f_B$ und ihre Diagonalisierung.
Dazu braucht man die Eigenwerte und -vektoren und startet deshalb mit der charakteristischen Gleichung

$$\chi_B(\lambda) = \det \begin{pmatrix} -2 - \lambda & 1 \\ 1 & -1 - \lambda \end{pmatrix} = 0\,.$$

Mit Sarrus bekommt man

$$\chi_B(\lambda) = \lambda^2 + 3\lambda + 1 = 0 \iff \lambda_{1,2} = \frac{-3 \pm \sqrt{5}}{2}.$$

Das charakteristische Polynom hätte man dabei auch mit (13.2) berechnen können, was im Fall $n = 2$ zu

$$\chi_B(\lambda) = \lambda^2 - \operatorname{Spur}(B)\lambda + \det(B)$$

wird. Hier ist $\operatorname{Spur}(B) = -3$, $\det(B) = 1$.
Zu den zwei Eigenwerten $\lambda_{1,2} = \frac{-3\pm\sqrt{5}}{2}$ kann man nun Eigenvektoren berechnen. Dazu setzt man diese Werte in das homogene lineare Gleichungssystem

$$(B - \lambda \mathbb{1}_2)x = 0$$

ein. Nach der Vorarbeit mit den Eigenwerten weiß man, das diese Gleichung genau für $\lambda = \lambda_{1,2}$ nichttriviale Lösungen besitzt.

$$\begin{pmatrix} -2-\lambda_1 & 1 \\ 1 & -1-\lambda_1 \end{pmatrix} = \begin{pmatrix} \frac{-1+\sqrt{5}}{2} & 1 \\ 1 & \frac{1+\sqrt{5}}{2} \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 0 \\ 1 & \frac{1+\sqrt{5}}{2} \end{pmatrix} \qquad \boxed{\lambda_1 = \frac{-3-\sqrt{5}}{2}}$$

$$\Rightarrow x = \mu \begin{pmatrix} \frac{1+\sqrt{5}}{2} \\ -1 \end{pmatrix} \text{ mit } \mu \neq 0$$

$$\begin{pmatrix} -2-\lambda_2 & 1 \\ 1 & -1-\lambda_2 \end{pmatrix} = \begin{pmatrix} \frac{-1-\sqrt{5}}{2} & 1 \\ 1 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 0 \\ 1 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \qquad \boxed{\lambda_2 = \frac{-3+\sqrt{5}}{2}}$$

$$\Rightarrow x = \nu \begin{pmatrix} \frac{1-\sqrt{5}}{2} \\ -1 \end{pmatrix} \text{ mit } \nu \neq 0$$

Die beiden Eigenvektoren

$$v_1 := \begin{pmatrix} \frac{1+\sqrt{5}}{2} \\ -1 \end{pmatrix} \text{ und } v_2 := \begin{pmatrix} \frac{1-\sqrt{5}}{2} \\ -1 \end{pmatrix}$$

zu den verschiedenen Eigenwerte λ_1 und λ_2 sind linear unabhängig, bilden also eine Basis des $\mathbb{R}^2$. Wechselt man von der kanonischen Basis auf die Basis $C := (v_1, v_2)$ aus Eigenvektoren, so wird aus der Abbildungsmatrix $B = [f]_E$ die sehr viel einfachere Matrix $[f]_C = \operatorname{diag}(\lambda_1, \lambda_2)$.

Satz 13.2.9. *Es seien K ein Körper, V ein endlichdimensionaler K-Vektorraum, $n \in \mathbb{N}$ und $A, B \in K^{n\times n}$ ähnlich. Dann gilt*

$$\chi_A(\lambda) = \chi_B(\lambda), \ \det A = \det B, \ \operatorname{Spur} A = \operatorname{Spur} B.$$

Beweis: Sind $A, B \in K^{n\times n}$ ähnlich, so gibt es ein $S \in \mathrm{GL}(n, K)$ mit $A = S^{-1}BS$. Damit folgt

$$\begin{aligned} \chi_A(\lambda) &= \det(A - \lambda \mathbb{1}_n) = \det(S^{-1}BS - \lambda S^{-1}\mathbb{1}_n S) \overset{12.5.1}{=} \\ &= \det(S^{-1}) \det(B - \lambda \mathbb{1}_n) \det(S) \\ &= (\det(S))^{-1} \det(S) \det(B - \lambda \mathbb{1}_n) = \\ &= \chi_B(\lambda). \end{aligned}$$

Da die charakteristischen Polynome gleich sind, stimmen auch deren Nullstellen, das sind die Eigenwerte von A bzw. B, entsprechend Ihrer Vielfachheit überein. Gemäß (13.2) sind $\det A$ und $\det B$ bzw. $\operatorname{Spur} A$ und $\operatorname{Spur} B$ entsprechende Koeffizienten von $\chi_A(\lambda)$ und $\chi_B(\lambda)$. Nachdem die Polynome gleich sind, stimmen sie also auch in diesen Koeffizienten überein. □

Satz 13.2.10. *Es seien K ein Körper, V ein endlichdimensionaler K-Vektorraum und $f \in \operatorname{End}(V)$.*

(i) Wenn f diagonalisierbar ist, so zerfällt $\chi_f(\lambda)$ in Linearfaktoren.

(ii) Wenn $\chi_f(\lambda)$ in Linearfaktoren zerfällt und seine Wurzeln (=Eigenwerte von f) paarweise verschieden sind, so ist f diagonalisierbar.

Beweis: (i) Ist f diagonalisierbar, so gibt es eine Basis B von V mit

$$[f]_B = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \text{ mit } \lambda_1, \dots, \lambda_n \in K.$$

Damit gilt für das charakteristische Polynom

$$\chi_f(\lambda) = \det(\mathrm{diag}(\lambda_1 - \lambda, \ldots, \lambda_n - \lambda)) = (\lambda_1 - \lambda) \cdot \ldots \cdot (\lambda_n - \lambda)$$

(ii) Ist nun $\chi_f(\lambda) = (\lambda_1 - \lambda) \cdot \ldots \cdot (\lambda_n - \lambda)$ mit paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_n \in K$, so hat zu jedem dieser Eigenwerte die Gleichung $(f - \lambda_i \cdot \mathrm{id}_V)(v) = 0$ bzw. $([f]_B - \lambda_i \cdot \mathbb{1}_n)v = 0$ mindestens eine nichttriviale Lösung v_i, einen Eigenvektor zu λ_i. Wenn man jetzt noch zeigen kann, dass $v_1, \ldots, v_n$ linear unabhängig sind, so hat man die gesuchte Basis aus Eigenvektoren.
Behauptung: Es seien $\lambda_1, \ldots, \lambda_m$ paarweise verschiedene Eigenwerte von f, $1 \le m \le n$ und $v_1, \ldots, v_m$ zugehörige Eigenvektoren. Dann gilt: $\{v_1, \ldots, v_m\}$ ist linear unabhängig.
Dieser Teil wird mit Induktion nach m gezeigt:
Für $m = 1$ ist $v_1 \neq 0$ linear unabhängig.
Nun sei $m > 1$ und $v_1, \ldots, v_{m-1}$ gemäß Induktionsvoraussetzung linear unabhängig. Man macht den üblichen Ansatz

$$\sum_{k=1}^{m} \nu_k v_k = 0$$

Dann sind zwei Fälle zu unterscheiden:
Ist $\underline{\nu_m = 0}$, also

$$\sum_{k=1}^{m-1} \nu_k v_k = 0\,,$$

so ist nach Induktionsvoraussetzung $\nu_k = 0$ für $k = 1, \ldots, m-1$, also zusammen $\nu_k = 0$ für $k = 1, \ldots, m$, d.h. $v_1, \ldots, v_m$ sind linear unabhängig.
Ist $\underline{\nu_m \neq 0}$, so kann man auflösen nach v_m und erhält

$$v_m = -\frac{1}{\nu_m} \sum_{k=1}^{m-1} \nu_k v_k \tag{13.3}$$

Da v_m Eigenvektor von f zum Eigenwert λ_m ist, folgt

$$\begin{aligned} f(v_m) = \lambda_m v_m \iff f\Big(-\frac{1}{\nu_m} \sum_{k=1}^{m-1} \nu_k v_k\Big) = \lambda_m v_m \iff \\ -\frac{1}{\nu_m} \sum_{k=1}^{m-1} \nu_k f(v_k) = \lambda_m v_m \overset{(13.3)}{\iff} -\frac{1}{\nu_m} \sum_{k=1}^{m-1} \nu_k \lambda_k v_k = -\lambda_m \frac{1}{\nu_m} \sum_{k=1}^{m-1} \nu_k v_k \\ \iff \sum_{k=1}^{m-1} (\lambda_m - \lambda_k) \nu_k v_k = 0 \end{aligned}$$

Da nach Induktionsvoraussetzung $v_1, \ldots, v_{m-1}$ linear unabhängig sind, folgt $(\lambda_m - \lambda_k)\nu_k = 0$ für $k = 1, \ldots m-1$. Wegen $\lambda_m \neq \lambda_k$ für $k = 1, \ldots, m-1$ heißt das $\nu_k = 0$ für $k = 0, \ldots, m-1$. Einsetzen in $\sum_{k=1}^{m} \nu_k v_k = 0$ führt dann aber auf $\nu_m v_m = 0$, was wegen $v_m \neq 0$ nicht sein kann. Die Annahme $\nu_m \neq 0$ ist also falsch. □

13.3 Algebraische und geometrische Vielfachheit

- Die Eigenvektoren zu einem Eigenwert $\lambda_i \in K$ sind gerade alle von 0 verschiedenen Lösungen der charakteristischen Gleichung $(f - \lambda_i \cdot \mathrm{id}_V)(v) = 0$ bzw. $([f]_B - \lambda_i \cdot \mathbb{1}_n)v = 0$.
- Die Menge aller Lösungen dieses linearen homogenen Gleichungssystems bilden bekanntlich einen Vektorraum (→Übung 4.4)
- Dieser Vektorraum ist mindestens eindimensional, denn λ_i wird ja gerade so bestimmt, dass das LGS nicht nur die Lösung 0 hat.

Definition 13.3.1. *Es seien K ein Körper, V ein K-Vektorraum und $f \in \mathrm{End}(V)$. Der Eigenraum von f zum Eigenwert λ_i ist*

$$E_f(\lambda_i) := \{v \in V \mid (f - \lambda_i \cdot \mathrm{id}_V)(v) = 0\} = \mathrm{Kern}(f - \lambda_i \cdot \mathrm{id}_V)\,.$$

Die Dimension $g_f(\lambda_i) := \dim(E_f(\lambda_i))$ heißt die geometrische Vielfachheit des Eigenwerts λ_i.

- Bevor man an die Berechnung von Eigenvektoren geht, braucht man erst einmal die Eigenwerte.
- Diese sind als Nullstellen des charakteristischen Polynoms je nach Größe von n und zugrunde liegendem Körper ziemlich schwierig zu bekommen.
- Der folgende *Fundamentalsatz der Algebra* gibt Auskunft darüber, wie viele Nullstellen man erwarten kann.
- Der erste vollständige Beweis für den Fundamentalsatz der Algebra wurde 1799 von *Johann Carl Friedrich Gauß (1777 - 1855)* im Rahmen seiner Dissertation angegeben.
- Der Beweis dieses Satzes kommt für Mathematikstudenten meistens in der Vorlesung Funktionentheorie (=höhere komplexe Analysis).
- Für diesen Beweis fehlen momentan noch die mathematischen Grundlagen. Sie müssen den Satz also vorerst glauben.

Satz 13.3.2. *Jedes nichtkonstante Polynom $p(z) \in \mathbb{C}[z]$ besitzt eine Nullstelle.*

- Hat $p(z)$ die Nullstelle $z_1 \in \mathbb{C}$, so ist $p(z)$ durch $z - z_1$ teilbar, d.h. es gibt ein $p_1(z) \in \mathbb{C}[z]$ mit $p(z) = (z - z_1) \cdot p_1(z)$.
- Ist $\deg p = n \in \mathbb{N}$, so ist $\deg p_1 = n - 1$.
- Ist $n \geq 2$, so ist p_1 nichtkonstant und man kann den Satz erneut anwenden, d.h. es gibt ein $z_2 \in \mathbb{C}$ und ein $p_2(z) \in \mathbb{C}[z]$ mit $p(z) = (z - z_1)(z - z_2) \cdot p_2(z)$.
- Dies hat zur Folge, dass ein Polynom $p(z) \in \mathbb{C}[z]$ mit dem Grad $\deg(p(z)) = n$ eine (bis auf die Reihenfolge) eindeutige Faktorisierung
$$p(z) = a(z - z_1)(z - z_2) \cdot \ldots \cdot (z - z_n) \tag{13.4}$$
mit $a \in \mathbb{C} \setminus \{0\}$ und $z_1, \ldots, z_n \in \mathbb{C}$ (nicht unbedingt verschieden) besitzt, d.h. es hat (wenn man die Vielfachheiten berücksichtigt) genau n komplexe Wurzeln.
- Ist $p(z) \in \mathbb{R}[z]$, so kann man wegen $\mathbb{R} \subset \mathbb{C}$ ebenfalls den Fundamentalsatz verwenden, nur werden die z_i aus (13.4) oft nicht in $\mathbb{R}$ liegen.
- Bekanntlich ist in $\mathbb{C}$ die Abbildung $z \mapsto \overline{z}$ (*konjugiert Komplexes*) ein *Körperautomorphismus*, d.h. es gilt für beliebige $z, w \in \mathbb{C}$: $\overline{z + w} = \overline{z} + \overline{w}$ und $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$.
- Damit folgt, dass die nichtreellen Nullstellen von $p(z) \in \mathbb{R}[z]$ jeweils paarweise auftreten: Ist $z_1 \in \mathbb{C} \setminus \mathbb{R}$ eine Nullstelle von $p(z) \in \mathbb{R}[z]$, so ist auch $\overline{z}_1 \in \mathbb{C} \setminus \mathbb{R}$ eine Nullstelle von $p(z)$.
- Die beiden Faktoren $z - z_1$ und $z - \overline{z}_1$ von $p(z)$ ergeben zusammen $(z - z_1)(z - \overline{z}_1) = z^2 - (z_1 + \overline{z}_1) \cdot z + z_1\overline{z}_1$.
- Wegen $z_1 + \overline{z}_1 = 2 \operatorname{Re} z_1, z_1\overline{z}_1 = |z_1|^2 \in \mathbb{R}$ ist dieses quadratische Polynom in $\mathbb{R}[x]$.
- Ein reelles Polynom 5ten Grades hat somit z.B. 1, 3 oder 5 reelle Nullstellen (entsprechend der Vielfachheit gezählt), die jeweils restlichen Nullstellen sind paarweise konjugierte, echt komplexe Nullstellen(das wird auch oft in der Analysis mit dem Zwischenwertsatz gezeigt).

Satz 13.3.3 (Merkregel für rationale Wurzeln)**.** *Es sei $p(z) = \sum_{i=0}^{n} p_i z^i \in \mathbb{Z}[z]$, also $p_i \in \mathbb{Z}$ für $i = 0, \ldots, n$ mit $n = \deg p \geq 1$, also $p_n \neq 0$. Hat $p(z)$ eine rationale Nullstelle $\frac{a}{b}$ (mit $a \in \mathbb{Z}$, $b \in \mathbb{N}$), so ist a ein Teiler des konstanten Terms p_0 und b ein Teiler des Leitkoeffizienten p_n.*

Beweis: Hat $p(z)$ die rationale Nullstelle $\frac{a}{b}$, so hat $p(z)$ den Linearfaktor $z - \frac{a}{b} \in \mathbb{Q}[z]$ bzw. $bz - a \in \mathbb{Z}[z]$. Ausmultiplizieren zeigt die Behauptung. □

Beispiel 13.3.4. *Gegeben sei das Polynom*

$$p(z) = z^5 - 3z^4 - z^3 + 3z^2 - 2z + 6 \in \mathbb{Z}[z]$$

Es ist $p_0 = 6$, also gemäß Merkregel 13.3.3: $a \in \{\pm 6, \pm 3, \pm 2, \pm 1\}$.
Es ist $\deg p = 5$ und $p_5 = 1$, also gemäß Merkregel 13.3.3: $b \in \{\pm 1\}$.
Zusammen liefert das $\frac{a}{b} \in \{\pm 6, \pm 3, \pm 2, \pm 1\}$.

Damit kann durch gezieltes Probieren recht schnell herausfinden, ob das Polynom solche Nullstellen hat. Im vorliegenden Fall ist z.B. 3 eine rationale Nullstelle, die man auf diese Art findet. Damit kann man den Linearfaktor $z-3$ durch Polynomdivision abspalten und erhält

$$\frac{p(z)}{z-3} = z^4 - z^2 - 2\,.$$

Macht man mit dem verbleibenden Polynom den gleichen Test nochmal, so stellt sich heraus, dass keiner der untersuchten rationalen Kandidaten eine Nullstelle ist, d.h. über den rationalen Zahlen hat das Polynom keine weiteren Linearfaktoren (es könnte aber sehr wohl noch in 2 quadratische Faktoren zerfallen). Damit ist man bei einem Polynom 4.Grades üblicherweise auch schon am Ende seines Lateins, denn die so genannten Cardanischen Formeln (nach Gerolamo Cardano (1501 - 1576)) für Polynome dritten und vierten Grades sind so aufwändig, dass man sie selbst in einigen Formelsammlungen nicht findet (ab dem Grad 5 gibt es keine Formeln mehr). Üblicherweise überlässt man diese Arbeit einem CAS (Computeralgebra-System).
Im vorliegenden Fall hat man Glück, weil es sich um eine so genannte biquadratische Gleichung handelt, d.h. man kann durch die Substitution $z^2 = x$ zu einer quadratischen Gleichung übergehen und diese dann mit der „Mitternachtsformel" faktorisieren:

$$\begin{aligned} x^2 - x - 2 =& 0 \iff x_{1,2} = \begin{cases} -1 \\ 2 \end{cases} \\ \Rightarrow x^2 - x - 2 =& (x+1)(x-2) \Rightarrow z^4 - z^2 - 2 = (z^2+1)(z^2-2) \end{aligned}$$

Damit hat man jetzt insgesamt

$$p(z) = (z-3)(z^2+1)(z^2-2)\,.$$

Das ist die vollständige Faktorisierung von $p(z)$ über $\mathbb{Q}$.
Von den beiden quadratischen Faktoren ist z^2+1 über den reellen Zahlen unzerlegbar, denn jedes Quadrat ist dort nichtnegativ.
Der Faktor z^2-2 lässt sich aber zerlegen in $(z-\sqrt{2})(z+\sqrt{2})$.
Die vollständige Faktorisierung von p über $\mathbb{R}$ lautet somit

$$p(z) = (z-3)(z-\sqrt{2})(z+\sqrt{2})(z^2+1)\,.$$

Nimmt man nun aber die komplexen Zahlen zu Hilfe, so zerfällt auch noch (gemäß der Aussage des Fundamentalsatzes) der letzte Faktor z^2+1 in die Linearfaktoren $(z-i)(z+i)$, so dass man insgesamt

$$p(z) = (z-3)(z-\sqrt{2})(z+\sqrt{2})(z-i)(z+i)$$

als vollständige Faktorisierung von p über $\mathbb{C}$ hat.
Das Polynom 5.Grades p hat also hat also über $\mathbb{C}$ die 5 einfachen Wurzeln $3, \sqrt{2}, -\sqrt{2}, i, -i$.
Das in diesem Beispiel behandelte Polynom tritt z.B. auf als das charakteristische Polynom der 5×5-Matrix

$$\begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3}\sqrt{2} & \frac{4}{3} & 0 & 0 \\ 0 & \frac{4}{3} & -\frac{1}{3}\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix}$$

Je nach zugrundeliegendem Körper hat sie also die Eigenwerte:

$$\begin{aligned} \text{über } \mathbb{Q} &: 3\,, \\ \text{über } \mathbb{R} &: 3, \pm\sqrt{2}\,, \\ \text{über } \mathbb{C} &: 3, \pm\sqrt{2}, \pm i\,. \end{aligned}$$

13.12.10

- Laut Fundamentalsatz besitzt ein Polynom $p(z) \in \mathbb{C}[z]$ mit dem Grad $\deg(p(z)) = n$ eine (bis auf Reihenfolge) eindeutige Faktorisierung (13.4).
- Fasst man hier gleiche Faktoren in Potenzen zusammen, so erhält man eine Zerlegung

$$p(z) = a(z-w_1)^{\nu_1}(z-w_2)^{\nu_2}\cdot\dots\cdot(z-w_\ell)^{\nu_\ell} \tag{13.5}$$

mit paarweise verschiedenen Wurzeln $w_1, \dots, w_\ell \in \mathbb{C}$ und Exponenten $\nu_1, \dots, \nu_\ell \in \mathbb{N}$ mit $\nu_1 + \dots + \nu_\ell = n = \deg(p)$. Auch diese Zerlegung ist bis auf Reihenfolge der Faktoren eindeutig bestimmt.

Definition 13.3.5. *In (13.5) heißt ν_i die Vielfachheit der Wurzel w_i von p. Ist $p(z) = \chi_f(z)$ das charakteristische Polynom eines Endomorphismus f, so ist $\lambda_i = w_i$ Eigenwert von f und ν_i heißt die algebraische Vielfachheit dieses Eigenwerts. i.Z. $a_f(\lambda_i) = \nu_i$.*

Definition 13.3.6. *Eine $n \times n$-Matrix der Gestalt*

$$J = \begin{pmatrix} c & 1 & 0 & 0 & \dots & 0 \\ 0 & c & 1 & 0 & & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & c & 1 & 0 \\ 0 & \dots & \dots & 0 & c & 1 \\ 0 & \dots & \dots & 0 & 0 & c \end{pmatrix}$$

mit $c \in K$ heißt nach Marie Ennemond Camille Jordan (1838 - 1922) Jordan-Block zu c.

Beispiel 13.3.7. *Da der Jordan-Block zu c in oberer Dreiecksgestalt vorliegt, bekommt man das charakteristische Polynom einfach und schon faktorisiert geliefert:*

$$\chi_J(\lambda) = \det \begin{pmatrix} c-\lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & c-\lambda & 1 & 0 & & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & c-\lambda & 1 & 0 \\ 0 & \dots & \dots & 0 & c-\lambda & 1 \\ 0 & \dots & \dots & 0 & 0 & c-\lambda \end{pmatrix} = (c-\lambda)^n ,$$

d.h. c ist der einzige Eigenwert von J mit der algebraischen Vielfachheit $a_J(c) = n$.
Zur Berechnung des zugehörigen Eigenraums setzt man $\lambda = c$ ein und betrachtet das LGS

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 1 & 0 \\ 0 & \dots & \dots & 0 & 0 & 1 \\ 0 & \dots & \dots & 0 & 0 & 0 \end{pmatrix} \cdot v = 0 .$$

Da dieses LGS schon in Zeilenstufenform vorliegt, ist klar, dass es eine eindimensionale Lösung gibt, nämlich

$$v = \mathbb{R} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad bzw. \quad E_J(c) = \mathrm{Lin}\left(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right) .$$

Die geometrische Vielfachheit des Eigenwerts c ist also $g_J(c) = 1$.
Mit diesem 1-dimensionalen Eigenraum bekommt man im Fall $n > 1$ sicher nicht genug linear unabhängige Vektoren für eine Basis von V zusammen, d.h. Jordan-Blöcke sind für $n > 1$ nicht diagonalisierbar.

Satz 13.3.8. *Es seien K ein Körper, V ein endlichdimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$. Der Endomorphismus f ist genau dann diagonalisierbar, wenn die beiden folgenden Punkte erfüllt sind*

(i) Das charakteristische Polynom $\chi_f(\lambda)$ zerfällt über K in Linearfaktoren.

(ii) Für jeden Eigenwert λ_i von f gilt:

$$\boxed{\text{Alg. Vielfachheit } a_f(\lambda_i) = g_f(\lambda_i) \text{ Geom. Vielfachheit}}$$

Beweis: „⇒“ (i) Ist f diagonalisierbar, so gibt es eine Basis B von V mit

$$\begin{aligned} [f]_B &= \mathrm{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{n_1 \text{ Stück}}, \underbrace{\lambda_2, \dots, \lambda_2}_{n_2 \text{ Stück}}, \dots, \underbrace{\lambda_k, \dots, \lambda_k}_{n_k \text{ Stück}}) \\ &\text{mit } \lambda_1, \dots, \lambda_k \in K \text{ für } k \le n \\ &\text{und } \lambda_i \neq \lambda_j \text{ für } i \neq j \\ &\text{und } n_1, \dots, n_k \in \mathbb{N} \text{ mit } n_1 + \dots + n_k = n . \end{aligned}$$

Das zugehörige charakteristische Polynom ist dann

$$\chi_f(\lambda) = (\lambda_1 - \lambda)^{n_1} \cdot \ldots \cdot (\lambda_k - \lambda)^{n_k} ,$$

was Teil (i) mit $n_1 = a_f(\lambda_1)$, …, $n_k = a_f(\lambda_k)$ zeigt. „$\Rightarrow$“ (ii) Der Einfachheit halber wird (ii) für λ_1 nachgewiesen. Für die anderen Eigenwerte geht das genauso. Wegen

$$[f]_B - \lambda_1 \mathbb{1}_n = \operatorname{diag}(\underbrace{0, \ldots, 0}_{a_f(\lambda_1)}, \underbrace{\lambda_2 - \lambda_1, \ldots, \lambda_2 - \lambda_1}_{a_f(\lambda_2)}, \ldots, \underbrace{\lambda_k - \lambda_1, \ldots, \lambda_k - \lambda_1}_{a_f(\lambda_k)})$$

und $\lambda_i \neq \lambda_j$ für $i \neq j$ ist der Rang $n - a_f(\lambda_1)$, d.h. die Lösung des LGS $([f]_B - \lambda_1 \mathbb{1}_n)v_1 = 0$, das ist gerade $E_f(\lambda_1)$, hat die Dimension $a_f(\lambda_1)$.
Somit ist die Vielfachheit $a_f(\lambda_1)$ des Eigenwerts λ_1 im charakteristischen Polynom (seine algebraische Vielfachheit) gleich der Dimension des zugehörigen Eigenraums (der geometrischen Vielfachheit $g_f(\lambda_1)$).
„$\Leftarrow$“ Nach (i) gibt es die Eigenwerte $\lambda_1, \ldots, \lambda_k$ mit den Vielfachheiten $a_f(\lambda_1), \ldots, a_f(\lambda_k)$ und $a_f(\lambda_1) + \cdots + a_f(\lambda_k) = n$, d.h. zählt man die Eigenwerte entsprechend ihrer Vielfachheit, so gibt es gerade n Stück. Die zugehörigen Eigenräume $E_f(\lambda_i)$ haben nach (ii) die Dimension $g_f(\lambda_i) = a_f(\lambda_i)$. Man kann also zu jedem dieser Eigenräume eine entsprechend große Basis angeben:

$$\begin{aligned}
(v_1^{(1)}, \ldots, v_{a_f(\lambda_1)}^{(1)}) &\text{ sei eine geordnete Basis von } E_f(\lambda_1) ,\\
(v_1^{(2)}, \ldots, v_{a_f(\lambda_2)}^{(2)}) &\text{ sei eine geordnete Basis von } E_f(\lambda_2) ,\\
&\vdots\\
(v_1^{(k)}, \ldots, v_{a_f(\lambda_k)}^{(k)}) &\text{ sei eine geordnete Basis von } E_f(\lambda_k) .
\end{aligned}$$

Es genügt zu zeigen, dass

$$\{v_1^{(1)}, \ldots, v_{a_f(\lambda_1)}^{(1)}, v_1^{(2)}, \ldots, v_{a_f(\lambda_2)}^{(2)}, \ldots, v_1^{(k)}, \ldots, v_{a_f(\lambda_k)}^{(k)}\}$$

eine Basis von V ist. Da das n Vektoren sind, reicht der Nachweis der linearen Unabhängigkeit. Dazu betrachtet man wie üblich die Linearkombination

$$c_1^{(1)} v_1^{(1)} + \cdots + c_{a_f(\lambda_1)}^{(1)} v_{a_f(\lambda_1)}^{(1)} + \cdots + c_1^{(k)} v_1^{(k)} + \cdots + c_{a_f(\lambda_k)}^{(k)} v_{a_f(\lambda_k)}^{(k)} = 0 \tag{13.6}$$

Wendet man die Abbildung $f - \lambda_\ell \mathrm{id}_V$ auf einen Eigenvektor $v^{(j)} \in E_f(\lambda_j)$ an, so ergibt sich

$$(f - \lambda_\ell \mathrm{id}_V)(v^{(j)}) = \lambda_j v^{(j)} - \lambda_\ell v^{(j)} = (\lambda_j - \lambda_\ell) v^{(j)}$$

Macht man das mehrfach hintereinander, so ergibt sich

$$\begin{aligned}
(f - \lambda_2 \mathrm{id}_V) \circ \cdots \circ (f - \lambda_k \mathrm{id}_V)(v^{(j)}) &= (\lambda_j - \lambda_2) \cdot \ldots \cdot (\lambda_j - \lambda_k) v^{(j)}\\
&\begin{cases} = 0 & \text{falls } j > 1 ,\\ \neq 0 & \text{falls } j = 1. \end{cases}
\end{aligned} \tag{13.7}$$

Dies kann man ausnutzen, indem man die Abbildung

$$(f - \lambda_2 \mathrm{id}_V) \circ \ldots \circ (f - \lambda_k \mathrm{id}_V)$$

beidseitig auf Gleichung (13.6) anwendet. Zusammen mit (13.7) führt das auf

$$\begin{aligned}
(\lambda_1 - \lambda_2) \cdot \ldots \cdot (\lambda_1 - \lambda_k)(c_1^{(1)} v_1^{(1)} + \cdots + c_{a_f(\lambda_1)}^{(1)} v_{a_f(\lambda_1)}^{(1)}) = 0 &\iff\\
c_1^{(1)} v_1^{(1)} + \cdots + c_{a_f(\lambda_1)}^{(1)} v_{a_f(\lambda_1)}^{(1)} = 0 &\iff\\
c_1^{(1)} = \cdots = c_{a_f(\lambda_1)}^{(1)} = 0&
\end{aligned}$$

Analog kann man etwa $(f - \lambda_1 \mathrm{id}_V)(f - \lambda_3 \mathrm{id}_V) \circ \cdots \circ (f - \lambda_k \mathrm{id}_V)$ beidseitig auf (13.6) anwenden und erhält

$$c_1^{(2)} = \cdots = c_{a_f(\lambda_2)}^{(2)} = 0$$

usw. d.h. alle $c_{\nu_i}^{(j)}$ verschwinden für $j = 1, \ldots, k$ und jeweils $i = 1, \ldots, \nu_j$, d.h.
$\{v_1^{(1)}, \ldots, v_{a_f(\lambda_1)}^{(1)}, v_1^{(2)}, \ldots, v_{a_f(\lambda_2)}^{(2)}, \ldots, v_1^{(k)}, \ldots, v_{a_f(\lambda_k)}^{(k)}\}$ ist eine Basis von V aus Eigenvektoren, f ist also diagonalisierbar! $\square$

13.4 Jordan-Normalform

Beispiel 13.4.1. *Der Jordan-Block*

$$J = \begin{pmatrix} c & 1 & 0 & 0 & \dots & 0 \\ 0 & c & 1 & 0 & & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & c & 1 & 0 \\ 0 & \dots & \dots & 0 & c & 1 \\ 0 & \dots & \dots & 0 & 0 & c \end{pmatrix}$$

aus 13.3.7 hat den n-fachen Eigenwert c, d.h. es ist $a_J(c) = n$. Zu diesem gibt es aber nur einen eindimensionalen Eigenraum $E_J(c)$, d.h. es gilt

$$\boxed{\textit{Geom. Vielfachheit } g_J(c) = 1 < \textit{Alg. Vielfachheit } a_J(c) = n}$$

für $n > 1$, d.h. J ist nicht diagonalisierbar.

Folgerung 13.4.2. *Aus dem Beweis von 13.3.8 folgt, dass für jeden Eigenwert λ gilt*

$$\boxed{1 \leq \textit{Geometrische Vielfachheit } g_f(\lambda) \leq \textit{Algebraische Vielfachheit } a_f(\lambda)}$$

Beispiel 13.4.3. *Gegeben sei der Endomorphismus $f \in \mathrm{End}(\mathbb{R}^3)$ durch die Abbildungsmatrix*

$$\begin{pmatrix} 10 & 18 & 42 \\ -8 & -17 & -47 \\ 2 & 5 & 15 \end{pmatrix} .$$

Das charakteristische Polynom ist

$$\chi_f(\lambda) = -\lambda^3 + 8\lambda^2 - 20\lambda + 16 = -(\lambda - 4)(\lambda - 2)^2 ,$$

*f hat den doppelten Eigenwert 2, also $a_f(2) = 2$ und den einfachen Eigenwert 4, also $a_f(4) = 1$.
Die zugehörigen Eigenräume sind*

$$E_f(2) = \mathbb{R}\begin{pmatrix} 6 \\ -5 \\ 1 \end{pmatrix} , \; E_f(4) = \mathbb{R}\begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} ,$$

*d.h. die geometrischen Vielfachheiten der beiden Eigenwerte sind $g_f(2) = 1 = g_f(4)$.
Für den doppelten Eigenwert 2 ist also die geometrische Vielfachheit zu klein, d.h. f ist nach dem vorhergehenden Satz nicht diagonalisierbar. Es gibt somit keine Basis des $\mathbb{R}^3$ aus Eigenvektoren von f. Trotzdem kann es sich lohnen, wenn man so viele linear unabhängige Eigenvektoren nimmt, wie man bekommen kann (im vorliegenden Fall 2), und zu einer Basis des $\mathbb{R}^3$ ergänzt. Im vorliegenden Fall kann man z.B. noch als dritten Basisvektor e_2 ergänzen.
Bezüglich der neuen Basis*

$$B = \left(\begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} , \begin{pmatrix} 6 \\ -5 \\ 1 \end{pmatrix} , \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right)$$

gilt dann

$$[f]_B = \begin{pmatrix} 4 & 0 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix} ,$$

*was immer noch deutlich besser als die ursprünglich gegebene Matrix ist.
Noch etwas schöner wird die Darstellung, wenn man die beiden Eigenvektoren wie folgt zu einer Basis ergänzt*

$$B' = \left(\begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} , \begin{pmatrix} 6 \\ -5 \\ 1 \end{pmatrix} , \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} \right) ,$$

denn bezüglich dieser Basis gilt

$$[f]_{B'} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} .$$

- Die so genannte *Jordan-Normalform* (letzte Matrix im vorhergehenden Beispiel) ist in vielen Fällen die bestmögliche Darstellung.

Definition 13.4.4. *Es seien K ein Körper und $n \in \mathbb{N}$. Eine Matrix $A \in K^{n \times n}$ heißt Jordan-Matrix, falls es ein $m \in \mathbb{N}$ und Jordan-Blöcke (→13.3.6) $J_1, J_2, \ldots, J_m$ gibt, so dass A von der folgenden Blockdiagonalgestalt ist:*

$$A = \begin{pmatrix} \boxed{J_1} & & & \mathbf{0} \\ & \boxed{J_2} & & \\ & & \ddots & \\ \mathbf{0} & & & \boxed{J_m} \end{pmatrix}$$

- Die Diagonaleinträge in den Matrizen $J_1, J_2, \ldots, J_m$ einer Jordan-Matrix A (→13.4.4) müssen nicht verschieden sein:

Beispiel 13.4.5.

$$\begin{pmatrix} \boxed{2} & & \mathbf{0} \\ & \boxed{\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}} & \\ \mathbf{0} & & \boxed{\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}} \end{pmatrix}, \begin{pmatrix} \boxed{2} & & \mathbf{0} \\ & \boxed{2} & \\ \mathbf{0} & & \boxed{\begin{smallmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{smallmatrix}} \end{pmatrix}, \begin{pmatrix} \boxed{\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}} & \mathbf{0} \\ \mathbf{0} & \boxed{\begin{smallmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{smallmatrix}} \end{pmatrix}$$

sind verschiedene (nicht alle) 5×5-Jordan-Matrizen mit dem 5-fachen Eigenwert 2.

Definition 13.4.6. *Es seien K ein Körper, V ein n-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$. Eine Basis $B = (b_1, b_2, \ldots, b_n)$ von V heißt Jordan-Basis zu f, falls $[f]_B$ eine Jordan-Matrix ist.*

Satz 13.4.7. *Es seien K ein Körper, V ein n-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$. Zerfällt $\chi_f(\lambda)$ über K in Linearfaktoren, so besitzt f eine Jordan-Basis.*

Beweis: Für diesen wichtigen Satz gibt es unzählige Beweise. Auch der kürzeste mir bekannte Beweis[siehe etwa Hartl, 1988] erfordert aber etwa eine Vorlesungswoche und ist noch dazu nicht konstruktiv. Ich gebe hier deshalb nur einen Algorithmus zur Berechnung und keinen Beweis an. □

- In einer Jordan-Normalform stehen in der Diagonalen alle Eigenwerte entsprechend ihrer Vielfachheit.

- Betrachtet man etwa die Matrizen in 13.4.5, so sieht man, dass jeweils die erste Spalte eines Jordan-Blocks zu einem Eigenvektor in der zugehörigen Jordan-Basis gehört.

- Man liest für jeden Eigenwert λ_i ab:

$$\boxed{g_A(\lambda_i) = \text{Anzahl der Jordan-Blöcke zu } \lambda_i.}$$

Definition 13.4.8. *Es seien K ein Körper, $n \in \mathbb{N}$ und $A \in K^{n \times n}$ eine Matrix mit dem Eigenwert $\lambda \in K$. Es sei $A_\lambda := A - \lambda \mathbb{1}_n$ und*

$$H_A(\lambda, k) := \mathrm{Kern}\left(A_\lambda^k\right) \text{ für } k \in \{0, 1, 2, \ldots\}.$$

$H_A(\lambda, k)$ heißt Hauptraum k-ter Stufe von A zum Eigenwert λ, die Elemente von $H_A(\lambda, k) \setminus H_A(\lambda, k-1)$ heißen Hauptvektoren k-ter Stufe von A zum Eigenwert λ.
Insbesondere ist der Hauptraum 1-ter Stufe identisch mit dem Eigenraum, i.Z. $H_A(\lambda, 1) = E_A(\lambda)$, die Hauptvektoren 1-ter Stufe sind Eigenvektoren.

16.12.10

Satz 13.4.9. *Es seien K ein Körper, $n \in \mathbb{N}$ und $A \in K^{n \times n}$ eine Matrix mit dem Eigenwert $\lambda \in K$. Dann gibt es ein $\ell \in \mathbb{N}$ mit*

$$0 = \dim(H_A(\lambda, 0)) < \dim(H_A(\lambda, 1)) < \cdots < \dim(H_A(\lambda, \ell))$$

Ab da wird die Folge der Dimensionen stationär:

$$\dim(H_A(\lambda, \ell)) = \dim(H_A(\lambda, \ell + 1)) = \ldots$$

Beweis: Für jede lineare Abbildung $f_B : K^n \to K^n$ mit $B \in K^{n\times n}$ gilt $\mathrm{Kern}(f_B^i) \subset \mathrm{Kern}(f_B^{i+1})$ bzw. $\mathrm{Kern}(B^i) \subset \mathrm{Kern}(B^{i+1})$, denn für $x \in \mathrm{Kern}(B^i)$ gilt

$$B^{i+1} \cdot x = B \cdot B^i \cdot x = B \cdot 0 = 0 \Rightarrow x \in \mathrm{Kern}(B^{i+1})$$

also $\dim \mathrm{Kern}(B^i) \leq \dim \mathrm{Kern}(B^{i+1})$. Gilt für ein ℓ mal $\dim \mathrm{Kern}(B^\ell) = \dim \mathrm{Kern}(B^{\ell+1})$, also $\mathrm{Kern}(B^\ell) = \mathrm{Kern}(B^{\ell+1})$, so folgt sogar $B^{\ell+1} \cdot x = 0 \iff B^\ell \cdot x = 0$ und damit

$$\begin{aligned} x \in \mathrm{Kern}(B^{\ell+2}) &\iff B^{\ell+2} \cdot x = B \cdot B^{\ell+1} \cdot x = 0 \iff \\ &\iff B^{\ell+1} \cdot x = B \cdot B^\ell \cdot x = 0 \iff x \in \mathrm{Kern}(B^{\ell+1}), \end{aligned}$$

also insbesondere $\dim \mathrm{Kern}(B^{\ell+2}) = \dim \mathrm{Kern}(B^{\ell+1})$, und somit mit Induktion $\dim \mathrm{Kern}(B^{\ell+k}) = \dim \mathrm{Kern}(B^\ell)$ für $k \in \mathbb{N}$.Da alle genannten Räume Unterräume des K^n sind, sind alle hier betrachteten Dimensionen $\leq n$, d.h. die Folge muss in jedem Fall stationär werden. □

- Man kann zeigen, dass die Dimension von $H_A(\lambda, \ell)$ gerade die algebraische Vielfachheit $a_A(\lambda)$ ist.
- Der Hauptraum ℓter Stufe umfasst alle Haupträume anderer Stufen. Er wird deshalb einfach *Hauptraum* von A zum Eigenwert λ genannt (oder *verallgemeinerter Eigenraum*).
- Die Basisvektoren für den zu λ gehörigen Teil der Jordan-Normalform (das ist eine $a_A(\lambda) \times a_A(\lambda)$-Matrix) stammen aus dem Hauptraum zu λ.

Definition 13.4.10. *Es seien K ein Körper, $n \in \mathbb{N}$ und $A \in K^{n\times n}$ eine Matrix mit dem Eigenwert $\lambda \in K$. Für $k \in \{0, 1, 2, \dots\}$ seien $A_\lambda := A - \lambda \mathbb{1}_n$ und $r_k(A, \lambda) := \mathrm{Rang}(A_\lambda^k)$ (oder kurz r_k, wenn keine Verwechslung möglich ist). Weiterhin sei für $k = 1, 2 \dots$:*

$$c_k(A, \lambda) := r_{k+1}(A, \lambda) + r_{k-1}(A, \lambda) - 2r_k(A, \lambda)$$

(bzw. wieder kurz c_k).

Folgerung 13.4.11. *Für jeden Eigenwert λ ist $r_k(A, \lambda) = n - \dim H_A(\lambda, k)$, insbesondere $r_1(A, \lambda) = n - g_A(\lambda)$, d.h. zusammen mit 13.4.9 folgt*

$$n = r_0(A, \lambda) > r_1(A, \lambda) > \dots > r_\ell(A, \lambda) = r_{\ell+1}(A, \lambda) = \dots$$

Satz 13.4.12. *Es seien K ein Körper, $n \in \mathbb{N}$ und $A \in K^{n\times n}$ eine Matrix, deren charakteristisches Polynom $\chi_A(\lambda)$ über K in Linearfaktoren zerfällt. Dann ist für jeden Eigenwert $\lambda_i \in K$ von A*

$$\boxed{c_k(A, \lambda_i) = \text{Anzahl der Jordan-Blöcke der Länge } k \text{ zum Eigenwert } \lambda_i}$$

in einer Jordan-Normalform von A.

Folgerung 13.4.13. *Zusammen mit einer früheren Bemerkung zur Anzahl der Jordan-Blöcke zu einem Eigenwert λ_i heißt das*

$$g_A(\lambda_i) = \sum_{k=1}^{n} c_k(A, \lambda_i)$$

Algorithmus 1: Berechnung der JNF

```
Eingabe : A ∈ K^{n×n}, sodass χ_A(λ) über K in Linearfaktoren zerfällt.
Ausgabe : T ∈ GL(n, K), sodass T^{-1}AT Jordan-Matrix ist.
Beginn
  Setze B := (); Berechne E := {λ ∈ K | λ ist Eigenwert von A};
  für alle λ ∈ E tue
    Berechne c_ℓ = c_ℓ(A, λ) gemäß 13.4.10 für ℓ = 1, 2, ...;
    Setze D_ℓ := ∅ für ℓ = 1, 2, ..., max{ℓ | c_ℓ ≠ 0};
    für ℓ = max{ℓ | c_ℓ ≠ 0}, ..., 2, 1 tue
      Berechne Basen B_{ℓ-1} von H_A(λ, ℓ-1) und B_ℓ von H_A(λ, ℓ);
      Wähle b_1, ..., b_{c_ℓ} ∈ B_ℓ, sodass B_{ℓ-1} ∪ D_ℓ ∪ {b_1, ..., b_{c_ℓ}} lin. unabh.;
      für i = 1, ..., c_ℓ tue
        für j = 1, ..., ℓ tue
          Füge b_i an B an;
          Setze D_{ℓ-j+1} ← D_{ℓ-j+1} ∪ {b_i};
          b_i ← (A - λ · 𝟙_n) b_i;
  Drehe die Reihenfolge von B um;
  T :=Matrix mit den Vektoren aus B als Spalten
```

Beispiel 13.4.14. *Es sei*

$$A := \begin{pmatrix} \frac{134}{125} & \frac{87}{125} & -\frac{329}{150} & \frac{172}{75} & -\frac{49}{150} \\ -\frac{188}{125} & \frac{391}{125} & -\frac{149}{50} & \frac{57}{25} & \frac{31}{50} \\ \frac{8}{75} & -\frac{2}{25} & \frac{131}{45} & -\frac{26}{45} & \frac{1}{45} \\ \frac{16}{75} & -\frac{4}{25} & \frac{52}{45} & \frac{68}{45} & -\frac{13}{45} \\ \frac{16}{75} & -\frac{4}{25} & \frac{22}{45} & \frac{8}{45} & \frac{62}{45} \end{pmatrix} \in \mathbb{R}^{5\times 5}$$

Für diese Matrix gilt

$$\chi_A(z) = -z^5 + 10\,z^4 - 40\,z^3 + 80\,z^2 - 80\,z + 32 = -(z-2)^5\ ,$$

d.h. es ist $a_A(2) = 5$ *und* A *besitzt gemäß* *13.4.7* *eine Jordan-Basis. Die zugehörige Jordan-Normalform hat fünfmal die* 2 *in der Hauptdiagonalen. Zuerst betrachtet man wie üblich* $A_2 := A - 2\cdot\mathbb{1}_5$ *und berechnet daraus* $r_1(A,2) := \mathrm{Rang}(A_2) = 3$*, also* $g_A(2) = 5-3 = 2$*. Aus der erreichten Zeilenstufenform liest man nun den Hauptraum 1.Stufe bzw. Eigenraum*

$$H_A(2,1) := \mathrm{Kern}(A_2) = E_A(2) = \mathrm{Lin}\left(\begin{pmatrix} \frac{15}{16} \\ 0 \\ \frac{1}{2} \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{3}{4} \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\right)$$

ab. Es gibt also $g_A(2) = 2$ *Jordan-Blöcke zum Eigenwert* 2*, die damit die Größen* 1 *und* 4 *oder* 2 *und* 3 *haben müssen. Nun betrachtet man* $A_2^2 := (A - 2\cdot\mathbb{1}_5)^2$ *und berechnet deren Rang* $r_2(A,2) := \mathrm{Rang}(A_2^2) = 1$*. Der zugehörige* $5-1$*-dimensionale Kern dieser Matrix ist der Hauptraum 2.Stufe*

$$H_A(2,2) = \mathrm{Lin}\left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}\right)$$

Wegen $A_2^3 := (A - 2\cdot\mathbb{1}_5)^3 = 0$ *ist der Rang* $r_3(A,2) := \mathrm{Rang}(A_3^2) = 0$ *und somit der Hauptraum 3.Stufe, oder einfach Hauptraum,* $H_A(2,3) = \mathbb{R}^5$*. Damit ist auch* $A_2^k = 0$ *für* $k \geq 3$ *und somit folgt*

$$r_0 = 5 > r_1 = 3 > r_2 = 1 > 0 = r_3 = r_4 = \ldots$$

Mit *13.4.10* *folgt*

$$\begin{aligned} c_1(A,2) =& 5 + 1 - 2\cdot 3 = 0 \\ c_2(A,2) =& 3 + 0 - 2\cdot 1 = 1 \\ c_3(A,2) =& 1 + 0 - 2\cdot 0 = 1 \\ c_4(A,2) =& 0 + 0 - 2\cdot 0 = 0 \\ c_5(A,2) =& 0 + 0 - 2\cdot 0 = 0\,, \ldots \end{aligned}$$

d.h. die Jordan-Normalform hat 1 *Jordan-Block der Größe* 2 *und* 1 *Jordan-Block der Größe* 3*, sieht also aus wie das 3.Beispiel in* *13.4.5*. *Man liest* $\max\{\ell \mid c_\ell \neq 0\} = 3$ *ab, setzt also* $D_\ell := \emptyset$ *für* $\ell = 1,2,3$*. Die* ℓ*-Schleife des Algorithmus startet deshalb mit* $\ell = 3$ *mit den Basen* B_3 *und* B_2 *von* $H_A(2,3)$ *und* $H_A(2,2)$*:*

$$\overbrace{\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right\}}^{B_3=}, \overbrace{\left\{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}\right\}}^{B_2=}$$

Nun kommt der wesentliche Teil des Algorithmus: Es sind $b_1, \ldots, b_{c_\ell} \in B_\ell$ zu wählen, sodass $B_{\ell-1} \cup D_\ell \cup \{b_1, \ldots, b_{c_\ell}\}$ linear unabhängig ist. Wegen $c_3 = 1$ und $D_3 = \emptyset$, ist hier also ein $b_1 \in B_3$ zu wählen, sodass $B_2 \cup \{b_1\}$ linear unabhängig ist. $b_1 = e_3$ leistet das Gewünschte.
Mit diesem $b_1 \in H_A(2,3) \setminus H_A(2,2)$ kommt man nun in die Doppelschleife für i und j. Dort wird b_1 zuerst zu der bisher leeren geordneten Basis B hinzugefügt, d.h. wir haben jetzt $B = (e_3)$ und außerdem merkt man sich b_1 in der bisher leeren Menge D_3, d.h. jetzt ist $D_3 = \{e_3\}$, damit man in späteren Durchläufen noch weiß, dass dieser Vektor aus $H_A(2,3)$ schon verbraucht ist.
In der inneren j-Schleife werden nun die b_i mehrfach mit $A - \lambda \cdot \mathbb{1}_n$ multipliziert. Das Ergebnis wird im Programm jeweils auch wieder b_i genannt, was vielleicht etwas verwirrend ist. Zur Unterscheidung der Vektoren werden die (nur hier) mit $b_i, b_i', b_i'', \ldots$ bezeichnet. Man nennt $b_i, b_i', b_i'', \ldots$ auch eine Jordan-Kette zu λ.
Wegen $b_i \in H_A(\lambda, k) = \mathrm{Kern}\left((A - \lambda \cdot \mathbb{1}_n)^k\right)$ folgt

$$b_i' = (A - \lambda \cdot \mathbb{1}_n) \cdot b_i \in \mathrm{Kern}\left((A - \lambda \cdot \mathbb{1}_n)^{k-1}\right).$$

Für die berechneten Vektoren gilt

$$\begin{aligned} b_i' &= A \cdot b_i - \lambda \cdot b_i \iff b_i' + \lambda \cdot b_i = A \cdot b_i \\ b_i'' &= A \cdot b_i' - \lambda \cdot b_i' \iff b_i'' + \lambda \cdot b_i' = A \cdot b_i' \\ &\vdots \end{aligned}$$

Der letzte berechnete Vektor in dieser Kette ist in $\mathrm{Kern}(A - \lambda \cdot \mathbb{1}_n) = E_A(\lambda)$.
Nimmt man also $\ldots, b_i'', b_i', b_i$ in der Reihenfolge in eine Basis, so gehört zu diesen Basisvektoren der Jordan-Block

$$\begin{pmatrix} \lambda & 1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & 0 & & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda & 1 & 0 \\ 0 & \ldots & \ldots & 0 & \lambda & 1 \\ 0 & \ldots & \ldots & 0 & 0 & \lambda \end{pmatrix}$$

Im vorliegenden Fall ergibt sich

$$b_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \in H_A(2,3) \setminus H_A(2,2) \Rightarrow b_1' = \begin{pmatrix} -\frac{329}{150} \\ -\frac{149}{50} \\ \frac{41}{45} \\ \frac{52}{45} \\ \frac{22}{45} \end{pmatrix} \in H_A(2,2) \setminus H_A(2,1) \Rightarrow b_1'' = \begin{pmatrix} \frac{34}{75} \\ \frac{4}{25} \\ \frac{8}{45} \\ \frac{16}{45} \\ \frac{16}{45} \end{pmatrix} \in H_A(2,1) \setminus H_A(2,0)$$

Auch b_1' und b_1'' merkt man sich in den bisher leeren Mengen D_2 und D_1, d.h. jetzt ist $D_2 = \{b_1'\}$ und $D_1 = \{b_1''\}$. Außerdem fügt man b_1' und b_1'' in der berechneten Reihenfolge der geordneten Basis B zu, d.h. jetzt ist $B = (b_1, b_1', b_1'')$.
Damit ist die ℓ-Schleife des Algorithmus einmal durchlaufen und es geht mit $\ell = 2$ weiter. Zuerst wird eine Basis B_1 von $H_A(2,1) = \mathrm{Kern}(A - 2 \cdot \mathbb{1}_5) = E_A(2)$ *berechnet, etwa*

$$B_1 = \{ \begin{pmatrix} \frac{15}{16} \\ 0 \\ \frac{1}{2} \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{3}{4} \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \}$$

Wegen $c_2(A,2) = 1$ ist nun $b_1 \in B_2$ zu wählen, sodass $B_1 \cup D_2 \cup \{b_1\}$ linear unabhängig ist.
Mit den bereits bekannten D_2 und B_2 heißt dies

$$b_1 \in \{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \} \wedge \{ \begin{pmatrix} \frac{15}{16} \\ 0 \\ \frac{1}{2} \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{3}{4} \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{329}{150} \\ -\frac{149}{50} \\ \frac{41}{45} \\ \frac{52}{45} \\ \frac{22}{45} \end{pmatrix}, b_1 \} \ \textit{lin. unabh.}$$

Der Vektor $b_1 = {}^t\!\begin{pmatrix}0 & 0 & 1 & 1 & 0\end{pmatrix}$ *leistet das Gewünschte.*
Dieser Vektor wird an die geordneten Basis B *angehängt und zu der Menge* D_2 *hinzugefügt.*
Außerdem wird $b_1' = (A - 2 \cdot \mathbb{1}_5) \cdot b_1 \in E_A(2)$ *berechnet und an* B *angehängt und zu* D_1 *hinzugefügt. Die 3 inneren Schleifen für* ℓ, i *und* j *sind damit abgearbeitet, denn für* $\ell = 1$ *ist wegen* $c_1(A, 2) = 0$ *nichts mehr zu tun.*
Es ist der folgende Endstand erreicht

$$D_1 = \left\{ \begin{pmatrix} \frac{1}{10} \\ -\frac{7}{10} \\ \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}, \begin{pmatrix} \frac{34}{75} \\ \frac{4}{25} \\ \frac{8}{45} \\ \frac{16}{45} \\ \frac{16}{45} \end{pmatrix} \right\}, D_2 = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{329}{150} \\ -\frac{149}{50} \\ \frac{41}{45} \\ \frac{52}{45} \\ \frac{22}{45} \end{pmatrix} \right\}, D_3 = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$B = \left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{329}{150} \\ -\frac{149}{50} \\ \frac{41}{45} \\ \frac{52}{45} \\ \frac{22}{45} \end{pmatrix}, \begin{pmatrix} \frac{34}{75} \\ \frac{4}{25} \\ \frac{8}{45} \\ \frac{16}{45} \\ \frac{16}{45} \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{10} \\ -\frac{7}{10} \\ \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} \right)$$

Die zu berechnende Jordan-Basis ergibt sich jetzt durch Umkehr der Reihenfolge aus B *und die Transformationsmatrix* T *für den Basiswechsel hat die neuen Basisvektoren als Spalten, also*

$$T = \begin{pmatrix} \frac{1}{10} & 0 & \frac{34}{75} & -\frac{329}{150} & 0 \\ -\frac{7}{10} & 0 & \frac{4}{25} & -\frac{149}{50} & 0 \\ \frac{1}{3} & 1 & \frac{8}{45} & \frac{41}{45} & 1 \\ \frac{2}{3} & 1 & \frac{16}{45} & \frac{52}{45} & 0 \\ \frac{2}{3} & 0 & \frac{16}{45} & \frac{22}{45} & 0 \end{pmatrix}$$

Mit diesem T *gilt endlich*

$$T^{-1} A T = \begin{pmatrix} \boxed{\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}} & \mathbf{0} \\ \mathbf{0} & \boxed{\begin{smallmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{smallmatrix}} \end{pmatrix}.$$

20.12.10

Beispiel 13.4.15. *Wie so oft bei Algorithmen, gibt es natürlich alle möglichen Sonderfälle, in denen das einfacher und schneller geht. Ein besonderer Fall wird in diesem Beispiel gezeigt: Es sei*

$$A := \begin{pmatrix} \frac{1517}{1875} & \frac{452}{625} & -\frac{299}{750} & \frac{431}{150} & -\frac{689}{250} \\ -\frac{998}{625} & \frac{1961}{625} & -\frac{219}{250} & \frac{161}{50} & -\frac{427}{250} \\ \frac{4}{1125} & -\frac{26}{375} & \frac{127}{75} & -\frac{29}{45} & \frac{73}{225} \\ \frac{8}{1125} & -\frac{52}{375} & \frac{18}{25} & \frac{17}{45} & \frac{71}{225} \\ \frac{8}{1125} & -\frac{52}{375} & \frac{4}{75} & \frac{2}{45} & \frac{221}{225} \end{pmatrix} \in \mathbb{R}^{5 \times 5}$$

Es ist

$$\chi_A(z) = -z^5 + 7z^4 - 19z^3 + 25z^2 - 16z + 4 = -(z-2)^2 (z-1)^3,$$

d.h. es ist $a_A(2) = 2$, $a_A(1) = 3$. *Die Berechnung der zugehörigen Eigenräume bzw. Haupträume 1.Stufe liefert*

$$E_A(2) = \mathbb{R} \begin{pmatrix} -\frac{87}{20} \\ -\frac{141}{20} \\ 1/2 \\ 1 \\ 1 \end{pmatrix}, \quad E_A(1) = \mathbb{R} \begin{pmatrix} \frac{51}{40} \\ \frac{9}{20} \\ 1/2 \\ 1 \\ 1 \end{pmatrix}$$

Da beide Eigenräume eindimensional sind, gibt es zu jedem Eigenwert einen Jordan-Block, d.h. die Gestalt der Jordan-Normalform steht fest:

$$\begin{pmatrix} \boxed{\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}} & \mathbf{0} \\ \mathbf{0} & \boxed{\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{smallmatrix}} \end{pmatrix}$$

Ist $B = \{b_1, b_2, b_3, b_4, b_5\}$ eine zu dieser Jordan-Normalform gehörige Jordan-Basis, so gilt $b_1 \in E_A(2)$ und $b_3 \in E_A(1)$. An den weiteren Spalten der Jordan-Normalform liest man nun ab

$$Ab_2 = b_1 + 2b_2\,,\; Ab_4 = b_3 + b_4\,,\; Ab_5 = b_4 + b_5$$

Somit kann man geeignete Hauptvektoren als Lösungen einfacher linearer Gleichungssysteme berechnen:

$$\begin{aligned} Ab_2 = b_1 + 2b_2 \quad &\Longleftrightarrow \quad (A - 2\mathbb{1}_5)b_2 = b_1 \quad \Longleftrightarrow \quad A_2 b_2 = b_1 \\ Ab_4 = b_3 + b_4 \quad &\Longleftrightarrow \quad (A - \mathbb{1}_5)b_4 = b_3 \quad \Longleftrightarrow \quad A_1 b_4 = b_3 \\ Ab_5 = b_4 + b_5 \quad &\Longleftrightarrow \quad (A - \mathbb{1}_5)b_5 = b_4 \quad \Longleftrightarrow \quad A_1 b_5 = b_4\,. \end{aligned}$$

Diese inhomogenen linearen Gleichungssysteme haben jeweils mindestens 1-dimensionale Lösungsräume, aus denen man sich jeweils nur eine Lösung auswählt. Hier erhält man z.B.

$$b_2 = \begin{pmatrix} -\frac{3}{4} \\ -\frac{29}{4} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \; b_4 = \begin{pmatrix} \frac{375}{32} \\ 0 \\ \frac{35}{4} \\ \frac{125}{8} \\ \frac{55}{4} \end{pmatrix}, \; b_5 = \begin{pmatrix} -\frac{28467}{256} \\ -\frac{12303}{128} \\ \frac{945}{64} \\ 0 \\ -\frac{765}{32} \end{pmatrix}$$

und damit die Transformationsmatrix

$$T = \begin{pmatrix} -\frac{87}{20} & -3/4 & \frac{51}{40} & \frac{375}{32} & -\frac{28467}{256} \\ -\frac{141}{20} & -\frac{29}{4} & \frac{9}{20} & 0 & -\frac{12303}{128} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{35}{4} & \frac{945}{64} \\ 1 & 0 & 1 & \frac{125}{8} & 0 \\ 1 & 0 & 1 & \frac{55}{4} & -\frac{765}{32} \end{pmatrix}$$

Dieses Verfahren funktioniert hier nur deshalb, weil es zu beiden Eigenwerten jeweils nur einen Jordan-Block und somit jeweils nur eine Jordan-Kette gibt.
Somit kann man bei einem Eigenvektor starten und die Jordan-Kette in anderer Richtung als im Algorithmus durchlaufen.
Hat man dagegen mehrdimensionale Eigenräume, so geht dieses Verfahren nicht mehr so direkt, weil man nicht weiß, bei welchem Eigenvektor man starten kann. So ist z.B. für

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

der Eigenraum $E_A(2) = \mathrm{Lin}\,(e_1 + e_3, e_3)$, aber zu keinem der beiden Vektoren $b = e_1 + e_3$ oder $b = e_3$ das LGS $A_2 x = b$ lösbar. 23.12.10

14 Bilinearformen

14.1 Matrixdarstellung

- Es seien $\varphi : V \times V \to K; (x,y) \mapsto \varphi(x,y)$ eine Bilinearform (→12.1.1) auf einem K-Vektorraum V endlicher Dimension und $B = (b_1, \dots, b_n)$ eine geordnete Basis von V.
- Betrachtet man analog zu linearen Abbildungen die Bilder der Basisvektoren $a_{ij} := \varphi(b_i, b_j)$, so gilt dann wegen der Bilinearität für zwei beliebige Vektoren $v = \sum_{i=1}^n v_i b_i$ und $w = \sum_{j=1}^n w_j b_j$ aus V
$$\varphi(v,w) = \varphi(\sum_{i=1}^n v_i b_i, \sum_{j=1}^n w_j b_j) = \sum_{i=1}^n v_i \sum_{j=1}^n a_{ij} w_j \,, \tag{14.1}$$
d.h. φ ist durch die Bilder $a_{ij} := \varphi(b_i, b_j)$ der Basisvektoren bereits eindeutig festgelegt.

Definition 14.1.1. *Es seien K ein Körper, $n \in \mathbb{N}$, V ein n-dimensionaler K-Vektorraum, $B = (b_1, \dots, b_n)$ eine geordnete Basis von V und $\varphi : V \times V \to K; (x,y) \mapsto \varphi(x,y)$ eine Bilinearform. Die Matrix*
$$G_B(\varphi) := (\varphi(b_i, b_j))_{1 \le i,j \le n}$$
heißt Grammatrix (auch Formmatrix oder Strukturmatrix) von φ.

- (14.1) zeigt nun, dass man mit Hilfe der Grammatrix schreiben kann:
$$\varphi(v,w) = {}^t\!\left(v_{/B}\right) \cdot G_B(\varphi) \cdot w_{/B} \,.$$
- Wegen der bekannten Rechengesetze für Matrizen prüft man nun leicht nach, dass anders herum für jede beliebige Matrix $A \in K^{n \times n}$ die Abbildung $(v,w) \mapsto {}^t\!\left(v_{/B}\right) \cdot A \cdot w_{/B}$ eine Bilinearform des K^n ist.
- Diese heißt dann *die zu A gehörige Bilinearform.*

Beispiel 14.1.2. *Mit der Grammatrix $A = \mathbb{1}_n$ im $\mathbb{R}^n$ erhält man das so genannte Standard-Skalarprodukt (oder kanonische reelle Skalarprodukt)*
$${}^t v A w = {}^t v \mathbb{1}_n w = {}^t v w = v_1 w_1 + v_2 w_2 + \dots + v_n w_n \,,$$
d.h. dieses Skalarprodukt ist eine spezielle Bilinearform.

Satz 14.1.3. *Eine Bilinearform φ eines K-Vektorraumes V heißt nicht ausgeartet, wenn für festes $v \in V$ gilt*
$$\varphi(v,w) = 0 \text{ für alle } w \in V \Rightarrow v = 0 \,.$$

Satz 14.1.4. *Es seien K ein Körper, $n \in \mathbb{N}$, V ein n-dimensionaler K-Vektorraum, $B = (b_1, \dots, b_n)$ eine geordnete Basis von V und $\varphi : V \times V \to K; (x,y) \mapsto \varphi(x,y)$ eine Bilinearform. Dann gilt*
$$\operatorname{Rang} G_B(\varphi) = n \iff \varphi \text{ ist nicht ausgeartet.}$$

Beweis: „⇐“Angenommen, $\operatorname{Rang} G_B(\varphi) = m < n$. Dann ist mit $w \in V$ die Menge der Vektoren $G_B(\varphi)\, w_{/B}$ ein m-dimensionaler Teilraum des K^n. Ist $C = (c_1, \dots, c_m)$ eine Basis dieses Raumes, so ist $\varphi(v,w) = 0$ für alle $w \in V$ gleichbedeutend mit
$${}^t\!\left(v_{/B}\right) \cdot \begin{pmatrix} c_1 & c_2 & \dots & c_m \end{pmatrix} = 0 \iff \begin{pmatrix} {}^t(c_1) \\ \vdots \\ {}^t(c_m) \end{pmatrix} \cdot v_{/B} = 0 \,.$$
Da die Koeffizientenmatrix dieses LGS den Rang $m < n$ hat, gibt es einen $n-m$-dimensionalen Lösungsraum für die v's, also insbesondere immer von 0 verschiedene Lösungen, d.h. φ ist ausgeartet.
„⇒“ Ist $\operatorname{Rang} G_B(\varphi) = n$, so ist die Menge aller $G_B(\varphi)\, w_{/B}$ ganz V. Deshalb ist $\varphi(v,w) = 0$ für alle $w \in V$ gleichbedeutend mit ${}^t\!\left(v_{/B}\right) \cdot \mathbb{1}_n = 0$, d.h. es gibt nur die Lösung $v = 0$ und φ ist nicht ausgeartet. □

14.2 Basiswechsel

- Wie bei linearen Abbildungen interessiert man sich nun für das Verhalten der Grammatrix beim Basiswechsel und für Basen bezüglich derer die Grammatrix besonders einfach aussieht.

- Der Basiswechsel von einer Basis C von V zu einer Basis B von V wird dabei wie gewohnt von einer (invertierbaren) Transformationsmatrix $T := {}_B[\mathrm{id}_V]_C$ beschrieben:

- $$\begin{aligned} v_{/B} &= {}_B[\mathrm{id}_V]_C v_{/C}\,, \quad w_{/B} = {}_B[\mathrm{id}_V]_C w_{/C} \Rightarrow \\ \varphi(v,w) &= {}^t\!\left(v_{/B}\right) \cdot G_B(\varphi) \cdot w_{/B} = \\ &= {}^t\!\left({}_B[\mathrm{id}_V]_C \cdot v_{/C}\right) \cdot G_B(\varphi) \cdot \left({}_B[\mathrm{id}_V]_C \cdot w_{/C}\right) = \\ &= {}^t\!\left(v_{/C}\right) \cdot \underbrace{{}^t\!\left({}_B[\mathrm{id}_V]_C\right) \cdot G_B(\varphi) \cdot {}_B[\mathrm{id}_V]_C}_{=G_C(\varphi)} \cdot w_{/C}\,. \end{aligned}$$

  oder kurz $G_C(\varphi) = {}^tT \cdot G_B(\varphi) \cdot T$.

Definition 14.2.1. *Zwei Matrizen $A, B \in K^{n\times n}$ heißen kongruent, wenn es ein invertierbares $T \in K^{n\times n}$ mit $A = {}^tT \cdot B \cdot T$ gibt.*

Satz 14.2.2. *Die Kongruenz von Matrizen ist eine Äquivalenzrelation.*

Beweis: Einfache Übung. □

- Genauso wie man sich bei den Äquivalenzklassen ähnlicher Matrizen möglichst einfache Vertreter aussucht (Diagonalgestalt, Dreiecksgestalt oder Jordan-Normalform), stellt sich auch für Äquivalenzklassen kongruenter Matrizen das Normalformenproblem.

- Mit Hilfe von Elementarmatrizen kann man jetzt einiges zu Kongruenzumformungen sagen:

- Ist T eine invertierbare Matrix, so kann man diese als Produkt von Elementarmatrizen schreiben.

- Die Multiplikation einer gegebenen Matrix B von rechts mit der Matrix T bewirkt eine Folge von elementaren Spaltenumformungen und die Multiplikation dieser Matrix mit ${}^tT$ von links bewirkt die analogen elementaren Zeilenumformungen.

- Dieses Wissen nutzt man aus, um eine gegebene Matrix A auf eine einfachere, aber kongruente Gestalt zu bringen:

- Es seien K ein Körper und $A \in K^{n\times n}$.

- An der gegebenen Matrix A werden identische elementare Zeilen- und Spaltenumformungen vorgenommen.

- Die Zeilenumformungen werden gleichzeitig in einer gleichgroßen Einheitsmatrix mitprotokolliert (analog könnte man natürlich auch die elementaren Spaltenumformungen mitprotokollieren):

- $$(A|\mathbb{1}_n) \overset{\substack{\text{Identische elem. Zeilen-} \\ \text{und Spaltenumf. an } A, \\ \text{an } \mathbb{1}_n \text{ nur Zeilenumf.}}}{\rightsquigarrow} (A'|T) \Rightarrow A' = T \cdot A \cdot {}^tT$$

- Ist A die Grammatrix einer Bilinearform φ bezüglich der Basis B, also $A = G_B(\varphi)$, so ist $A' = G_C(\varphi)$ und ${}^tT = {}_B[\mathrm{id}]_C$, d.h. man kann aus T direkt die neue Basis ablesen.

- Für eine spezielle Klasse von Bilinearformen wird das Normalformenproblem für kongruente Matrizen durch den so genannten *Sylvesterschen Trägheitssatz* sehr schön beantwortet werden.

Beispiel 14.2.3.

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{array}\right) \begin{array}{ccc} Z_2 \leftarrow Z_2 - 4Z_1 & \textit{und} & S_2 \leftarrow S_2 - 4S_1 \\ Z_3 \leftarrow Z_3 - 7Z_1 & \textit{und} & S_3 \leftarrow S_3 - 7S_1 \\ & \rightsquigarrow & \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & -2 & -4 & 1 & 0 & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & -6 & -12 & -7 & 0 & 1 \end{array}\right) \begin{array}{ccc} Z_3 \leftarrow Z_3 - 2Z_2 & \underset{\rightsquigarrow}{\textit{und}} & S_3 \leftarrow S_3 - 2S_2 \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & -2 & 0 & 1 & 0 & 0 \\ 0 & -3 & 0 & -4 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right) \Rightarrow$$

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix}}_{=T} \cdot \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}}_{=A=G_B(\varphi)} \cdot \underbrace{\begin{pmatrix} 1 & -4 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}}_{={}^tT={}_B[\mathrm{id}]_C} = \underbrace{\begin{pmatrix} 1 & -2 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{=A'=G_C(\varphi)}$$

14.3 Quadratische Formen

Definition 14.3.1. *Es seien K ein Körper, V ein K-Vektorraum und $\varphi : V \times V \to K$ eine Bilinearform. Die Bilinearform φ heißt symmetrisch, wenn gilt*

$$\varphi(u, v) = \varphi(v, u) \quad \forall u, v \in V\,.$$

Satz 14.3.2. *Es seien K ein Körper, V ein endlichdimensionaler K-Vektorraum und $\varphi : V \times V \to K$ eine Bilinearform. Die Bilinearform φ ist genau dann symmetrisch, wenn die zugehörige Grammatrix bezüglich einer beliebigen Basis B von V symmetrisch ist: φ symmetrisch $\iff G_B(\varphi) = {}^t\big(G_B(\varphi)\big)$.*

Beweis:

$$\forall u, w \in V:$$

$$\left.\begin{array}{rcl} \varphi(u, w) & = & {}^tu \cdot G_B(\varphi) \cdot w = {}^t\big({}^tu \cdot G_B(\varphi) \cdot w\big) = {}^tw \cdot {}^t\big(G_B(\varphi)\big) \cdot u \\ \varphi(w, u) & = & {}^tw \cdot G_B(\varphi) \cdot u \end{array}\right\} \Rightarrow$$

$$G_B(\varphi) = {}^t\big(G_B(\varphi)\big)\,.$$

□

Definition 14.3.3. *Es seien K ein Körper, V ein K-Vektorraum und $\varphi : V \times V \to K$ eine symmetrische Bilinearform. Dann heißt*

$$q_\varphi : \begin{cases} V \to K \\ v \mapsto \varphi(v, v) \end{cases}$$

die zu φ gehörige quadratische Form.

- Aus der Bilinearität von φ folgt für die quadratische Form $q_\varphi(\mu v) = \mu^2 q_\varphi(v)$ für alle $\mu \in K$, $v \in V$; daher der Name.
- Für alle $u, v \in V$ gilt

$$\begin{aligned} q_\varphi(u + v) &= \varphi(u + v, u + v) = \varphi(u, u + v) + \varphi(v, u + v) = \\ &= \varphi(u, u) + \varphi(u, v) + \varphi(v, u) + \varphi(v, v) = \\ &= q_\varphi(u) + q_\varphi(v) + 2\varphi(u, v)\,. \end{aligned} \tag{14.2}$$

- Ist V endlichdimensional und $A = (a_{ij}) = G_B(\varphi)$ die zu φ gehörige symmetrische Formmatrix bezüglich einer festen Basis B von V, so ist

$$q(v) = {}^tvAv = \sum_{i=1}^{n} a_{ii} v_i^2 + \sum_{1 \le i < j \le n} 2a_{ij} v_i v_j\,.$$

- Dies ist ein *Polynom in mehreren Variablen*, also eine Summe von so genannten *Termen* der Gestalt $a v_1^{i_1} v_2^{i_2} \cdot \ldots \cdot v_r^{i_r}$ mit $a \in K$, $r \in \mathbb{N}$ und $i_j \in \mathbb{N}_0$ für $j = 1, \ldots, r$ und Unbestimmten $v_1, \ldots, v_r$.

- Der *Grad* eines solchen Terms ist $i_1 + \dots + i_r$.
- Haben alle Terme eines Polynoms gleichen Grad, so heißt das Polynom *homogen*.
- Eine quadratische Form liefert also ein homogenes Polynom vom Grad 2 (oder quadratisches Polynom).
- Aus $q(v) = \sum_{i=1}^n a_{ii}v_i^2 + \sum_{1\leq i<j\leq n} 2a_{ij}v_iv_j$ liest man ab, dass sich jedes homogene quadratische Polynom umschreiben lässt in die Gestalt ${}^t\!vAv$, also eine quadratische Form ist:
- Die Koeffizienten der *reinquadratischen Terme* v_i^2 liefern für $i = 1, \dots, n$ die Diagonalelemente a_{ii} der Formmatrix.
- Die Koeffizienten der *gemischtquadratischen Terme* v_iv_j werden für $1 \leq i < j \leq n$ je zur Hälfte in a_{ij} und a_{ji} übernommen.

Beispiel 14.3.4. *Das homogene quadratische Polynom* $1x^2 + 1y^2 + 2z^2 + 2xy + 4xz - 1yz$ *ist identisch mit der quadratischen Form*

$$q(X) = \begin{pmatrix} x & y & z \end{pmatrix} A \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{mit} \quad A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & -\frac{1}{2} \\ 2 & -\frac{1}{2} & 2 \end{pmatrix}.$$

Satz 14.3.5. *Es seien K ein Körper mit $1 + 1 \neq 0$, V ein K-Vektorraum und $q : V \to K$ eine quadratische Form. Dann ist*

$$\varphi(u, v) = \frac{1}{2}\left(q(u + v) - q(u) - q(v)\right)$$

die zu q zugehörige symmetrische Bilinearform, d.h. $q = q_\varphi$.

Beweis: Ist $q = q_\psi$ mit der Bilinearform $\psi : V \times V \to K$, so folgt für alle $u, v \in V$

$$\begin{aligned} \varphi(u, v) &= \tfrac{1}{2}\left(q(u + v) - q(u) - q(v)\right) \overset{(14.2)}{=} \\ &= \tfrac{1}{2}\left(q(u) + q(v) + 2\psi(u, v) - q(u) - q(v)\right) = \psi(u, v) \end{aligned}$$

also $\varphi = \psi$. □

Beispiel 14.3.6. *Für symmetrische Bilinearformen kann man mit Kongruenztransformationen eine besonders schöne Form erreichen:*

$$\left(\begin{array}{ccc|ccc} 1 & 4 & 7 & 1 & 0 & 0 \\ 4 & 6 & 8 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{array}\right) \quad \begin{array}{ccc} Z_2 \leftarrow Z_2 - 4Z_1 & \mathit{und} & S_2 \leftarrow S_2 - 4S_1 \\ Z_3 \leftarrow Z_3 - 7Z_1 & \mathit{und} & S_3 \leftarrow S_3 - 7S_1 \\ & \rightsquigarrow & \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -10 & -20 & -4 & 1 & 0 \\ 0 & -20 & -40 & -7 & 0 & 1 \end{array}\right) \quad \begin{array}{ccc} Z_3 \leftarrow Z_3 - 2Z_2 & \underset{\rightsquigarrow}{\mathit{und}} & S_3 \leftarrow S_3 - 2S_2 \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -10 & 0 & -4 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right) \quad \begin{array}{ccc} Z_2 \leftarrow \frac{Z_2}{\sqrt{10}} & \underset{\rightsquigarrow}{\mathit{und}} & S_2 \leftarrow \frac{S_2}{\sqrt{10}} \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -\frac{4}{\sqrt{10}} & \frac{1}{\sqrt{10}} & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right) \Rightarrow$$

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{4}{\sqrt{10}} & \frac{1}{\sqrt{10}} & 0 \\ 1 & -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 4 & 7 \\ 4 & 6 & 8 \\ 7 & 8 & 9 \end{pmatrix} \cdot \begin{pmatrix} 1 & -\frac{4}{\sqrt{10}} & 1 \\ 0 & \frac{1}{\sqrt{10}} & -2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

10.1.11

Satz 14.3.7 (Trägkeitssatz von Sylvester). [15] *Es seien $K = \mathbb{R}$ und φ eine symmetrische Bilinearform des n-dimensionalen $\mathbb{R}$-Vektorraumes V. Dann existiert eine Basis B mit*

$$G_B(\varphi) = \begin{pmatrix} \mathbb{1}_{n_+} & 0 & 0 \\ 0 & -\mathbb{1}_{n_-} & 0 \\ 0 & 0 & 0_{n_0} \end{pmatrix}.$$

[15] Benannt nach *James Joseph Sylvester (1814 - 1897)*, der u.a. in seiner Arbeit On a New Class of Theorems (1850) zum ersten Mal das Wort Matrix im heutigen Sinne verwendete. Dieser Satz hier war allerdings schon davor *Carl Gustav Jacob Jacobi (1804 - 1851)* bekannt.

Dabei seien $\mathbb{1}_{n_+}$ bzw. $\mathbb{1}_{n_-}$ Einheitsmatrizen der Größe n_+ bzw. n_- und 0_{n_0} eine $n_0 \times n_0$-Nullmatrix. Die restlichen Nullen bezeichnen Nullmatrizen passender Größe. Die Zahlen $n_+, n_-, n_0 \in \mathbb{N}$ $(n_+ + n_- + n_0 = n)$ sind durch die Bilinearform eindeutig bestimmt.

Beweis: Der Beweis wird durch Induktion nach der Dimension n von V geführt. Für $n = 0$ ist die Aussage trivial. Nun sei $n \geq 1$ und die Behauptung für $n-1$ bewiesen.Ist $\varphi(v,v) = 0$ für alle $v \in V$, so gilt wegen 14.3.5 auch

$$\varphi(v,w) = \frac{1}{2}(\varphi(v+w, v+w) - \varphi(v,v) - \varphi(w,w)) = 0$$

für alle $v, w \in V$, d.h. φ ist die Nullform und $G_B(\varphi) = 0$ für jede Basis B von V.
Gibt es dagegen ein $v_1 \in V$ mit $\varphi(v_1, v_1) \neq 0$, so setzt man $w_1 := \frac{1}{\sqrt{|\varphi(v_1,v_1)|}} v_1$. Wegen der Bilinearität von φ folgt

$$\varphi(w_1, w_1) = \frac{\varphi(v_1, v_1)}{|\varphi(v_1, v_1)|} = \pm 1 \,.$$

Nun sei $W := \{w \in V \mid \varphi(w_1, w) = 0\}$. Da φ eine Bilinearform von V ist, ist $\varphi(w_1, w)$ eine Linearform von V, also W der Kern dieser Linearform. Insbesondere ist W ein Untervektorraum von V. Wegen $\varphi(w_1, w_1) = \pm 1$ ist $\varphi(w_1, w)$ nicht die Nullform, d.h. es ist $\operatorname{Bild} \varphi(w_1, w) = \mathbb{R}$ und deshalb $\dim W = n - 1$.
In Übung 8.9 wurde $\operatorname{Lin}(w_1) \oplus W = V$ gezeigt.
Nach Induktionsvoraussetzung hat W eine Basis $(w_2, \ldots, w_n)$ mit den geforderten Eigenschaften. Deshalb ist $B = (w_1, w_2, \ldots, w_n)$ eine Basis von V mit den gesuchten Eigenschaften. □

14.4 Definitheit

Definition 14.4.1. *Es seien $K = \mathbb{R}$ und φ eine symmetrische Bilinearform des n-dimensionalen $\mathbb{R}$-Vektorraumes V. Das nach 14.3.7 eindeutig bestimmte Tripel (n_+, n_-, n_0) heißt die Signatur der symmetrischen Bilinearform φ. Eine reelle symmetrische Bilinearform mit $n_+ = n = \dim V$ heißt positiv definit. Die Matrix $G_B(\varphi)$ aus dem Trägheitssatz heißt die Sylvester-Normalform von φ.*

Beispiel 14.4.2. *Die in 14.3.6 betrachtete Bilinearform hat die Signatur $(1,1,1)$ und ist deshalb nicht positiv definit.*
Die Summe $n_+ + n_-$ ist der Rang jeder Grammatrix von φ, hier also 2.

- Eine symmetrische Bilinearform heißt nach bisheriger Definition *positiv definit*, wenn sie als Sylvester-Normalform die Einheitsmatrix (Signatur $(\dim V, 0, 0)$) besitzt.
- Bezüglich der zugehörigen Basis B gilt dann

  $$\varphi(v,v) = {}^t\!\left(v_{/B}\right) \cdot \mathbb{1}_n \cdot v_{/B} = |v_1|^2 + \cdots + |v_n|^2 \,,$$

  d.h. $\varphi(v,v)$ ist immer positiv und nur für $v = 0$ gilt $\varphi(v,v) = 0$.
- Damit wird die bisherige Definition über die Normalform nun erweitert auf Bilinearformen über beliebigen reellen Vektorräumen:

Definition 14.4.3. *Eine symmetrische Bilinearform φ in einem reellen Vektorraum V (nicht zwingend endlichdimensional) heißt positiv definit, wenn gilt $\varphi(v,v) > 0 \; \forall v \in V \setminus \{0\}$. Analog heißt φ positiv semidefinit, wenn gilt $\varphi(v,v) \geq 0 \; \forall v \in V$, negativ definit, wenn $\varphi(v,v) < 0 \; \forall v \in V \setminus \{0\}$ und negativ semidefinit, wenn $\varphi(v,v) \leq 0 \; \forall v \in V$. Hat φ keine diese Eigenschaften, d.h. gibt es $v \in V$ mit $\varphi(v,v) > 0$ und $w \in V$ mit $\varphi(w,w) < 0$, so heißt φ indefinit.*

- Für eine positiv semidefinite Bilinearform gilt $n_- = 0$, bzw. $n_+ + n_0 = n$.
- Für eine negativ definite Bilinearform ist $n = n_-$, für eine negativ semidefinite $n_+ = 0$, bzw. $n_- + n_0 = n$.
- Für den Nachweis der positiven Definitheit sind damit schon zwei Kriterien bekannt:
- Direktes Nachrechnen von $\varphi(v,v) > 0$ für $v \neq 0$ (z.B. durch *quadratische Ergänzung*),
- Umrechnung auf die Sylvester-Normalform mit Hilfe von Kongruenzumformungen.
- Dabei muss am Schluss nicht unbedingt die genaue Sylvester-Normalform mit $\pm 1, 0$ in der Diagonalen sein; irgendeine Diagonalgestalt und Betrachtung der Vorzeichen in der Diagonalen reicht schon.

Satz 14.4.4. *Es sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum mit Basis B und φ eine symmetrische Bilinearform. Dann besitzt $A := G_B(\varphi)$ genau n reelle (nicht notwendig verschiedene, entsprechend ihrer Vielfachheit gezählte) Eigenwerte, d.h. jede reelle symmetrische $n \times n$-Matrix hat n reelle Eigenwerte!*

Beweis: Ist φ eine symmetrische Bilinearform mit der ebenfalls symmetrischen Formmatrix $A \in \mathbb{R}^{n\times n}$, so hat dieses A nach dem Fundamentalsatz der Algebra 13.3.2 n komplexe Eigenwerte. Ist $\lambda \in \mathbb{C}$ solch ein Eigenwert und $v \in \mathbb{C}^n$ ein zugehöriger Eigenvektor, so gilt

$$Av = \lambda v \iff \overline{Av} = \overline{\lambda v} \iff \overline{A}\overline{v} = {}^t\!A\overline{v} = \overline{\lambda}\overline{v} \iff {}^t(\overline{v})\, A = \overline{\lambda}\, {}^t(\overline{v}) \ .$$

Damit folgt

$$\varphi(\overline{v}, v) = {}^t(\overline{v})\, Av = \left\{ \begin{matrix} {}^t(\overline{v})\,(Av) = {}^t(\overline{v})\,(\lambda v) = \lambda\, {}^t(\overline{v})\, v \\ ({}^t(\overline{v})\, A)v = (\overline{\lambda}\, {}^t(\overline{v}))v = \overline{\lambda}\, {}^t(\overline{v})\, v \end{matrix} \right\} \Rightarrow \overline{\lambda} = \lambda \Rightarrow \lambda \in \mathbb{R}\,,$$

denn ${}^t(\overline{v})\, v = |v_1|^2 + \cdots + |v_n|^2 \neq 0$. □

Satz 14.4.5. *Es sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Dann besitzt das charakteristische Polynom $\chi_A(\lambda) = \sum_{i=0}^{n} \alpha_i \lambda^i$ nur reelle Koeffizienten α_i für $i = 0, \ldots, n$ und n reelle Nullstellen (also reelle Eigenwerte $\lambda_1, \ldots, \lambda_n$) und A ist genau dann positiv definit, wenn eine der folgenden Bedingungen erfüllt ist:*

(1) $\lambda_1, \ldots, \lambda_n > 0$,

(2) $(-1)^j \alpha_j > 0$ *für* $j = 0, \ldots, n-1$,

(3) $\det(A_k) > 0$ *für* $k = 1, \ldots, n$, *wobei A_k die Teilmatrix von A sei, die aus den ersten k Zeilen und Spalten von A besteht.* $\det(A_k)$ *heißt Hauptunterdeterminante oder Hauptminor von A. Dieser Punkt ist auch unter dem Namen Hurwitz-Kriterium bekannt (nach Adolf Hurwitz (1859 - 1919))*

- *Vorsicht*: Nur das 1. Kriterium kann durch Übergang von $>$ zu $\geq$ auf den Nachweis von positiver Semidefinitheit angewendet werden!
- Eine Matrix A ist genau dann negativ definit, wenn $-A$ positiv definit ist.
- Damit lassen sich die aufgelisteten Kriterien auch zum Nachweis der negativen Definitheit verwenden.
- Der Beweis dieses Satzes erfordert einige Hilfsmittel, die in der Vorlesung noch nicht behandelt wurden, und wird deshalb vorerst weggelassen.

15 Euklidische Vektorräume

15.1 Skalarprodukt

Definition 15.1.1. *Es seien V ein $\mathbb{R}$-Vektorraum und $\varphi : V \times V \to \mathbb{R}$ eine positiv definite symmetrische Bilinearform. Dann heißt φ Skalarprodukt auf V und V Euklidischer Vektorraum. Ist eine Bilinearform φ sogar ein Skalarprodukt, so schreibt man auch $\langle x|y\rangle$ statt $\varphi(x,y)$ (in der Schule oder bei Ingenieuren oft $x \cdot y$). Um zu betonen, dass man einen Euklidischen Vektorraum hat, schreibt man auch (V, φ) bzw. $(V, \langle\ |\ \rangle)$.*

Beispiel 15.1.2. *Das bereits aus der Schule bekannte so genannte reelle kanonische Skalarprodukt $\langle v|w\rangle := {}^t v \mathbb{1}_n w = {}^t v w$ im $\mathbb{R}^n$ ist ein Skalarprodukt gemäß obiger Definition, der $(\mathbb{R}^n, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum.*

Beispiel 15.1.3. *Es sei $V = \mathbb{R}[x]_3$ der Vektorraum der reellen Polynome vom Grad höchstens 3. Dann ist*

$$\langle p|q\rangle := \int_{-1}^{1} p(x)q(x)\,dx$$

ein Skalarprodukt, $(V, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum.

15.2 Norm

Definition 15.2.1. *Es sei $(V, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum. Die Abbildung*

$$\|\ \| : \begin{cases} V & \to & \mathbb{R} \\ x & \mapsto & \sqrt{\langle x|x\rangle} \end{cases}$$

heißt die zu $\langle\ |\ \rangle$ gehörige Norm von V (auch Betrag oder Länge genannt).

Beispiel 15.2.2. *Zu dem reellen kanonischen Skalarprodukt im $\mathbb{R}^n$ erhält man so die übliche Euklidische Länge*

$$\|x\| = \sqrt{\langle x|x\rangle} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}\,.$$

Satz 15.2.3 (Die Cauchy-Schwarz-sche Ungleichung)**.** *Es sei $(V, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum. Für $x, y \in V$ gilt*

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

Das Gleichheitszeichen gilt genau dann, wenn x und y linear abhängig sind.

Beweis: Zum Beweis betrachtet man die Norm des Vektors

$$v = \langle x|x\rangle y - \langle y|x\rangle x\,.$$

Es gilt:

$$\begin{aligned}
0 \leq \|v\|^2 &= \langle \langle x|x\rangle y - \langle y|x\rangle x | \langle x|x\rangle y - \langle y|x\rangle x\rangle = \\
&= \langle \langle x|x\rangle y | \langle x|x\rangle y\rangle - \langle \langle y|x\rangle x | \langle x|x\rangle y\rangle - \langle \langle x|x\rangle y | \langle y|x\rangle x\rangle + \langle \langle y|x\rangle x | \langle y|x\rangle x\rangle = \\
&= \langle x|x\rangle^2 \langle y|y\rangle - \langle y|x\rangle\langle x|x\rangle\langle x|y\rangle - \langle x|x\rangle\langle y|x\rangle^2 + \langle y|x\rangle^2\langle x|x\rangle = \\
&= \langle x|x\rangle\langle x|x\rangle\langle y|y\rangle - \langle y|x\rangle\langle x|x\rangle\langle x|y\rangle - \langle x|x\rangle\langle y|x\rangle^2 + \langle y|x\rangle\langle x|y\rangle\langle x|x\rangle = \\
&= \langle x|x\rangle^2\langle y|y\rangle - \langle x|x\rangle\langle y|x\rangle^2 \iff \\
0 \leq & \langle x|x\rangle\langle y|y\rangle - \langle y|x\rangle^2 \iff \\
|\langle y|x\rangle|^2 \leq \|x\|^2\,\|y\|^2 & \iff |\langle y|x\rangle| \leq \|x\|\,\|y\|\ .
\end{aligned}$$

Gilt in der letzten Zeile das Gleichheitszeichen, so folgt in der ersten Zeile $\|v\| = 0$ und somit $v = \langle x|x\rangle y - \langle y|x\rangle x = 0$ also x und y linear abhängig. Sind umgekehrt x und y linear abhängig, etwa $x = \lambda y$, so bekommt man

$$\begin{aligned}
|\langle x|y\rangle| &= |\langle \lambda y|y\rangle| = |\lambda\langle y|y\rangle| = |\lambda|\,\|y\|^2 \\
\|x\| \cdot \|y\| &= \|\lambda y\| \cdot \|y\| = |\lambda|\,\|y\|^2\ ,
\end{aligned}$$

also die Gleichheit. □

Satz 15.2.4. *Es sei* $(V, \langle\ |\ \rangle)$ *ein Euklidischer Raum und* $\|\ \|$ *die zu* $\langle\ |\ \rangle$ *gehörige Norm. Dann gilt:*

(N_1) $\|x\| \geq 0 \quad \forall x \in V,\ \|x\| = 0 \iff x = 0,$

(N_2) $\|\lambda x\| = |\lambda| \cdot \|x\| \quad \forall x \in V,\ \lambda \in \mathbb{R},$

(N_3) $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$ *(Dreiecksungleichung)*

Beweis: Die ersten beiden Eigenschaften folgen direkt aus der Bilinearität und positiven Definitheit des Skalarprodukts, die dritte aus der CSU:

$$\begin{aligned}\|x+y\|^2 =& \langle x+y|x+y\rangle = \langle x|x\rangle + \langle x|y\rangle + \langle y|x\rangle + \langle y|y\rangle = \\ =& \langle x|x\rangle + \langle x|y\rangle + \langle x|y\rangle + \langle y|y\rangle = \|x\|^2 + 2\langle x|y\rangle + \|y\|^2 \\ (\|x\| + \|y\|)^2 =& \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2\ .\end{aligned}$$

Wegen $\langle x|y\rangle \leq |\langle x|y\rangle| \leq \|x\|\,\|y\|$ (CSU) folgt somit auch

$$\|x+y\|^2 \leq (\|x\| + \|y\|)^2 \iff \|x+y\| \leq \|x\| + \|y\|$$

wegen der Monotonie des Quadrats bzw. der Wurzel für positive Werte. Wegen der Bemerkung am Ende des Beweises der CSU kann die Gleichheit hier nur dann eintreten, wenn x und y linear abhängig sind. □

- (N_3) wird üblicherweise *Dreiecksungleichung* genannt, was durch die nebenstehende Skizze klar sein sollte.
- Die Eigenschaften (N_1)-(N_3) werden allgemein zur Definition einer Norm verwendet.
- Es gibt Normen, die sich nicht durch ein Skalarprodukt $\langle\ |\ \rangle$ als $\|x\| = \sqrt{\langle x|x\rangle}$ definieren lassen:
- $\|x\| := \max\{|x_i|\,|\ 1 \leq i \leq n\}$ für $\dim V = n$ ist z.B. solch eine Norm.
- Diese Normen sind aber in der Linearen Algebra meist nicht von Inter-
13.1.11 esse.

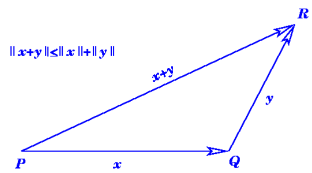


Abbildung 29:
Dreiecksungleichung

15.3 Metrik

Definition 15.3.1. *Es sei* $(V, \langle\ |\ \rangle)$ *ein Euklidischer Vektorraum. Die Abbildung*

$$d : \begin{cases} V \times V & \to & \mathbb{R} \\ (x, y) & \mapsto & \|x - y\| \end{cases}$$

heißt die zu $\langle\ |\ \rangle$ *gehörige Metrik von* V*.*

Folgerung 15.3.2. *Es seien* $(V, \langle\ |\ \rangle)$ *ein Euklidischer Vektorraum und* d *die zu* $\langle\ |\ \rangle$ *gehörige Metrik. Dann gilt:*

(D_1) $d(P, Q) \geq 0,\ d(P, Q) = 0 \iff P = Q,$

(D_2) $d(P, Q) = d(Q, P),$

(D_3) $d(P, R) \leq d(P, Q) + d(Q, R)$

für alle Punkte $P, Q, R \in V$*.*

- Die Eigenschaften (D_1)-(D_3) werden allgemein zur Definition einer Metrik verwendet.
- Es gibt Metriken, die sich nicht durch eine Norm $\|\ \|$ als $d(x, y) = \|x - y\|$ definieren lassen.
- Diese sind aber in der Linearen Algebra meist nicht von Interesse.

15.4 Winkel und Orthogonalität

- Wegen der CSU 15.2.3 ist $-1 \leq \frac{\langle x|y\rangle}{\|x\|\|y\|} \leq 1$.
- Deshalb ist der Arcuscosinus dieses Ausdruckes definiert, was man nun verwendet, um den Winkelbegriff beruhend auf Skalarprodukt und Norm allgemein in Euklidischen Vektorräumen zu definieren:

Definition 15.4.1. *Es sei V ein Euklidischer Vektorraum mit dem Skalarprodukt $\langle\ |\ \rangle$ und der zugehörigen Norm $\|\ \|$. Der Winkel zwischen zwei von dem Nullvektor verschiedenen Vektoren x und y von V ist*

$$\sphericalangle(x,y) := \arccos \frac{\langle x|y\rangle}{\|x\|\,\|y\|}\,.$$

- Wegen der gezeigten Eigenschaften von Skalarprodukt und Norm folgen für den so definierten Winkel die erwarteten Eigenschaften.
- $\sphericalangle(x,y) = \sphericalangle(y,x)\,,\ \sphericalangle(\lambda x,y) = \sphericalangle(x,y)\,,\ \forall x,y \in V, \lambda \in \mathbb{R}, \lambda > 0$
- Weiterhin hat der so definierte Winkel (auch in höherer Dimension) die erwarteten additiven Eigenschaften
- Für den Beweis braucht man den Cosinus mit seinen Additionstheoremen (siehe z.B. Fischer [2000], S.277).

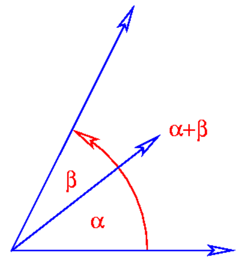


Abbildung 30: Winkel-Additivität

Definition 15.4.2. *Es sei $(V, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum. Zwei Vektoren $x, y \in V$ heißen orthogonal bezüglich $\langle\ |\ \rangle$, wenn $\langle x|y\rangle = 0$ ist.*
Sind $v_1, \ldots, v_m \in V \setminus \{0\}$ so heißt $\{v_1, \ldots, v_m\}$ ein Orthogonalsystem, falls $\langle v_i|v_j\rangle = 0$ für beliebige $i \neq j$ mit $1 \leq i,j \leq m$.
Ist zusätzlich $\langle v_i|v_i\rangle = 1$ für $1 \leq i \leq m$, so spricht man von einem Orthonormalsystem.
Mit Hilfe des Kronecker-Symbols kann man ein Orthonormalsystem kurz durch $\langle v_i|v_j\rangle = \delta_{ij}$ für beliebige $1 \leq i,j \leq m$ beschreiben.

Satz 15.4.3. *Es sei $(V, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum.*
Jedes Orthogonalsystem $\{v_1, \ldots, v_m\}$ in V ist linear unabhängig.

Beweis: Mit dem üblichen Ansatz $\sum_{i=1}^m \mu_i v_i = 0$ folgt für ein beliebiges k mit $1 \leq k \leq m$ wegen der Bilinearität des Skalarprodukts:

$$0 = \langle v_k|\sum_{i=1}^m \mu_i v_i\rangle = \sum_{i=1}^m \mu_i \langle v_k|v_i\rangle = \mu_k \langle v_k|v_k\rangle$$

Wegen $\langle v_k|v_k\rangle > 0$ heißt das $\mu_k = 0$. □

- Ist $m = n = \dim V$, so spricht man deshalb auch von einer *Orthogonalbasis* bzw. einer *Orthonormalbasis.*
- Ausgehend von einer beliebigen Basis $\{v_1, \ldots, v_n\}$ eines endlichdimensionalen Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$ kann nun eine Orthogonal- oder sogar Orthonormalbasis berechnet werden.
- Der erste Schritt dazu ist die orthogonale Zerlegung eines gegebenen Vektors b in eine Summe aus einem Vektor $b_a^{\|}$, der parallel zu einem anderen Vektor a ist (=*orthogonale Projektion*) von b auf a) und einen Vektor $b_a^{\perp}$ orthogonal zu a.

Satz 15.4.4.

Es seien $(V, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum, $O, A, B \in V$ Punkte mit $O \neq A$, $a = \overrightarrow{OA}$, $b = \overrightarrow{OB}$. Dann existiert $L \in OA$ mit $\overrightarrow{LB} \perp a$ und es gilt

$$b_a^{\|} := \overrightarrow{OL} = \frac{\langle a|b\rangle}{\|a\|^2} a$$

$$b_a^{\perp} := \overrightarrow{LB} = b - b_a^{\|}$$

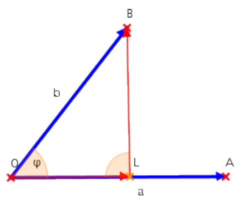


Abbildung 31: Orthogonale Zerlegung

Beweis:

$$\langle a|b_a^\perp\rangle = \langle a|b - \frac{\langle a|b\rangle}{\|a\|^2}a\rangle = \langle a|b\rangle - \frac{\langle a|b\rangle}{\|a\|^2}\underbrace{\langle a|a\rangle}_{=\|a\|^2} = 0 \Rightarrow b_a^\perp \perp a$$

Da offensichtlich $b_a^\parallel \| a$ und $b_a^\parallel + b_a^\perp = b$ gilt, folgt die Behauptung. □

15.5 Orthogonalisierung

- Geht man von einer Basis $\{v_1, \ldots, v_n\}$ aus, so nimmt man z.B. $w_1 := v_1$ als ersten Vektor der zu bestimmenden Orthogonalbasis und $w_2 := v_2 - \frac{\langle w_1|v_2\rangle}{\|w_1\|^2}w_1$.
- Nach 15.4.4 gilt dann $w_1 \perp w_2$ und nach 15.4.3 ist $\{w_1, w_2\}$ linear unabhängig.
- Wegen $w_1, w_2 \in \mathrm{Lin}\,(v_1, v_2)$ folgt $\mathrm{Lin}\,(w_1, w_2) = \mathrm{Lin}\,(v_1, v_2)$, d.h. $\{w_1, w_2\}$ ist eine Orthogonalbasis von $\mathrm{Lin}\,(v_1, v_2)$.
- Die gleiche Grundidee wird nun weiter verfolgt, um eine Orthogonalbasis von $\mathrm{Lin}\,(v_1, v_2, v_3)$ zu bestimmen.
- Dazu wird die orthogonale Projektion von v_3 in die von w_1 und w_2 aufgespannte Ebene berechnet und dann von v_3 subtrahiert:

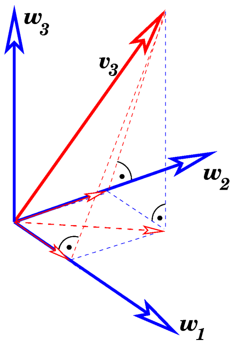


Abbildung 32: Orthogonale Projektion

- Die Projektion von v_3 in $\mathrm{Lin}\,(w_1, w_2)$ ist die Summe der Projektionen von v_3 auf w_1 und von v_3 auf w_2:
- Der Vektor

  $$(v_3)^\parallel_{\mathrm{Lin}(w_1,w_2)} := \frac{\langle w_1|v_3\rangle}{\|w_1\|^2}w_1 + \frac{\langle w_2|v_3\rangle}{\|w_2\|^2}w_2$$

  liegt offensichtlich in $\mathrm{Lin}\,(w_1, w_2)$.
- Der Vektor

  $$(v_3)^\perp_{\mathrm{Lin}(w_1,w_2)} := w_3 = v_3 - (v_3)^\parallel_{\mathrm{Lin}(w_1,w_2)}$$

  ist orthogonal zu jedem Vektor aus $\mathrm{Lin}\,(w_1, w_2)$,
- Dies zeigt man durch Skalarprodukt mit den Basisvektoren w_1 und w_2:

$$\begin{aligned}\langle w_1|w_3\rangle &= \langle w_1|v_3\rangle - \frac{\langle w_1|v_3\rangle}{\|w_1\|^2}\underbrace{\langle w_1|w_1\rangle}_{\|w_1\|^2} - \frac{\langle w_2|v_3\rangle}{\|w_2\|^2}\underbrace{\langle w_1|w_2\rangle}_{=0} = \\ &= \langle w_1|v_3\rangle - \langle w_1|v_3\rangle = 0\end{aligned}$$

- und analog $\langle w_2|w_3\rangle = 0$.
- Diese Grundidee lässt sich rekursiv fortsetzen und endet bei einer orthogonalen Basis $\{w_1, \ldots, w_n\}$ von $\{v_1, \ldots, v_n\}$.
- Das gesamte Verfahren ist nach *Jorgen Pedersen Gram (1850 - 1916)* und *Erhard Schmidt (1876 - 1959)* benannt:

Satz 15.5.1 (Gram-Schmidtsches Orthogonalisierungsverfahren)**.** *Es sei $(V, \langle\ |\ \rangle)$ ein endlichdimensionaler Euklidischer Vektorraum mit einer Basis $\{v_1, \ldots, v_n\}$. Dann ist $\{w_1, \ldots, w_n\}$ mit*

$$w_i := v_i - \sum_{j=1}^{i-1}\frac{\langle w_j|v_i\rangle}{\|w_j\|^2}w_j \quad \textit{für} \quad i = 1, \ldots, n$$

eine Orthogonalbasis von V bezüglich $\langle\ |\ \rangle$, bzw. $(\frac{w_1}{\|w_1\|}, \ldots, \frac{w_n}{\|w_n\|})$ eine Orthonormalbasis von V.

Beweis: Der Beweis wird durch Induktion nach i geführt. Der Fall $i = 1$ ist klar, der Beweis für die Fälle $i = 2$ und $i = 3$ ist in der Herleitung gemacht worden. Der noch fehlende Induktionsschritt von $i-1$ nach i ist nun leicht zu führen, indem man in obiger Gleichung für w_i beidseitig das Skalarprodukt mit w_ℓ für $1 \leq \ell \leq i-1$ bildet. Da nach Induktionsvoraussetzung $w_j \perp w_\ell$ für $1 \leq j, \ell \leq i-1$ und $j \neq \ell$ gilt, führt das auf

$$\langle w_\ell | w_i \rangle = \langle w_\ell | v_i \rangle - \sum_{j=1}^{i-1} \frac{\langle w_j | v_i \rangle}{\|w_j\|^2} \langle w_\ell | w_j \rangle = \langle w_\ell | v_i \rangle - \frac{\langle w_\ell | v_i \rangle}{\|w_\ell\|^2} \langle w_\ell | w_\ell \rangle = 0$$

□

Folgerung 15.5.2. *Jeder endlichdimensionale Euklidische Vektorraum $(V, \langle\ |\ \rangle)$ besitzt eine Orthogonalbasis (und natürlich auch eine Orthonormalbasis).*

Beweis: Man wende das Gram-Schmidtsche Orthogonalisierungsverfahren auf eine beliebige Basis von V an (und normiere gegebenenfalls). □

Folgerung 15.5.3. *Jede Orthogonalbasis eines Untervektorraumes U eines endlichdimensionalen Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$ lässt sich zu einer Orthogonalbasis von V ergänzen (analog für Orthonormalbasen).*

Beweis: Man ergänze eine Orthogonalbasis von U zu einer Basis von V und wende das Orthogonalisierungsverfahren auf diese Basis an. Die bereits orthogonalen Basisvektoren von U werden dabei nicht verändert (analog für Orthonormalbasen). □

Satz 15.5.4. *Es sei $B = (b_1, \ldots, b_n)$ eine Orthonormalbasis eines endlichdimensionalen Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$. Dann gilt für beliebiges $v \in V$:*

$$v_{/B} = \begin{pmatrix} \langle b_1 | v \rangle \\ \langle b_2 | v \rangle \\ \vdots \\ \langle b_n | v \rangle \end{pmatrix}.$$

Beweis: Ist $v = \sum_{i=1}^{n} v_i b_i$ die eindeutige Basisdarstellung von v bezüglich B, so zeigt beidseitiges Skalarprodukt mit einem b_j mit $1 \leq j \leq n$

$$\langle b_j | v \rangle = \langle b_j | \sum_{i=1}^{n} v_i b_i \rangle = \sum_{i=1}^{n} v_i \langle b_j | b_i \rangle = v_j \langle b_j | b_j \rangle = v_j \,.$$ □

15.6 Orthogonale Unterräume

Definition 15.6.1. *Zwei Untervektorräume U_1, U_2 eines Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$ heißen orthogonal, in Zeichen $U_1 \perp U_2$, wenn $u_1 \perp u_2$ für alle $u_1 \in U_1$ und $u_2 \in U_2$ gilt.*

Definition 15.6.2. *Ist U ein Untervektorraum eines Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$, so sei das orthogonale Komplement von U*

$$U^\perp := \{v \in V \mid \langle u | v \rangle = 0 \quad \textit{für alle} \quad u \in U\}$$

Satz 15.6.3. *Es sei U ein Untervektorraum eines endlichdimensionalen Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$. Dann gilt ($\oplus$ =direkte Summe →Übung 8.9)*

$$V = U \oplus U^\perp \,, \; U^{\perp\perp} = U \,, \; \dim V = \dim U + \dim U^\perp$$

Beweis: Da V endlichdimensional vorausgesetzt ist, sind dies auch U und $U^\perp$. Man kann sich also nach 15.5.3 eine endliche Orthonormalbasis $\{b_1, \ldots, b_m\}$ von U nehmen und diese zu einer Orthonormalbasis $B := \{b_1, \ldots, b_m, b_{m+1}, \ldots, b_n\}$ von V ergänzen. Für jedes $v \in U^\perp$ gilt dann gemäß 15.5.4

$$v_{/B} = \begin{pmatrix} \langle b_1 | v \rangle \\ \vdots \\ \langle b_m | v \rangle \\ \langle b_{m+1} | v \rangle \\ \vdots \\ \langle b_n | v \rangle \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \langle b_{m+1} | v \rangle \\ \vdots \\ \langle b_n | v \rangle \end{pmatrix}$$

d.h. v ist eine Linearkombination von $b_{m+1}, \dots, b_n$, also

$$U^\perp = \mathrm{Lin}\,(b_{m+1}, \dots, b_n)\ .$$

Daraus folgt direkt der erste und dritte Teil der Behauptung. Teil 2 folgt aus der Definition von $U^\perp$. □

15.7 Abstände von Teilräumen

Definition 15.7.1. *Es sei $(V, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum, M eine beliebige Teilmenge von V und $p \in V$ ein Punkt. Der Abstand $d(p, M)$ von p und M ist definiert durch*

$$d(p, M) := \inf\{d(p, m) \mid m \in M\}\,.$$

- Von besonderem Interesse sind dabei die Abstände, bei denen M ein affiner Teilraum ist:

Satz 15.7.2 (Abstand Punkt-Gerade)**.** *Es seien $(V, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum, $u, v, p \in V$, $v \neq 0$ und $G = \{u + \lambda v \mid \lambda \in \mathbb{R}\}$ eine Gerade. Dann gilt:*

$$d(p, G) = \sqrt{\|u - p\|^2 - \frac{\langle u - p|v\rangle^2}{\|v\|^2}}\,.$$

Dieser Wert wird erreicht für

$$\lambda = \frac{\langle p - u|v\rangle}{\|v\|^2},\ \text{ d.h. im Geradenpunkt } \ell_{p,G} = u + \frac{\langle p - u|v\rangle}{\|v\|^2} v\,.$$

Das Infimum aus 15.7.1 ist also ein Minimum. Die Verbindungsgerade von p und $\ell_{p,G}$ steht orthogonal auf G, d.h. der kürzeste Abstand ist das Lot, $\ell_{p,G}$ der Lotfußpunkt.

Beweis: in der Übung □

- Da man jetzt weiß, dass der kürzeste Abstand das Lot ist, kann man diesen Abstand z.B. auch so berechnen (vgl. Skizze):

$$d(p, G) = \left\|(p - u)_v^\perp\right\|\ ,$$

was nach Einsetzen der Formel aus 15.4.4 auf den gleichen Ausdruck führt.

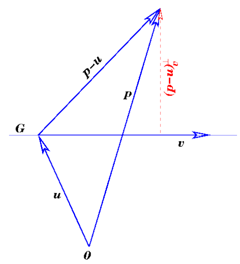


Abbildung 33: Abstand Punkt-Gerade

- Ist $U = \mathrm{Lin}\,(u_1, \dots, u_k)$ mit $1 < k < n - 1$, so kann man wegen $V = U \oplus U^\perp$ (→15.6.3) jedes $p \in V$ eindeutig zerlegen in

$$p = \underbrace{p_U^\parallel}_{\in U} + \underbrace{p_U^\perp}_{\in U^\perp}\,.$$

- Bereits bei der Herleitung des Gram-Schmidtschen Orthogonalisierungsverfahrens war gezeigt worden, dass die orthogonale Projektion $p_U^\parallel$ eines Punktes p in einen Unterraum U gerade die Summe der orthogonalen Projektionen von p auf die Basisvektoren von U ist, wenn diese ein Orthogonalsystem bilden, also

$$p_U^\parallel = \sum_{j=1}^{k} \frac{\langle u_j|p\rangle}{\|u_j\|^2} u_j\,.$$

- Dies kann man nun zur Berechnung des Abstandes eines Punktes zu U verwenden:

Satz 15.7.3. *Es sei U ein endlichdimensionaler Unterraum eines Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$ mit der Orthogonalbasis $\{u_1, \ldots, u_k\}$. Dann gilt*

$$d(p, U) = \sqrt{\|p\|^2 - \sum_{j=1}^{k} \frac{\langle u_j|p\rangle^2}{\|u_j\|^2}}\,.$$

Beweis: Für beliebiges $u \in U$ gilt

$$\begin{aligned}\|p-u\|^2 = \left\|p_U^{\|} + p_U^{\perp} - u\right\|^2 &= \langle p_U^{\|} + p_U^{\perp} - u | p_U^{\|} + p_U^{\perp} - u\rangle = \\ &= \langle p_U^{\|} - u | p_U^{\|} - u\rangle + \langle p_U^{\perp} | p_U^{\perp}\rangle = \\ &= \left\|p_U^{\|} - u\right\|^2 + \|p_U^{\perp}\|^2 \geq \|p_U^{\perp}\|^2\end{aligned}$$

Das Gleichheitszeichen wird dabei für $p_U^{\|} = u$ angenommen, $\|p_U^{\perp}\|$ ist also das Minimum von $\|p-u\|$. Zusammen mit $\|p_U^{\perp}\|^2 + \left\|p_U^{\|}\right\|^2 = \|p\|^2 \iff \|p_U^{\perp}\|^2 = \|p\|^2 - \left\|p_U^{\|}\right\|^2$ (Pythagoras) folgt also

$$\begin{aligned}d(p,U) = \|p_U^{\perp}\| = \sqrt{\|p\|^2 - \left\|p_U^{\|}\right\|^2} &= \sqrt{\|p\|^2 - \left\|\sum_{j=1}^{k} \frac{\langle u_j|p\rangle}{\|u_j\|^2} u_j\right\|^2} = \\ &= \sqrt{\|p\|^2 - \sum_{j=1}^{k} \frac{\langle u_j|p\rangle^2}{\|u_j\|^2}}\,.\end{aligned}$$

□

15.8 Winkel zwischen Teilräumen

- Schneiden sich affine Teilräume eines Euklidischen Vektorraumes, so kann man in manchen Fällen auch von einem *Schnittwinkel* sprechen.
- Eine sinnvolle Definition muss dabei jeweils unabhängig von den speziellen Basisvektoren der Richtungsräume sein:

Definition 15.8.1. *Der Winkel zwischen sich schneidenden Geraden $g_1 = u_1 + \mathbb{R}w_1$ und $g_2 = u_2 + \mathbb{R}w_2$ in einem Euklidischen Vektorraum $(V, \langle\ |\ \rangle)$ sei*

$$\sphericalangle(g_1, g_2) := \begin{cases} \sphericalangle(w_1, w_2) & \text{falls } \langle w_1|w_2\rangle \geq 0 \\ \sphericalangle(-w_1, w_2) & \text{sonst}\,. \end{cases}$$

- Zusammen mit der Definition des Winkels zwischen Vektoren folgt

Folgerung 15.8.2 (Winkel zwischen Geraden)**.**

$$0 \leq \sphericalangle(u_1 + \mathbb{R}w_1, u_2 + \mathbb{R}w_2) = \arccos \frac{|\langle w_1|w_2\rangle|}{\|w_1\|\,\|w_2\|} \leq \frac{\pi}{2}\,.$$

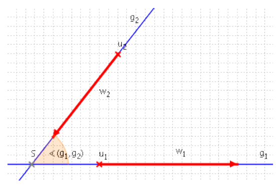


Abbildung 34: Winkel zwischen Geraden 17.1.11

Definition 15.8.3.

Es seien $(V, \langle\ |\ \rangle)$ ein endlichdimensionaler Euklidischer Vektorraum, $H : \langle N|x-p\rangle = 0$ eine Hyperebene (also $p, N \in V$, $N \neq 0$) und $g = u + \mathbb{R}w$ eine Gerade (also $u, w \in V$, $w \neq 0$). Der Winkel zwischen der Hyperebene H und der Geraden g wird definiert als $\frac{\pi}{2}-$der Winkel zwischen den Geraden $\mathbb{R}w$ und $\mathbb{R}N$:

$$\sphericalangle(g, H) := \frac{\pi}{2} - \sphericalangle(\mathbb{R}w, \mathbb{R}N) = \frac{\pi}{2} - \arccos \frac{|\langle N|w\rangle|}{\|N\|\,\|w\|}\,.$$

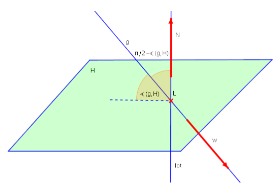


Abbildung 35: Winkel zwischen Gerade und Hyperebene

Definition 15.8.4.

Es seien $(V, \langle\ |\ \rangle)$ ein endlichdimensionaler Euklidischer Vektorraum und $H : \langle N|x-p\rangle = 0$, $H' : \langle N'|x-p'\rangle = 0$ zwei Hyperebenen (also $p, p', N, N' \in V$, $N, N' \neq 0$). Der Winkel zwischen den Hyperebene H und H' wird definiert als der Winkel zwischen den Geraden $\mathbb{R}N$ und $\mathbb{R}N'$:

$$\sphericalangle(H, H') := \sphericalangle(\mathbb{R}N, \mathbb{R}N') = \arccos \frac{|\langle N|N'\rangle|}{\|N\|\,\|N'\|}\,.$$

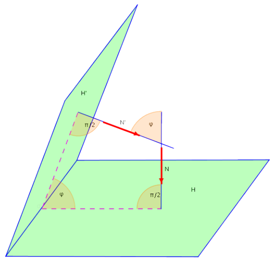


Abbildung 36: Winkel zwischen Gerade und Hyperebene

15.9 Volumen

Definition 15.9.1. *Für $m \in \mathbb{N}$ und Vektoren $a_1, \ldots, a_m$ eines Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$ heißt*

$$P(a_1, \ldots, a_m) := \{\lambda_1 a_1 + \cdots + \lambda_m a_m \mid 0 \leq \lambda_i \leq 1 \ \text{ für } \ i = 1, \ldots, m\}$$

(m-dimensionales, falls $a_1, \ldots, a_m$ linear unabhängig sind) Parallelotop oder Parallelepiped, im $\mathbb{R}^3$ auch Spat.

- Um einen sinnvollen Volumenbegriff für Parallelotope einzuführen, verwendet man Grundideen der mathematischen Maßtheorie, wie z.B.:
  - Sind $M_1, M_2, \ldots, M_n$ endlich viele paarweise disjunkte Parallelotope, so ist das Volumen von deren Vereinigungsmenge die Summe der einzelnen Volumina („Additivität“).
  - Die Verschiebung eines beliebigen Parallelotops ändert nicht dessen Volumen („Bewegungsinvarianz“).
  - Ein gleichseitiges orthogonales Parallelotop mit der Kantenlänge 1 (ON-System $a_1, \ldots, a_m$) hat das Volumen 1 („Normiertheit“).
- Für Strecken (=1-dimensionale Parallelotope) ist solch ein Volumenbegriff bereits bekannt, nämlich die Länge („Metrik“).
- Um diesen Volumenbegriff in höhere Dimension zu übertragen, verwendet man das nach *Bonaventura Francesco Cavalieri (1598 - 1647)* benannte Cavalieri-Prinzip.
- Dieses besagt:

  Raumgebilde sind inhaltsgleich, wenn jeweils in gleicher Höhe geführte Schnitte inhaltsgleiche Gebilde ergeben.
- Um das Cavalieri-Prinzip zu verwenden, zerlegt man m-dimensionale Parallelotope in „Schnitte“ $m-1$-dimensionaler Parallelotope:

$$P(a_1, \ldots, a_m) = \bigcup_{0 \leq \lambda_m \leq 1} P(a_1, \ldots, a_{m-1}) + \lambda_m a_m\,.$$

Definition 15.9.2. *Für $m \in \mathbb{N}$ und Vektoren $a_1, \ldots, a_m$ eines Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$ bezeichne* $\mathrm{Vol}_{\langle\ |\ \rangle}(a_1, \ldots, a_m)$ *(oder einfach* $\mathrm{Vol}\,(a_1, \ldots, a_m)$*, wenn das Skalarprodukt klar ist) das Volumen des Parallelotops $P(a_1, \ldots, a_m)$ bezüglich des Skalarproduktes $\langle\ |\ \rangle$. Dieses Volumen wird rekursiv definiert durch* $\mathrm{Vol}\,(a_1) = \|a_1\| = \sqrt{\langle a_1|a_1\rangle}$ *und für $m \geq 2$:*

$$\mathrm{Vol}\,(a_1, \ldots, a_m) := \mathrm{Vol}\,(a_1, \ldots, a_{m-1}) \cdot d(a_m, \mathrm{Lin}\,(a_1, \ldots, a_{m-1}))\,.$$

Satz 15.9.3. *Es seien $(V, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum, $m \in \mathbb{N}$ und $a_1, \ldots, a_m \in V$. Dann ist*

$$\mathrm{Vol}_{\langle\ |\ \rangle}(a_1, \ldots, a_m) = \sqrt{\det\left((\langle a_i|a_j\rangle)_{1 \leq i,j \leq m}\right)}\,.$$

- Den Satz beweist man mit Induktion aus 15.9.2 und mit Hilfe von 15.7.3 ([siehe z.B. Koecher, 1997]).

Folgerung 15.9.4. *Bezüglich des Standardskalarproduktes $\langle \mid \rangle$ im $\mathbb{R}^n$ gilt*

$$(\langle a_i | a_j \rangle)_{1 \leq i,j \leq m} = {}^t A \cdot A \,,$$

wobei A die Matrix mit den Spalten $a_1, a_2, \ldots, a_m$ sei.
Damit folgt für $m = n$:

$$\mathrm{Vol}_{\langle \mid \rangle}(a_1, a_2, \ldots, a_m) = \sqrt{\det({}^t A \cdot A)} = \sqrt{\det(A)^2} = |\det(A)| \,,$$

also der bereits bei der Einführung der Determinante erwähnte Spezialfall.

- Neben den Parallelotopen sind besonders das von zwei Vektoren aufgespannte *Dreieck* und das von drei Vektoren aufgespannte *Tetraeder* (=Pyramide aus Dreiecksflächen) für die Flächen- bzw. Volumenmessung interessant.
- Das Dreieck hat die halbe Fläche des von den gleichen Vektoren aufgespannten Parallelogramms, das Tetraeder ein Sechstel des Volumens des von den gleichen Vektoren aufgespannten Spats.

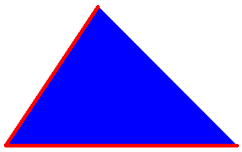


Abbildung 37: $F_\Delta = \frac{1}{2}\,\mathrm{Vol}\,(a_1, a_2)$

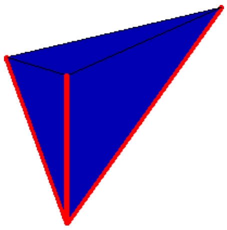


Abbildung 38: $V_{\mathrm{Tetra}} = \frac{1}{6}\,\mathrm{Vol}\,(a_1, a_2, a_3)$

16 Unitäre Vektorräume

16.1 Sesquilinearformen

- Das kanonische reelle Skalarprodukt lässt sich nicht direkt auf $\mathbb{C}$ übertragen: Mit $v = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ und $w = \left(\begin{smallmatrix}0\\i\end{smallmatrix}\right)$ gilt etwa
$$^t ww = \begin{pmatrix}0 & i\end{pmatrix}\begin{pmatrix}0\\i\end{pmatrix} = -1 \quad ^t vw = \begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}0\\i\end{pmatrix} = i$$
d.h. die Eigenschaften des Skalarproduktes, die darauf aufbauende Begriffe wie Norm und Winkel erst ermöglichen, sind nicht erfüllt:
- Es ist nicht $^t ww \geq 0$,
- und somit kann man weder die Norm mit diesem Produkt definieren, noch die CSU 15.2.3 formulieren.
- Die Eigenschaft $\langle w|w\rangle \geq 0$ kann man auf den komplexen Fall hinüberretten, wenn man statt mit $^t vw$ mit $\langle v|w\rangle := {}^t(\overline{v})\, w$ rechnet.
- Dieses so genannte *kanonische komplexe Skalarprodukt* ist allerdings keine Bilinearform mehr, denn es gilt ja
$$\langle \lambda v|w\rangle := {}^t\!\left(\overline{\lambda v}\right) w = \overline{\lambda}\, {}^t(\overline{v})\, w = \overline{\lambda}\langle v|w\rangle \text{ und } \overline{\lambda} \neq \lambda \text{ für } \lambda \notin \mathbb{R}\,.$$
- Dies führt auf die folgende Definition:

Definition 16.1.1. *Es sei V ein $\mathbb{C}$-Vektorraum. Eine Abbildung $\varphi : V \times V \to \mathbb{C}$ heißt Sesquilinearform*[16]*, wenn gilt*

(SF_1) *φ ist semilinear in der ersten Komponente, d.h.*
$$\varphi(u+v,w) = \varphi(u,w) + \varphi(v,w) \quad \forall u,v,w \in V$$
$$\varphi(c\cdot u,w) = \overline{c}\cdot\varphi(u,w) \quad \forall u,w \in V\,,\ c \in \mathbb{C}$$

(SF_2) *φ ist linear in der zweiten Komponente, d.h.*
$$\varphi(u,v+w) = \varphi(u,v) + \varphi(u,w) \quad \forall u,v,w \in V$$
$$\varphi(u,c\cdot w) = c\cdot\varphi(u,w) \quad \forall u,w \in V\,,\ c \in \mathbb{C}$$

- In der Literatur ist nicht einheitlich, in welcher Komponente der Sesquilinearform die volle Linearität, und in welcher die „halbe“ Linearität (man findet auch die Bezeichnung *konjugiert linear* oder *semilinear* dafür) vorliegt. Das macht auch keinen wesentlichen Unterschied, man sollte nur darauf gefasst sein.

16.2 Matrixdarstellung und Basiswechsel

- Wie eine Bilinearform kann man auch eine Sesquilinearform durch ihre *Grammatrix* beschreiben. Die Herleitung kann man vom reellen Fall fast wörtlich übernehmen, muss nur jeweils wegen $\varphi(cu,v) = \overline{c}\varphi(u,v)$ den Strich in die erste Komponente bringen:
- Es sei φ eine Sesquilinearform auf einem $\mathbb{C}$-Vektorraum V endlicher Dimension und $B = (b_1,\ldots,b_n)$ eine geordnete Basis von V. Mit $a_{ij} := \varphi(b_i,b_j)$ und $A = (a_{ij})$ gilt:
$$\varphi(v,w) = {}^t\!\left(\overline{v}_{/B}\right)\cdot A\cdot w_{/B}\,,$$
oder kurz $\varphi(v,w) = {}^t(\overline{v})\cdot A\cdot w$, wenn B feststeht (Bezeichnung auch wieder $G_B(\varphi) = A$).
- Das kanonische komplexe Skalarprodukt bekommt man daraus mit der Einheitsmatrix als Grammatrix.
- Auch beim Basiswechsel lässt sich alles mit einem Strich zusätzlich auf den komplexen Fall übertragen:
$$G_C(\varphi) = {}^t\!\left(\overline{{}_B[\mathrm{id}_V]_C}\right)\cdot G_B(\varphi)\cdot {}_B[\mathrm{id}_V]_C = {}^t(\overline{T})\cdot G_B(\varphi)\cdot T$$
mit der Transformationsmatrix $T = {}_B[\mathrm{id}_V]_C$.

[16] sesquilinear heißt $1\frac{1}{2}$-mal linear

16.3 Hermitesche Formen

- Analog zur symmetrischen Bilinearform definiert man

Definition 16.3.1. *Es seien V ein $\mathbb{C}$-Vektorraum und $\varphi : V \times V \to \mathbb{C}$ eine Sesquilinearform. Die Sesquilinearform φ heißt hermitesch (kurz: hermitesche Form), wenn gilt*

$$\varphi(u,v) = \overline{\varphi(v,u)} \quad \forall u,v \in V\,.$$

Folgerung 16.3.2. *Für eine Hermitesche Form gilt insbesondere $\varphi(v,v) = \overline{\varphi(v,v)} \quad \forall v \in V$, d.h $\varphi(v,v) \in \mathbb{R}$!*

- Etwas mühsamer zu sehen, aber auch richtig, ist die Umkehrung dieser Aussage [siehe z.B. Lorenz, 2005, , S.105]:

Satz 16.3.3. *Gilt für eine Sesquilinearform φ, dass $\varphi(u,u) \in \mathbb{R}$ für alle $u \in V$, so ist φ Hermitesch.*

- Bei einer Hermiteschen Form gilt

$$\begin{aligned}
&\forall v,w \in V:\\
&\left.\begin{array}{rcl}
\varphi(v,w) &=& {}^t(\overline{v})\, G_B(\varphi)\, w = {}^t\!\left({}^t(\overline{v})\, G_B(\varphi)\, w\right) = {}^t w\, {}^t\!\left(G_B(\varphi)\right) \overline{v}\\
\overline{\varphi(w,v)} &=& {}^t(\overline{\overline{w}}) \cdot \overline{G_B(\varphi)} \cdot \overline{v} = {}^t w \cdot \overline{G_B(\varphi)} \cdot \overline{v}
\end{array}\right\} \Rightarrow\\
&G_B(\varphi) = {}^t\!\left(\overline{G_B(\varphi)}\right).
\end{aligned}$$

- Solch eine Matrix heißt *hermitesch*.
- Eine Hermitesche Form φ liegt also genau dann vor, wenn ihre Grammatrix $G_B(\varphi)$ Hermitesch ist.
- Für ${}^t(\overline{A})$ findet man auch die Kurz-Schreibweisen A^* oder A^H. A^* heißt auch die *zu A adjungierte Matrix*.
- Eine Matrix ist also Hermitesch, wenn sie mit ihrer *Adjungierten* übereinstimmt.
- Für die Adjungierte beweist man leicht die folgenden Analogien zur Transponierten:
- ${}^t(AB) = {}^tB\,{}^tA \rightsquigarrow (AB)^* = B^*A^*$,
- $\left({}^tA\right)^{-1} = {}^t\!\left(A^{-1}\right) \rightsquigarrow (A^*)^{-1} = \left(A^{-1}\right)^*$.
- Wegen 16.3.2 kann man analog zu 14.3.3 auch im komplexen Fall zu einer Hermiteschen Form φ des $\mathbb{C}$-Vektorraumes V die zugehörige *quadratische Form*

$$q_\varphi : \begin{cases} V \to \mathbb{R}\\ v \mapsto \varphi(v,v) \end{cases}$$

betrachten.

- Es gilt $q(\mu v) = \mu\overline{\mu}q(v) = |\mu|^2\, q(v)$ für alle $\mu \in \mathbb{C}$ und $v \in V$.
- Auch der Trägheitssatz von Sylvester 14.3.7 und die Definition der Definitheit 14.4.1 lassen sich damit auf den komplexen Fall übertragen:

Satz 16.3.4. *Es sei $K = \mathbb{C}$ und φ eine Hermitesche Form des n-dimensionalen $\mathbb{C}$-Vektorraumes V. Dann existiert eine Basis B, so dass*

$$G_B(\varphi) = \begin{pmatrix} \mathbb{1}_{n_+} & 0 & 0\\ 0 & -\mathbb{1}_{n_-} & 0\\ 0 & 0 & 0_{n_0} \end{pmatrix}.$$

Dabei seien $\mathbb{1}_{n_+}$ bzw. $\mathbb{1}_{n_-}$ Einheitsmatrizen der Größe n_+ bzw. n_- und 0_{n_0} eine $n_0 \times n_0$-Nullmatrix. Die restlichen Nullen bezeichnen Nullmatrizen passender Größe. Die Zahlen $n_+, n_-, n_0 \in \mathbb{N}$ ($n_+ + n_- + n_0 = n$) sind durch die Hermitesche Form eindeutig bestimmt.

Definition 16.3.5. *Es seien $K = \mathbb{C}$ und φ eine Hermitesche Form des n-dimensionalen $\mathbb{C}$-Vektorraumes V. Das nach 16.3.4 eindeutig bestimmte Tripel (n_+, n_-, n_0) heißt die Signatur der Hermiteschen Form φ. Eine Hermitesche Form mit $n_+ = n = \dim V$ heißt positiv definit. Die Matrix $G_B(\varphi)$ aus dem Trägheitssatz heißt die Sylvester-Normalform von φ.*
Wie im reellen Fall lässt sich das verallgemeinern: eine Hermitesche Form φ in einem komplexen Vektorraum V (nicht zwingend endlichdimensional) heißt positiv definit, wenn gilt $\varphi(v,v) > 0\ \forall v \in V \setminus \{0\}$, positiv semidefinit, wenn gilt $\varphi(v,v) \geq 0\ \forall v \in V$, negativ definit, wenn $\varphi(v,v) < 0\ \forall v \in V \setminus \{0\}$ und negativ semidefinit, wenn $\varphi(v,v) \leq 0\ \forall v \in V$.

- Vorsicht: Bei einer Hermiteschen Matrix über $\mathbb{C}$ bedeuten Kongruenzumformungen, dass man elementare Zeilenumformungen macht und die jeweils dazugehörigen *konjugierten* Spaltenumformungen:

Beispiel 16.3.6.

$$\left(\begin{array}{ccc|ccc} 1 & -i & -1-i & 1 & 0 & 0 \\ i & 3 & 2i & 0 & 1 & 0 \\ -1+i & -2i & 5 & 0 & 0 & 1 \end{array}\right) \begin{array}{c} Z_2 \leftarrow Z_2 - iZ_1 \\ Z_3 \leftarrow Z_3 + (1-i)Z_1 \\ \rightsquigarrow \text{ und } \rightsquigarrow \\ S_2 \leftarrow S_2 + iS_1 \\ S_3 \leftarrow S_3 + (1+i)S_1 \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -1+3i & -i & 1 & 0 \\ 0 & -1-3i & 3 & 1-i & 0 & 1 \end{array}\right) \begin{array}{c} Z_3 \leftarrow Z_3 + (\frac{1}{2} + \frac{3}{2}i)Z_2 \\ \rightsquigarrow \text{ und } \rightsquigarrow \\ S_3 \leftarrow S_3 + (\frac{1}{2} - \frac{3}{2}i)S_2 \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & -i & 1 & 0 \\ 0 & 0 & -2 & \frac{5}{2} - \frac{3}{2}i & \frac{1}{2} + \frac{3}{2}i & 1 \end{array}\right) \begin{array}{c} Z_2 \leftarrow \frac{Z_2}{\sqrt{2}} \\ \rightsquigarrow \\ S_2 \leftarrow \frac{S_2}{\sqrt{2}} \end{array} \quad \text{und} \quad \begin{array}{c} Z_3 \leftarrow \frac{Z_3}{\sqrt{2}} \\ \rightsquigarrow \\ S_3 \leftarrow \frac{S_3}{\sqrt{2}} \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{i\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & -1 & \frac{5\sqrt{2}}{4} - \frac{3\sqrt{2}}{4}i & \frac{\sqrt{2}}{4} + \frac{3\sqrt{2}}{4}i & \frac{\sqrt{2}}{2} \end{array}\right) \Rightarrow$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = S \cdot \begin{pmatrix} 1 & -i & -1-i \\ i & 3 & 2i \\ -1+i & -2i & 5 \end{pmatrix} \cdot {}^t\overline{S} \quad \text{mit}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{i\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{5\sqrt{2}}{4} - \frac{3\sqrt{2}}{4}i & \frac{\sqrt{2}}{4} + \frac{3\sqrt{2}}{4}i & \frac{\sqrt{2}}{2} \end{pmatrix} = S = {}^t\overline{T} \Rightarrow$$

$$T = {}_B[\mathrm{id}_V]_C = {}^t\overline{S} = \begin{array}{c} \begin{pmatrix} 1 & \frac{i\sqrt{2}}{2} & \frac{5\sqrt{2}}{4} + \frac{3\sqrt{2}}{4}i \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{4} - \frac{3\sqrt{2}}{4}i \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{pmatrix} \\ \begin{array}{ccc} c_1/_B & c_2/_B & c_3/_B \end{array} \end{array}$$

20.1.11

Satz 16.3.7. *Es sei V ein n-dimensionaler $\mathbb{C}$-Vektorraum mit Basis B und φ eine Hermitesche Form. Dann besitzt $A := G_B(\varphi)$ genau n reelle (nicht notwendig verschiedene, entsprechend ihrer Vielfachheit gezählte) Eigenwerte, d.h. jede Hermitesche $n \times n$-Matrix hat n reelle Eigenwerte!*

Beweis: Analog zu 14.4.4→selber machen. □

16.4 Norm, Metrik, CSU

Definition 16.4.1. *Es sei V ein $\mathbb{C}$-Vektorraum und $\varphi : V \times V \to \mathbb{R}$ eine positiv definite Hermitesche Form. Dann heißt φ Skalarprodukt auf V und V unitärer Vektorraum.*

Definition 16.4.2. *Ist $(V, \langle\ |\ \rangle)$ ein unitärer Vektorraum, so definiert man analog zu 15.2.1 die Norm*

$$\|\ \| : \begin{cases} V & \to \quad \mathbb{R} \\ x & \mapsto \quad \sqrt{\langle x|x\rangle} \end{cases}$$

und analog zu 15.3.1 die zu $\langle\ |\ \rangle$ gehörige Metrik von V

$$d : \begin{cases} V \times V & \to \quad \mathbb{R} \\ (x,y) & \mapsto \quad \|x-y\| \end{cases}$$

- Die so definierte Norm erfüllt sinngemäß 15.2.4 (unitär statt Euklidisch und $\lambda \in \mathbb{C}$), die Metrik 15.3.2.
- Auch die Cauchy-Schwarz-sche Ungleichung 15.2.3 gilt im komplexen Fall (im Beweis sind einige Striche für das Konjugierte hinzuzufügen):

Satz 16.4.3. *Es sei $(V, \langle \ | \ \rangle)$ ein unitärer Vektorraum. Für $x, y \in V$ gilt*

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

Das Gleichheitszeichen gilt genau dann, wenn x und y linear abhängig sind.

- Da in einem unitären Vektorraum üblicherweise $\frac{\langle x|y\rangle}{\|x\|\|y\|}$ nicht reell ist, kann man dort den Winkel nicht analog zu 15.4.1 definieren.
- Man kann aber auch im komplexen Fall von Orthogonalität reden: Zwei Vektoren $x, y \in V$ heißen auch hier *orthogonal* bezüglich $\langle \ | \ \rangle$, wenn $\langle x|y\rangle = 0$ ist.
- Anwendungen der Orthogonalität, wie etwa 15.5.4, gelten auch im komplexen Fall.
- Auch orthogonale Teilräume kann man entsprechend 15.6.1 einführen und z.B. mit dem Gram-Schmidtschen Orthogonalisierungsverfahren 15.5.1 Orthogonalbasen berechnen.
- Wegen dieser weitgehenden Übereinstimmung zwischen Euklidischen und unitären Vektorräumen spricht man als Überbegriff von einem *Innenproduktraum* mit dem *inneren Produkt* $\langle \ | \ \rangle$.
- In der Analysis (dann meist der unendlichdimensionale Fall) spricht man nach *David Hilbert (1862 - 1943)* auch von einem *Prähilbertraum.*
- Für den zugrundeliegenden Körper $\mathbb{R}$ oder $\mathbb{C}$ eines Prähilbertraumes schreibt man meist einfach $\mathbb{K}$, um nicht dauernd die beiden Körper auflisten zu müssen.
- Betrachtet man in einem Prähilbertraum $(V, \langle \ | \ \rangle)$ das Skalarprodukt $\langle v|w\rangle$ zweier Vektoren $v, w \in V$, so ist bekanntlich die Abbildung
$$\ell_v : \begin{cases} V \to \mathbb{K} \\ w \mapsto \langle v|w\rangle \end{cases}$$
eine Linearform, also ein Element von V^*.
- Deshalb hat sich vor allem in der Quantenmechanik die Diracsche Schreibweise $\langle v| = \ell_v$ für die Linearform und $|w\rangle = w$ für den Vektor aus V eingebürgert.
- $\langle \ |$ wird dabei *Bra* genannt, $| \ \rangle$ *Ket*, denn zusammen ergeben die das spitze Klammernpaar $\langle \ | \ \rangle$ (engl.: bracket).

17 Normale Endomorphismen

17.1 Adjungierte

- Im Folgenden wird studiert, wie sich Endomorphismen mit einem gegebenen Skalarprodukt „vertragen“.

Beispiel 17.1.1. *Es sei* $\langle v|w\rangle = {}^t(\overline{v})\,w$ *das komplexe kanonische Skalarprodukt im* $\mathbb{C}^n$ *und* $f(x) = Ax \in \mathrm{End}(\mathbb{C}^n)$*. Dann gilt*

$$\langle f(v)|w\rangle = {}^t\!\left(\overline{f(v)}\right) w = {}^t(\overline{Av})\, w = {}^t(\overline{v})\ {}^t(\overline{A})\, w = {}^t(\overline{v})\, A^* w = \langle v|f^*(w)\rangle\,,$$

wobei $f^* \in \mathrm{End}(\mathbb{C}^n)$ *durch die zu* A *adjungierte Matrix* A^* *gegeben sei.*

Definition 17.1.2. *Es seien* $(V, \langle\ |\ \rangle)$ *ein Innenproduktraum,* $f \in \mathrm{End}(V)$*. Der Endomorphismus* $f^* \in \mathrm{End}(V)$ *heißt zu* f *adjungiert, wenn gilt*

$$\langle f(v)|w\rangle = \langle v|f^*(w)\rangle \quad \forall v, w \in V\,.$$

Satz 17.1.3. *Es sei* $(V, \langle\ |\ \rangle)$ *ein endlichdimensionaler Innenproduktraum. Zu jedem Endomorphismus* $f \in \mathrm{End}(V)$ *gibt es eine eindeutige Adjungierte* $f^* \in \mathrm{End}(V)$*.*

Beweis: Zuerst wird die Existenz von f^* gezeigt: Dazu wählt man sich eine Orthonormalbasis $B = (b_1, \ldots, b_n)$ von V. Für $v \in V$ gilt dann nach 15.5.4

$$v_{/B} = \begin{pmatrix} \langle b_1|v\rangle \\ \vdots \\ \langle b_n|v\rangle \end{pmatrix} = \begin{pmatrix} \overline{\langle v|b_1\rangle} \\ \vdots \\ \overline{\langle v|b_n\rangle} \end{pmatrix} \quad \text{bzw.} \quad v = \sum_{i=1}^n \overline{\langle v|b_i\rangle} b_i\,.$$

und somit $f(v) = \sum_{i=1}^n \overline{\langle v|b_i\rangle} f(b_i)$, also

$$\begin{aligned}\langle f(v)|w\rangle &= \langle \sum_{i=1}^n \overline{\langle v|b_i\rangle} f(b_i)|w\rangle = \sum_{i=1}^n \langle v|b_i\rangle \langle f(b_i)|w\rangle = \\ &= \sum_{i=1}^n \langle v|\langle f(b_i)|w\rangle b_i\rangle = \langle v|\sum_{i=1}^n \langle f(b_i)|w\rangle b_i\rangle = \\ &= \langle v|f^*(w)\rangle\end{aligned}$$

Ist nun $g \in \mathrm{End}(V)$ ein weiterer Endomorphismus mit

$$\langle f(v)|w\rangle = \langle v|f^*(w)\rangle = \langle v|g(w)\rangle \quad \forall v, w \in V\,,$$

so folgt

$$\begin{aligned}\langle v|f^*(w)\rangle - \langle v|g(w)\rangle &= 0 \iff \langle v|f^*(w) - g(w)\rangle = 0 \iff \\ \langle v|(f^* - g)(w)\rangle &= 0 \quad \forall v, w \in V\,.\end{aligned}$$

Da ein Skalarprodukt insbesondere nicht ausgeartet ist, folgt aus

$$\langle v|(f^* - g)(w)\rangle = 0 \quad \forall v \in V\,,$$

dass

$$(f^* - g)(w) = 0 \quad \forall w \in V, \text{ also } f^* - g = 0 \iff f^* = g\,,$$

was die Eindeutigkeit zeigt. □

- Nachdem jetzt gezeigt ist, dass es zu jedem Endomorphismus $f \in \mathrm{End}(V)$ eine eindeutige Adjungierte $f^* \in \mathrm{End}(V)$ gibt, kann man zu der eine Abbildungsmatrix berechnen.

- Dazu wählt man sich eine geordnete *Orthonormalbasis* $B = (b_1, \ldots, b_n)$ des endlichdimensionalen Innenproduktraumes V.

- Aus dem letzten Beweis liest man dann für $w \in V$ ab

$$f^*(w) = \sum_{i=1}^{n} \langle f(b_i)|w\rangle b_i, \text{ also } f^*(b_j) = \sum_{i=1}^{n} \langle f(b_i)|b_j\rangle b_i \iff$$

$$f^*(b_j)_{/B} = \begin{pmatrix} \langle f(b_1)|b_j\rangle \\ \vdots \\ \langle f(b_n)|b_j\rangle \end{pmatrix},$$

- Für die Abbildungsmatrix heißt das

$$[f^*]_B = \begin{pmatrix} \langle f(b_1)|b_1\rangle & \dots & \langle f(b_1)|b_n\rangle \\ \vdots & & \vdots \\ \langle f(b_n)|b_1\rangle & \dots & \langle f(b_n)|b_n\rangle \end{pmatrix}$$

- Es gilt gemäß 15.5.4: $f(b_j) = \sum_{i=1}^{n} \langle b_i|f(b_j)\rangle b_i = \sum_{i=1}^{n} \overline{\langle f(b_j)|b_i\rangle} b_i$
- Für die Abbildungsmatrix heißt das

$$[f]_B = \begin{pmatrix} \overline{\langle f(b_1)|b_1\rangle} & \dots & \overline{\langle f(b_n)|b_1\rangle} \\ \vdots & & \vdots \\ \overline{\langle f(b_1)|b_n\rangle} & \dots & \overline{\langle f(b_n)|b_n\rangle} \end{pmatrix}$$

- Damit ist gezeigt:

$$[f^*]_B = ([f]_B)^* \text{ für eine Orthonormalbasis } B\,. \tag{17.1}$$

- *Vorsicht*: Wegen der Verwendung von 15.5.4 ist das ohne die Bedingung für die Basis i.A. falsch!

17.2 Orthogonale und unitäre Endmorphismen

Definition 17.2.1. *Es sei* $(V, \langle\ |\ \rangle)$ *ein Euklidischer Vektorraum. Ein Endomorphismus* $f \in \mathrm{End}(V)$ *heißt orthogonal, wenn gilt*

$$\langle v|w\rangle = \langle f(v)|f(w)\rangle \quad \forall v, w \in V\,.$$

Definition 17.2.2. *Es sei* $(V, \langle\ |\ \rangle)$ *ein unitärer Vektorraum. Ein Endomorphismus* $f \in \mathrm{End}(V)$ *heißt unitär, wenn gilt*

$$\langle v|w\rangle = \langle f(v)|f(w)\rangle \quad \forall v, w \in V\,.$$

- Mit der zu f adjungierten Abbildung f^* liest man aus der Definition ab

$$\langle f(v)|f(w)\rangle = \langle v|f^*(f(w))\rangle = \langle v|w\rangle \quad \forall v, w \in V\,.$$

- Das heißt

$$f^*(f(w)) = w \quad \forall w \in V \iff f^* \circ f = \mathrm{id} \iff f^* = f^{-1}\,.$$

- Für die zugehörigen Abbildungs-Matrizen heißt das

$$[f^*]_B \cdot [f]_B = \mathbb{1}_n \overset{17.1}{\iff} ([f]_B)^* \cdot [f]_B = \mathbb{1}_n \iff ([f]_B)^* = ([f]_B)^{-1}\,. \tag{17.2}$$

- *Vorsicht*: Wegen der Verwendung von 17.1 gilt das wieder nur für eine ONB B!

Definition 17.2.3. *Eine Matrix* $A \in \mathbb{R}^{n\times n}$ *heißt orthogonal, wenn gilt*

$${}^tAA = \mathbb{1}_n \iff {}^tA = A^{-1}\,.$$

Eine Matrix $A \in \mathbb{C}^{n\times n}$ *heißt unitär, wenn gilt*

$$A^*A = \mathbb{1}_n \iff A^* = A^{-1}\,.$$

Folgerung 17.2.4. *Ein Endomorphismus f eines endlichdimensionalen Euklidischen Vektorraumes ist genau dann orthogonal, wenn die zugehörige Abbildungsmatrix bezüglich einer Orthonormalbasis orthogonal ist.*
Ein Endomorphismus f eines endlichdimensionalen unitären Vektorraumes ist genau dann unitär, wenn die zugehörige Abbildungsmatrix bezüglich einer Orthonormalbasis unitär ist.

- Aus ${}^tAA = \mathbb{1}_n$ folgt, dass eine quadratische reelle Matrix genau dann orthogonal ist, wenn ihre Spaltenvektoren (bzw. Zeilenvektoren) bezüglich des Standard-Skalarproduktes im $\mathbb{R}^n$ eine Orthonormalbasis bilden.

- Entsprechend folgt aus $A^*A = \mathbb{1}_n$, dass eine quadratische komplexe Matrix genau dann unitär ist, wenn ihre Spaltenvektoren (bzw. Zeilenvektoren) bezüglich des Standard-Skalarproduktes im $\mathbb{C}^n$ eine Orthonormalbasis bilden.

- Für eine orthogonale Matrix A folgt aus ${}^tAA = \mathbb{1}_n$ durch beidseitiges Anwenden der Determinante
$$\det(A)^2 \overset{12.3.1}{=} \det({}^tA)\cdot\det(A) \overset{12.5.1}{=} \det({}^tAA) = \det(\mathbb{1}_n) = 1\,, \tag{17.3}$$
d.h. es gilt $\det(A) = \pm 1$.

- Für eine unitäre Matrix A folgt aus $A^*A = \mathbb{1}_n$ durch beidseitiges Anwenden der Determinante
$$|\det(A)|^2 = \overline{\det(A)}\det(A) \overset{12.3.1}{=} \det({}^t(\overline{A}))\det(A) = \det(A^*)\det(A)$$
$$\overset{12.5.1}{=} \det(A^*A) = \det(\mathbb{1}_n) = 1\,,$$
d.h. es gilt $|\det(A)| = 1$.

- Es sei $(V, \langle\ |\ \rangle)$ ein Innenproduktraum und $f \in \mathrm{End}(V)$ orthogonal bzw. unitär.

- Ist λ ein Eigenwert von f (komplexe Eigenwerte gibt es immer) und $v \in V$ ein zugehöriger Eigenvektor, so gilt
$$\langle v|v\rangle = \langle f(v)|f(v)\rangle = \langle \lambda v|\lambda v\rangle = \overline{\lambda}\lambda\langle v|v\rangle = |\lambda|^2 \underbrace{\langle v|v\rangle}_{\neq 0} \Rightarrow |\lambda|^2 = 1 \tag{17.4}$$
d.h. die Eigenwerte von f haben den Betrag 1.

- *Vorsicht*: Dies heißt nicht, dass es überhaupt reelle Eigenwerte gibt!
- f hat insbesondere nicht den Eigenwert 0, d.h. $\mathrm{Kern}\, f = \{0\}$.
- Solch ein Endomorphismus ist nach 11.1.8 injektiv, im endlichdimensionalen Fall nach 11.2.6 sogar bijektiv.
- Ein orthogonaler oder unitärer Endomorphismus eines endlichdimensionalen Innenproduktraumes ist also ein Automorphismus.

Definition 17.2.5. *Es sei $(V, \langle\ |\ \rangle)$ ein Innenproduktraum. Ein $f \in \mathrm{End}(V)$ heißt*

- *längentreu, wenn $\|v\| = \|f(v)\|$ für alle $v \in V$,*
- *abstandstreu, wenn $d(v,w) = d(f(v), f(w))$ für alle $v, w \in V$,*
- *und im reellen Fall winkeltreu, wenn $\sphericalangle(v,w) = \sphericalangle(f(v), f(w))$ für alle $v, w \in V$.*

Folgerung 17.2.6. *Es sei $(V, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum. Jeder orthogonale Endomorphismus von V ist längen-, abstands- und winkeltreu.*
Es sei $(V, \langle\ |\ \rangle)$ ein unitärer Vektorraum. Jeder unitäre Endomorphismus von V ist längen- und abstandstreu.

Satz 17.2.7. *Es seien $(V, \langle\ |\ \rangle)$ ein Euklidischer Vektorraum und $f \in \mathrm{End}(V)$ längentreu. Dann ist f orthogonal.*
Es seien $(V, \langle\ |\ \rangle)$ ein unitärer Vektorraum und $f \in \mathrm{End}(V)$ längentreu. Dann ist f unitär.

Beweis: Für ein reelles Skalarprodukt $\langle\ |\ \rangle$ lautet die Polarisierungsformel 14.3.5
$$\langle v|w\rangle = \frac{1}{2}\left(\|v+w\|^2 - \|v\|^2 - \|w\|^2\right)$$
In der Übung 12.5 wird für den komplexen Fall die entsprechende Formel
$$\langle v|w\rangle = \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2 - i\,\|v+iw\|^2 + i\,\|v-iw\|^2\right)$$
nachgerechnet.
Damit folgt jeweils aus $\|v\| = \|f(v)\|$ für alle $v \in V$, dass sogar $\langle v|w\rangle = \langle f(v)|f(w)\rangle$ für alle $v, w \in V$ gilt. $\square$

Definition 17.2.8. *Es sei $B = (b_1, \ldots, b_n)$ eine Basis eines $\mathbb{R}$-Vektorraumes V. Vektoren $a_1, \ldots, a_n \in V$ heißen dann positiv orientiert bezüglich B, wenn $\det({a_1}_{/B}, \ldots, {a_n}_{/B}) > 0$ ist, negativ orientiert, wenn $\det({a_1}_{/B}, \ldots, {a_n}_{/B}) < 0$ ist.*

- Ist $f \in \mathrm{End}(V)$ mit $f(v)_{/B} = A \cdot v_{/B}$, so folgt $\det(f(a_1)_{/B}, \ldots, f(a_n)_{/B}) = \det(A \cdot {a_1}_{/B}, \ldots, A \cdot {a_n}_{/B}) = \det(A) \cdot \det({a_1}_{/B}, \ldots, {a_n}_{/B})$, d.h. für $\det(A) > 0$ hat $(f(a_1), \ldots, f(a_n))$ die gleiche Orientierung wie $(a_1, \ldots, a_n)$, für $\det(A) < 0$ die entgegengesetzte Orientierung.

Definition 17.2.9. *Es sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum. Ein $f \in \mathrm{End}(V)$ heißt orientierungstreu, wenn $\det(f) > 0$ ist.*

Definition 17.2.10. *Es sei (V, φ) ein Euklidischer Vektorraum. Ein $f \in \mathrm{End}(V)$ heißt volumentreu, wenn für jedes Parallelotop $P(a_1, \ldots, a_m)$ in V gilt:*

$$\mathrm{Vol}_\varphi(a_1, \ldots, a_m) = \mathrm{Vol}_\varphi(f(a_1), \ldots, f(a_m))$$

Satz 17.2.11. *Ein Endomorphismus f eines endlichdimensionalen Euklidischen Vektorraumes (V, φ) ist genau dann volumentreu, wenn $|\det(f)| = 1$ ist.*

Beweis: Aus Zeitgründen werden Details weggelassen. Man zeigt, dass sich bei Anwendung eines $f \in \mathrm{End}(V)$ auf ein *Parallelotop* $P(b_1, \ldots, b_m)$ das Volumen um $|\det f|$ ändert, also

$$\mathrm{Vol}_\varphi(f(b_1), \ldots, f(b_m)) = |\det(f)| \cdot \mathrm{Vol}_\varphi(b_1, \ldots, b_m) \ .$$

□

- Die folgenden Menge bilden jeweils zusammen mit der Matrizenmultiplikation Gruppen:

Definition	Bezeichnung	Eigenschaften
$\mathrm{GL}(n, K) :=$ $\{A \in K^{n\times n}\} \mid \det(A) \neq 0\}$	Lineare Gruppe	Invertierbare Endomorphismen
$\mathrm{SL}(n, K) :=$ $\{A \in \mathrm{GL}(n, K)\} \mid \det(A) = 1\}$	Spezielle lineare Gruppe	Volumen- und orientierungstreue Endomorphismen
$\mathrm{O}(n) :=$ $\{A \in \mathrm{GL}(n, \mathbb{R}) \mid {}^tAA = \mathbb{1}_n\}$	Orthogonale Gruppe	Abstandstreue Endomorphismen des $\mathbb{R}^n$
$\mathrm{SO}(n) :=$ $\{A \in \mathrm{O}(n) \mid \det(A) = 1\}$	Spezielle orthogonale Gruppe	Abstands- und orientierungstreue End. des $\mathbb{R}^n$
$\mathrm{U}(n) :=$ $\{A \in \mathrm{GL}(n, \mathbb{C}) \mid A^*A = \mathbb{1}_n\}$	Unitäre Gruppe	Abstandstreue Endomorphismen des $\mathbb{C}^n$
$\mathrm{SU}(n) :=$ $\{A \in \mathrm{U}(n) \mid \det(A) = 1\}$	Spezielle unitäre Gruppe	Abstands- und orientierungstreue End. des $\mathbb{C}^n$

17.3 Der Satz von Schur und seine Folgen

Definition 17.3.1. *Es seien $(V, \langle\ |\ \rangle)$ ein Innenproduktraum. Ein $f \in \mathrm{End}(V)$ heißt selbstadjungiert, wenn gilt*

$$\langle f(v)|w\rangle = \langle v|f(w)\rangle \quad \forall v, w \in V \iff f = f^*$$

und schiefadjungiert, falls $f = -f^*$.

Folgerung 17.3.2. *Für die Abbildungsmatrix A eines selbstadjungierten Endomorphismus bezüglich einer Orthonormalbasis gilt gemäß 17.1 $A = A^*$, d.h. im reellen Fall ist A symmetrisch, im komplexen Fall Hermitesch.*

Folgerung 17.3.3. *Das charakteristische Polynom eines selbstadjungierten Endomorphismus f eines n-dimensionalen Innenproduktraumes $(V, \langle\ |\ \rangle)$ hat gemäß 14.4.4 bzw. 16.3.7 n reelle Eigenwerte, d.h. f hat zumindest eine reelle Jordan-Normalform.*

Satz 17.3.4. *Es sei f ein selbstadjungierter Endomorphismus eines n-dimensionalen Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$. Dann gilt: Die Eigenräume von f zu paarweise verschiedenen Eigenwerten sind orthogonal.*

Beweis: Wird in der Übung gemacht. $\square$

Definition 17.3.5. *Es sei $(V, \langle\ |\ \rangle)$ ein endlichdimensionaler Innenproduktraum mit der Basis B, $f \in \operatorname{End}(V)$ mit $A := [f]_B$.*
Für $\mathbb{K} = \mathbb{R}$ heißt f bzw. A orthogonal trigonalisierbar, wenn es eine Orthonormalbasis C von V gibt, so dass $\tilde{A} := [f]_C$ Dreiecksgestalt hat, oder gleichbedeutend, wenn es eine orthogonale Transformationsmatrix $T = {}_B[\mathrm{id}]_C$ gibt, so dass $\tilde{A} := [f]_C$ Dreiecksgestalt hat, also

$$ {}^tT = T^{-1} \text{ (orthogonale Matrix) und } T^{-1}AT = {}^tTAT = \tilde{A}\,. $$

Entsprechend heißt f bzw. A für $\mathbb{K} = \mathbb{C}$ unitär trigonalisierbar, wenn es eine Orthonormalbasis C von V gibt, so dass $\tilde{A} := [f]_C$ Dreiecksgestalt hat, oder gleichbedeutend, wenn es eine unitäre Transformationsmatrix $T = {}_B[\mathrm{id}]_C$ gibt, so dass $\tilde{A} := [f]_C$ Dreiecksgestalt hat, also

$$ T^* = T^{-1} \text{ (unitäre Matrix) und } T^{-1}AT = T^*AT = \tilde{A}\,. $$

Entsprechend spricht man von orthogonaler bzw. unitärer Diagonalisierbarkeit, wenn $\tilde{A}$ sogar eine Diagonalmatrix ist.

Satz 17.3.6 (Satz von Schur)**.** *Ein Endomorphismus f eines n-dimensionalen Innenproduktraumes $(V, \langle\ |\ \rangle)$ ist für $\mathbb{K} = \mathbb{R}$ orthogonal trigonalisierbar, falls er n Eigenwerte in $\mathbb{R}$ hat, bzw. unitär trigonalisierbar für $\mathbb{K} = \mathbb{C}$.*

Beweis: Der Beweis verläuft per Induktion nach n. Für $n = 1$ ist nichts zu zeigen. Nun nimmt man an, dass die Behauptung für $n-1$ stimmt. Man kann sich einen Eigenwert $\lambda_1 \in \mathbb{K}$ und ein zugehöriges $b_1 \in E_f(\lambda_1) \setminus \{0\}$ nehmen (da kann man ohne Einschränkung der Allgemeinheit gleich einen normierten Vertreter nehmen, also einen mit $\|b_1\| = 1$). Gemäß 15.5.3 kann man $\{b_1\}$ zu einer geordneten Orthonormalbasis $C = (b_1, c_2, \dots, c_n)$ von V ergänzen. Bezüglich dieser Basis gilt

$$ [f]_C = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & & & \\ \vdots & & \boxed{D} & \\ 0 & & & \end{pmatrix} $$

mit einer $(n-1) \times (n-1)$-Matrix D. Nach dem Kästchensatz für Determinanten (→Übung 8.6) gilt

$$ \begin{aligned} \chi_f(\lambda) = \det &\begin{pmatrix} \lambda_1 - \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & \boxed{D - \lambda \mathbb{1}_{n-1}} & \\ 0 & & & \end{pmatrix} = \\ &= (\lambda_1 - \lambda) \cdot \det(D - \lambda \mathbb{1}_{n-1}) \Rightarrow \chi_D(\lambda) = \frac{\chi_f(\lambda)}{\lambda_1 - \lambda}\,. \end{aligned} $$

Da nach Voraussetzung $\chi_f(\lambda)$ über $\mathbb{K}$ in Linearfaktoren zerfällt, gilt dies also auch für $\chi_D(\lambda)$. Die Matrix D kann man deuten als Abbildungsmatrix eines Endomorphismus g von $W := \operatorname{Lin}(c_2, \dots, c_n)$ mit $V = \operatorname{Lin}(b_1) \oplus W$. Nach Induktionsvoraussetzung gibt es deshalb zu g eine geordnete Orthonormalbasis $(b_2, \dots, b_n)$ von W, bezüglich der g obere Dreiecksgestalt hat. Es gibt also eine Transformationsmatrix T mit

$$ T^{-1} \cdot D \cdot T = \begin{pmatrix} \lambda_2 & * & * & \dots & * \\ 0 & \lambda_3 & * & \dots & * \\ \vdots & \ddots & \ddots & & * \\ & & \ddots & \ddots & * \\ 0 & \dots & & 0 & \lambda_n \end{pmatrix} $$

Dabei sind $\lambda_2, \dots, \lambda_n$ die Eigenwerte von D, also bis auf λ_1 genau die Eigenwerte von f (da können auch mehrfache Eigenwerte dabei sein). Da T den Basiswechsel von der geordneten ON-Basis $(c_2, \dots, c_n)$ von W auf die geordnete ON-Basis $(b_2, \ldots, b_n)$ von W beschreibt, ist T nun aber unitär (bzw. im Reellen orthogonal), d.h. es ist $T^{-1} = T^*$. Mit den Rechengesetzen für Blockmatrizen (→Übung 8.6) folgt

$$\begin{pmatrix} 1 & 0 \ldots 0 \\ 0 & \\ \vdots & \boxed{T^*} \\ 0 & \end{pmatrix} \begin{pmatrix} \lambda_1 & * \ldots * \\ 0 & \\ \vdots & \boxed{D} \\ 0 & \end{pmatrix} \begin{pmatrix} 1 & 0 \ldots 0 \\ 0 & \\ \vdots & \boxed{T} \\ 0 & \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \ldots \quad * \\ 0 & \\ \vdots & \boxed{T^*DT} \\ 0 & \end{pmatrix} = \begin{pmatrix} \lambda_1 & * & * & \ldots & * \\ 0 & \lambda_2 & * & \ldots & * \\ \vdots & \ddots & \ddots & & * \\ & & \ddots & \ddots & * \\ 0 & \ldots & & 0 & \lambda_n \end{pmatrix}$$

□

Definition 17.3.7. *Es sei $(V, \langle \mid \rangle)$ ein Innenproduktraum. Ein Endomorphismus $f \in \mathrm{End}(V)$ heißt normal, wenn er mit seinem adjungierten Endomorphismus vertauschbar ist, d.h. wenn gilt*

$$f \circ f^* = f^* \circ f\,.$$

Entsprechend heißt eine Matrix $A \in \mathbb{K}^{n \times n}$ normal, wenn gilt

$$A \cdot A^* = A^* \cdot A\,.$$

Folgerung 17.3.8. *Wegen 17.1 folgt, dass eine Abbildung genau dann normal ist, wenn ihre Abbildungsmatrix bezüglich einer ON-Basis normal ist.*

Folgerung 17.3.9. *Normale Endomorphismen sind ein Oberbegriff für einige der bisher untersuchten Endomorphismen:*

$$\begin{array}{lcccl} f \text{ orthogonal (unitär)} & : & f^* = f^{-1} & \Rightarrow & f^* \circ f = \mathrm{id}_V = f \circ f^* \\ f \text{ selbstadjungiert} & : & f^* = f & \Rightarrow & f^* \circ f = f^2 = f \circ f^* \\ f \text{ schiefadjungiert} & : & f^* = -f & \Rightarrow & f^* \circ f = -f^2 = f \circ f^* \end{array} \tag{17.5}$$

d.h. orthogonale, unitäre, selbstadjungierte und schiefadjungierte Endomorphismen sind normal. Der folgende Satz rechtfertigt die Einführung dieses Sammelbegriffs:

Satz 17.3.10. *Ein normaler Endomorphismus f eines n-dimensionalen Innenproduktraumes $(V, \langle \mid \rangle)$ ist für $\mathbb{K} = \mathbb{R}$ orthogonal diagonalisierbar, falls er n Eigenwerten in $\mathbb{R}$ hat, bzw. unitär diagonalisierbar für $\mathbb{K} = \mathbb{C}$.*
Beweis: Nach dem Satz von Schur 17.3.6 gibt es eine ON-Basis C von V, so dass $[f]_C$ obere Dreiecksgestalt hat. Da $[f]_C$ normal ist, heißt das

$$\text{◹} \cdot \text{◺} = [f]_C \left([f]_C\right)^* = \left([f]_C\right)^* [f]_C = \text{◺} \cdot \text{◹}$$

Die Beweisidee sieht nun so aus, dass man diese Produkte von Dreiecksmatrizen auf beiden Seiten ausrechnet und die Elemente der Hauptdiagonalen miteinander vergleicht. Dazu sei $(a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}} := [f]_C$, also $\left([f]_C\right)^* = (\overline{a_{ji}})$.

Die Diagonalelemente von $[f]_C \left([f]_C\right)^*$ sind $\sum\limits_{i=1}^{j} \overline{a_{ij}} a_{ij}$ für $j = 1, \ldots, n$, die von $\left([f]_C\right)^* [f]_C$ sind $\sum\limits_{i=j}^{n} a_{ji} \overline{a_{ji}}$ für $j = 1, \ldots, n$. Gleichsetzen liefert also

$$\sum_{i=1}^{j} |a_{ij}|^2 = \sum_{i=1}^{j} \overline{a_{ij}} a_{ij} = \sum_{i=j}^{n} a_{ji} \overline{a_{ji}} = \sum_{i=j}^{n} |a_{ji}|^2$$

für $j = 1, \ldots, n$. Für $j = 1$ liest man ab $|a_{11}|^2 = \sum\limits_{i=1}^{n} |a_{1i}|^2 \iff \sum\limits_{i=2}^{n} |a_{1i}|^2 = 0$, also $a_{1i} = 0$ für $i = 2, \ldots, n$.

Für $j = 2$ liest man ab $\underbrace{|a_{12}|^2}_{=0} + |a_{22}|^2 = \sum\limits_{i=2}^{n} |a_{2i}|^2 \iff \sum\limits_{i=3}^{n} |a_{2i}|^2$, also $a_{2i} = 0$ für $i = 3, \ldots, n$.

…usw., für $j = n$ folgt schließlich $a_{n-1,n} = 0$.
Damit ist gezeigt, dass alle Nebendiagonalelemente der Matrix $[f]_C$ verschwinden, d.h. $[f]_C$ ist sogar eine Diagonalmatrix!

□

Folgerung 17.3.11. *Ist V ein n-dimensionaler Euklidischer Vektorraum und $f \in \mathrm{End}(V)$ selbstadjungiert, so ist die Abbildungsmatrix $A := [f]_B \in \mathbb{R}^{n \times n}$ bezüglich einer Orthonormalbasis B von V symmetrisch.*
Nach 17.5 sind f bzw. A normal und besitzen nach 14.4.4 n reelle Eigenwerte. Somit sind f bzw. A nach 17.3.10 orthogonal diagonalisierbar.

Folgerung 17.3.12. *Da nach 17.5 unitäre oder schiefadjungierte Endomorphismen eines unitären Vektorraums V ebenfalls normal sind und diese nach dem Fundamentalsatz 13.3.2 n Eigenwerte in $\mathbb{C}$ haben, sind sie (bzw. die bzgl. einer ON-Basis von V zugehörigen unitären oder schiefhermiteschen Matrizen) unitär diagonalisierbar.* 27.1.11

Beispiel 17.3.13. *Gegeben sei die Matrix*

$$A := \begin{pmatrix} 17 & 8 & 32 \\ 8 & 80 & -4 \\ 32 & -4 & 65 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

bzw. der Endomorphismus $f_A : x \mapsto Ax$ des $\mathbb{R}^3$ bzgl. der Standardbasis. Wegen $A = {}^tA$ ist A symmetrisch bzw. f_A selbstadjungiert, also auch normal nach 17.5.
Wegen 14.4.4 hat A drei reelle Eigenwerte. Im vorliegenden Fall rechnet man nach, dass $\lambda_{1,2} = 81$ und $\lambda_3 = 0$ ist.
Es ist günstig, zuerst den Eigenraum zu dem einfachen Eigenwert 0 zu berechnen. Hier ergibt sich

$$E_A(0) = \operatorname{Kern}(A) = \operatorname{Lin}\begin{pmatrix} -8 \\ 1 \\ 4 \end{pmatrix}$$

Da A nach 17.3.10 diagonalisierbar ist, gehört zu dem doppelten Eigenwert $\lambda_{1,2} = 81$ auch ein 2-dimensionaler Eigenraum.
Da A nach 17.3.10 orthogonal diagonalisierbar ist, sind die Eigenräume sogar orthogonal zueinander, hier also

$$E_A(81) \perp E_A(0) \text{ bzw. aus Dimensionsgründen } E_A(81) = E_A(0)^{\perp}.$$

Man kann also $E_A(81)$ wie üblich berechnen, sollte aber besser die Orthogonalität zu Hilfe nehmen:
Da z.B. $\left(\begin{smallmatrix} 0 \\ 4 \\ -1 \end{smallmatrix}\right)$ zu $E_A(0)$ orthogonal ist, liegt dieser Vektor in $E_A(81)$.
Ein 3.Eigenvektor von A bzw. 2.Eigenvektor aus $E_A(81)$ lässt sich natürlich ebenso bestimmen. Da aber insgesamt alle Eigenvektoren paarweise orthogonal sein sollen, muss der auch zu dem oben gewählten 1.Eigenvektor aus $E_A(81)$ orthogonal sein.
Im $\mathbb{R}^3$ geht das z.B. mit dem Vektorprodukt (in höherer Dimension mit dem Skalarprodukt):

$$\begin{pmatrix} -8 \\ 1 \\ 4 \end{pmatrix} \times \begin{pmatrix} 0 \\ 4 \\ -1 \end{pmatrix} = \begin{pmatrix} -17 \\ -8 \\ -32 \end{pmatrix} \Rightarrow E_A(81) = \operatorname{Lin}\left(\begin{pmatrix} 0 \\ 4 \\ -1 \end{pmatrix}, \begin{pmatrix} -17 \\ -8 \\ -32 \end{pmatrix}\right)$$

Für die neue ONB aus Eigenvektoren muss man noch normieren und erhält

$$\begin{pmatrix} -\frac{8}{9} & 0 & -\frac{1}{9}\sqrt{17} \\ \frac{1}{9} & \frac{4}{17}\sqrt{17} & -\frac{8}{153}\sqrt{17} \\ \frac{4}{9} & -\frac{1}{17}\sqrt{17} & -\frac{32}{153}\sqrt{17} \end{pmatrix}$$

Mit dieser orthogonalen Matrix T bzw. mit dieser ONB aus Eigenvektoren gilt

$$T^{-1}AT = {}^tTAT = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 81 & 0 \\ 0 & 0 & 81 \end{pmatrix}.$$

Die Wahl der Vektoren aus $E_A(81)$ ist natürlich nicht eindeutig. Mit

$$\begin{pmatrix} 1 \\ -8 \\ 4 \end{pmatrix} \perp E_A(0) = \operatorname{Lin}\left(\begin{pmatrix} -8 \\ 1 \\ 4 \end{pmatrix}\right) \Rightarrow \begin{pmatrix} 1 \\ -8 \\ 4 \end{pmatrix} \times \begin{pmatrix} -8 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 36 \\ 6 \\ 63 \end{pmatrix}$$

erhält man z.B.

$$T = \begin{pmatrix} -\frac{8}{9} & \frac{1}{9} & \frac{4}{9} \\ \frac{1}{9} & -\frac{8}{9} & \frac{4}{9} \\ \frac{4}{9} & \frac{4}{9} & \frac{7}{9} \end{pmatrix}.$$

Beispiel 17.3.14. *Gegeben sei die Matrix*

$$A := \begin{pmatrix} -64 & 8 & -32\,i \\ 8 & -1 & 4\,i \\ 32\,i & -4\,i & -16 \end{pmatrix} \in \mathbb{C}^{3\times 3}.$$

bzw. der Endomorphismus $f_A : x \mapsto Ax$ des $\mathbb{C}^3$ bzgl. der Standardbasis. Wegen $A = {}^t\!\left(\overline{A}\right) = A^*$ ist A Hermitesch bzw. f_A selbstadjungiert, also auch normal nach 17.5.
Wegen 16.3.7 hat A drei reelle Eigenwerte. Im vorliegenden Fall rechnet man nach, dass $\lambda_{1,2} = 0$ und $\lambda_3 = -81$ ist.
Wie im letzten Beispiel ist es günstig, zuerst den Eigenraum zu dem einfachen Eigenwert -81 zu berechnen. Hier ergibt sich

$$v_1 := \begin{pmatrix} 8\,i \\ -i \\ 4 \end{pmatrix} \Rightarrow E_A\,(-81) = \mathrm{Lin}\,(v_1)$$

Da A nach 17.3.10 diagonalisierbar ist, gehört zu dem doppelten Eigenwert $\lambda_{1,2} = 0$ auch ein 2-dimensionaler Eigenraum.
Da A nach 17.3.10 unitär diagonalisierbar ist, sind die Eigenräume sogar orthogonal zueinander, hier also

$$E_A\,(0) \perp E_A\,(-81) \;\; \textit{bzw. aus Dimensionsgründen } E_A\,(0) = E_A\,(-81)^{\perp}\ .$$

Man kann also $E_A\,(0)$ wie üblich berechnen, sollte aber besser wieder die Orthogonalität (jetzt komplexes Skalarprodukt!) zu Hilfe nehmen:
Da z.B. $v_2 := \left(\begin{smallmatrix} -i \\ 0 \\ 2 \end{smallmatrix}\right)$ zu $E_A\,(-81)$ orthogonal ist, liegt dieser Vektor in $E_A\,(0)$.
Ein 3.Eigenvektor von A bzw. 2.Eigenvektor aus $E_A\,(0)$ lässt sich natürlich ebenso bestimmen. Da aber insgesamt alle Eigenvektoren paarweise orthogonal sein sollen, muss der auch zu dem oben gewählten 1.Eigenvektor aus $E_A\,(0)$ orthogonal sein.
Im $\mathbb{C}^3$ geht das nicht mit dem Vektorprodukt! Man setzt also an $\langle v_1|v_3\rangle = \langle v_2|v_3\rangle = 0$. Dies ist ein einfaches LGS für v_3. Es folgt

$$v_3 = \begin{pmatrix} 2 \\ 20 \\ -i \end{pmatrix} \Rightarrow E_A\,(81) = \mathrm{Lin}\left(\begin{pmatrix} -i \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 20 \\ -i \end{pmatrix}\right)$$

Für die neue ONB aus Eigenvektoren muss man noch normieren und erhält

$$T := \begin{pmatrix} \frac{8}{9}\,i & -\frac{1}{5}\,i\sqrt{5} & \frac{2}{45}\,\sqrt{5} \\ -\frac{1}{9}\,i & 0 & \frac{4}{9}\,\sqrt{5} \\ \frac{4}{9} & \frac{2}{5}\,\sqrt{5} & -\frac{1}{45}\,i\sqrt{5} \end{pmatrix}$$

Mit dieser unitären Matrix T bzw. mit dieser ONB aus Eigenvektoren gilt

$$T^{-1}AT = T^*AT = \begin{pmatrix} -81 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Beispiel 17.3.15. *Gegeben sei die Matrix*

$$A := \frac{1}{9}\begin{pmatrix} -1 & 8 & -4 \\ 4 & 4 & 7 \\ 8 & -1 & -4 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

bzw. der Endomorphismus $f_A : x \mapsto Ax$ des $\mathbb{R}^3$ bzgl. der Standardbasis. Wegen $A \cdot {}^tA = \mathbb{1}_3$ ist A bzw. f_A orthogonal, also auch normal nach 17.5. Gemäß 17.4 haben alle Eigenwerte von f den Betrag 1 und nach 17.3 ist $\det A = \pm 1$. Da da charakteristische Polynom vom Grad 3 ist, gibt es mindestens einen reellen Eigenwert, der also ± 1 ist. Hier ist $\lambda_1 = 1$, $\lambda_{2,3} = \frac{1}{9}\left(-5 \pm 2i\sqrt{14}\right)$ und $\det A = 1$, also $f_A \in \mathrm{SO}(3)$, d.h. A ist (neben dem üblichen längen-, abstands-, winkel- und volumentreu) orientierungstreu aber nicht reell orthogonal diagonalisierbar. Die Elemente der $\mathrm{SO}(3)$ sind Drehungen, Drehachse ist der Eigenraum zum Eigenwert 1. Da nicht alle Eigenwerte von A in $\mathbb{R}$ liegen, ist A nur über $\mathbb{C}$ unitär diagonalisierbar, so wie das etwa im letzten Beispiel gezeigt wurde.

17.4 Normalformen normaler Endomorphismen

- Wenn man, wie im letzten Beispiel, einen normalen Endomorphismus hat, der aber nichtreelle Eigenwerte hat, so ist man trotzdem auch an einer möglichst einfachen reellen Darstellung interessiert.
- Solch eine einfache reelle Darstellung kann man in der Tat über die komplexe Diagonalform berechnen.
- Diese so genannten *Spektralsätze* (die Menge der Eigenwerte eines Endomorphismus heißt auch sein Spektrum) werden hier ohne Beweis angegeben:

Satz 17.4.1 (Spektralsatz). *Es sei f ein normaler Endomorphismus eines n-dimensionalen Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$ mit den Eigenwerten $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ und $\lambda_{r+1} = \alpha_1 + i\beta_1, \dots, \lambda_{r+s} = \alpha_s + i\beta_s \in \mathbb{C}$ (mit $r + s = n$ und $\beta_1, \dots, \beta_s \neq 0$, Eigenwerte jeweils entsprechend ihrer Vielfachheit nummeriert). Dann existiert eine Orthonormalbasis B von V, so dass*

$$[f]_B = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_r \end{matrix}} & & & \mathbf{0} \\ & \boxed{\begin{matrix} \alpha_1 & -\beta_1 \\ \beta_1 & \alpha_1 \end{matrix}} & & \\ & & \ddots & \\ \mathbf{0} & & & \boxed{\begin{matrix} \alpha_s & -\beta_s \\ \beta_s & \alpha_s \end{matrix}} \end{pmatrix}$$

- Hat das charakteristische Polynom des Endomorphismus f nur reelle Nullstellen, so ist f nach 17.3.10 orthogonal diagonalisierbar.
- Hat das reelle charakteristische Polynom von f dagegen auch nichtreelle Nullstellen (wie im Satz), so treten diese bekanntlich paarweise konjugiert komplex auf. In diesem Fall f nach 17.3.10 unitär diagonalisierbar.
- Sind $\alpha \pm i\beta$ zwei konjugiert komplexe Eigenwerte, dann gehören dazu auch konjugiert komplexe Eigenvektoren, denn mit der Linearität von f rechnet man leicht nach[17]

$$\begin{aligned} f(v + iw) =& (\alpha + i\beta)(v + iw) \Rightarrow \\ f(v - iw) =& (\alpha - i\beta)(v - iw) \text{ mit Vektoren } v, w \text{ aus dem reellen } V\,. \end{aligned}$$

- Die komplexen Eigenvektoren sind orthogonal zueinander, d.h. (komplexes Skalarprodukt)

$$\begin{aligned} 0 =& \langle v + iw | v - iw \rangle = \langle v|v \rangle + \langle v|-iw \rangle + \langle iw|v \rangle + \langle iw|-iw \rangle = \\ =& \langle v|v \rangle - \langle w|w \rangle - i(\langle v|w \rangle + \langle w|v \rangle) \Rightarrow \\ \langle v|v \rangle =& \langle w|w \rangle,\ \langle v|w \rangle = -\langle w|v \rangle\,. \end{aligned} \tag{17.6}$$

- Ist $v + iw$ sogar normiert, so gilt

$$\begin{aligned} 1 =& \langle v + iw | v + iw \rangle = \langle v|v \rangle + \langle v|iw \rangle + \langle iw|v \rangle + \langle iw|iw \rangle = \\ =& \langle v|v \rangle + \langle w|w \rangle + i(\langle v|w \rangle - \langle w|v \rangle) \Rightarrow \\ 1 =& \langle v|v \rangle + \langle w|w \rangle,\ \langle v|w \rangle = \langle w|v \rangle\,. \end{aligned} \tag{17.7}$$

- Aus 17.6 und 17.7 zusammen folgt

$$\langle v|v \rangle = \frac{1}{2} = \langle w|w \rangle,\ \langle v|w \rangle = 0 = \langle w|v \rangle\,,$$

  d.h. $\sqrt{2}v$ und $\sqrt{2}w$ sind reell und orthonormal.

- Aus

$$f(v) + if(w) = f(v + iw) = (\alpha + i\beta)(v + iw) = \alpha v - \beta w + i(\alpha w + \beta v)$$

  folgt durch Vergleich von Real- und Imaginärteil

$$f(v) = \alpha v - \beta w\,,\ f(w) = \alpha w + \beta v\,, \tag{17.8}$$

[17] Benutzt man f mit komplexem Argument, so ist das streng genommen eine neue Funktion $\tilde{f}$ auf $\tilde{V} := \{v + iw \mid v, w \in V\}$. Es stiftet aber keinen Schaden, wenn man die auch f nennt.

- Bezüglich des Basisfragments (w, v) der insgesamt aufzubauenden geordneten Basis des reellen V bekommt man gerade den Teilblock

$$\boxed{\begin{matrix}\alpha & -\beta \\ \beta & \alpha\end{matrix}}$$

der gesamten Abbildungsmatrix.

Folgerung 17.4.2. *Es sei f ein schiefadjungierter Endomorphismus eines n-dimensionalen Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$. Solch ein f hat nur rein komplexe Eigenwerte $\lambda_j = i\beta_j$ (→Übung 13.7; dabei kann durchaus auch $\beta = 0$ sein). Damit sieht die Normalform laut Spektralsatz so aus:*

$$[f]_B = \begin{pmatrix} \boxed{\begin{matrix} 0 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 0 \end{matrix}} & & & \mathbf{0} \\ & \boxed{\begin{matrix} 0 & -\beta_1 \\ \beta_1 & 0 \end{matrix}} & & \\ & & \ddots & \\ \mathbf{0} & & & \boxed{\begin{matrix} 0 & -\beta_s \\ \beta_s & 0 \end{matrix}} \end{pmatrix}$$

Folgerung 17.4.3. *Es sei f ein orthogonler Endomorphismus eines n-dimensionalen Euklidischen Vektorraumes $(V, \langle\ |\ \rangle)$. Gemäß 17.4 haben alle Eigenwerte von f den Betrag 1. Deshalb kann man bei denen statt $\lambda = \alpha + i\beta$ auch schreiben $\lambda = \cos\varphi + i \sin\varphi$. Somit sieht die Normalform so aus:*

$$[f]_B = \begin{pmatrix} \boxed{\begin{matrix} 1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1 \end{matrix}} & & & & \mathbf{0} \\ & \boxed{\begin{matrix} -1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & -1 \end{matrix}} & & & \\ & & \boxed{\begin{matrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{matrix}} & & \\ & & & \ddots & \\ \mathbf{0} & & & & \boxed{\begin{matrix} \cos\varphi_s & -\sin\varphi_s \\ \sin\varphi_s & \cos\varphi_s \end{matrix}} \end{pmatrix}$$

- Es seien $c = \cos(\varphi)$, $s = \sin(\varphi)$ mit $0 < \varphi < 2\pi$, $\varphi \neq \pi$.

- Dann liefert der Spektralsatz die folgenden orthogonalen Endomorphismen des $\mathbb{R}^2$ und $\mathbb{R}^3$:

Normalform	Typ	EWe	det	sonst
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\mathrm{id}_{\mathbb{R}^2}$ =Drehung um 0	1,1	1	
$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	*Spiegelung* an x-Achse	1,-1	-1	$f^2 = \mathrm{id}_{\mathbb{R}^2}$, *involutorisch*
$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	*Punktspiegelung* an 0 = Drehung um π	-1,-1	1	$f^2 = \mathrm{id}_{\mathbb{R}^2}$, *involutorisch*
$\begin{pmatrix} c & -s \\ s & c \end{pmatrix}$	*Drehung* um $\varphi \neq 0, \pi$	$c \pm i\,s$	1	

Abbildung 39: Normalformen orthogonaler Endomorphismen des $\mathbb{R}^2$

Normalform	Typ	EWe	det	sonst
$\begin{pmatrix} 1&0&0\\0&1&0\\0&0&1 \end{pmatrix}$	$\mathrm{id}_{\mathbb{R}^3}$=Drehung um 0	1,1,1	1	
$\begin{pmatrix} 1&0&0\\0&1&0\\0&0&-1 \end{pmatrix}$	*Spiegelung* an $x-y$-Ebene	1,1,-1	-1	$f^2 = \mathrm{id}_{\mathbb{R}^3}$, involutorisch
$\begin{pmatrix} 1&0&0\\0&-1&0\\0&0&-1 \end{pmatrix}$	*Drehung* um x-Achse, Winkel π	1,-1,-1	1	$f^2 = \mathrm{id}_{\mathbb{R}^3}$, involutorisch
$\begin{pmatrix} -1&0&0\\0&-1&0\\0&0&-1 \end{pmatrix}$	*Punktspiegelung* an 0, Drehspiegelung um π	-1,-1,-1	-1	$f^2 = \mathrm{id}_{\mathbb{R}^3}$, involutorisch
$\begin{pmatrix} 1&0&0\\0&c&-s\\0&s&c \end{pmatrix}$	*Drehung* um x-Achse, Winkel φ	$1, c \pm i\, s$	1	$\cos(\varphi) = \frac{\mathrm{Spur}\, A-1}{2}$
$\begin{pmatrix} -1&0&0\\0&c&-s\\0&s&c \end{pmatrix}$	*Drehspiegelung* x-Achse, $y-z$-Ebene, Winkel φ	$-1, c \pm i\, s$	-1	$\cos(\varphi) = \frac{\mathrm{Spur}\, A+1}{2}$

Abbildung 40: Normalformen orthogonaler Endomorphismen des $\mathbb{R}^3$

Beispiel 17.4.4. *Betrachtet wird nochmal die Matrix aus 17.3.15:*

$$A = \frac{1}{9}\begin{pmatrix} -1 & 8 & -4 \\ 4 & 4 & 7 \\ 8 & -1 & -4 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Es war bereits gezeigt worden, dass $f_A \in \mathrm{SO}(3)$ ist und die Eigenwerte $\lambda_1 = 1$, $\lambda_{2,3} = \frac{1}{9}\left(-5 \pm 2i\sqrt{14}\right)$ hat. Die zugehörigen Eigenräume sind

$$E_A(1) = \mathrm{Lin}\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \; E_A\left(\frac{1}{9}\left(-5 \pm 2i\sqrt{14}\right)\right) = \mathrm{Lin}\begin{pmatrix} -2 \pm 3i\sqrt{14} \\ -3 \mp 2i\sqrt{14} \\ 13 \end{pmatrix}.$$

$E_A(1)$ ist (als reeller Vektorraum) die Drehachse der Drehung f_A.
Die berechneten komplexen Eigenräume sind bezüglich des kanonischen komplexen Skalarprodukts im $\mathbb{C}^3$ orthogonal zueinander.
Gemäß Vorüberlegung stellt man sich aus dieser komplexen Basis eine reelle ON-Basis zusammen, indem man die reellen Eigenvektoren normiert übernimmt und von den konjugiert komplexen Eigenvektoren jeweils den normierten Real- und Imaginärteil übernimmt:

$$b_1 := \frac{1}{\sqrt{14}}\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \in E_A(1),$$

$$b_2 := \frac{1}{\sqrt{13}}\begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix} = \mathrm{Im}(v) \text{ von } v \text{ aus } E_A\left(\frac{1}{9}\left(-5 \pm 2i\sqrt{14}\right)\right),$$

$$b_3 := \frac{1}{\sqrt{182}}\begin{pmatrix} -2 \\ -3 \\ 13 \end{pmatrix} = \mathrm{Re}(v) \text{ von } v \text{ aus } E_A\left(\frac{1}{9}\left(-5 \pm 2i\sqrt{14}\right)\right).$$

Wechselt man zur neuen Basis $B = (b_1, b_2, b_3)$ des $\mathbb{R}^3$ bzw. transformiert man mit der Matrix $T := (b_1, b_2, b_3)$, so folgt mit 17.8

$$T^{-1}AT = {}^t TAT = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{5}{9} & -\frac{2}{9}\sqrt{14} \\ 0 & \frac{2}{9}\sqrt{14} & -\frac{5}{9} \end{pmatrix}$$

Die entspricht dem 5.Eintrag in Abbildung 40 mit $\cos(\varphi) = -\frac{5}{9}$, $\sin(\varphi) = \frac{2}{9}\sqrt{14}$. *Den* $\cos(\varphi)$ *kann man auch direkt aus der ursprünglichen Matrix A mit der Formel* $\cos(\varphi) = \frac{\operatorname{Spur} A - 1}{2}$ *ablesen.*

Abbildungsverzeichnis

Literatur

Gerd Fischer. *Lineare Algebra*. Grundkurs Mathematik, 12.Auflage. Vieweg Studium, Braunschweig/Wiesbaden, 2000. ISBN 3-528-87217-9.

Johann Hartl. Ein einfacher Induktionsbeweis für die Existenz der Jordan-Normalform. *Archiv der Mathematik*, 50:323–327, 1988. ISSN 0003-889X. URL http://dx.doi.org/10.1007/BF01190226. 10.1007/BF01190226.

Max Koecher. *Lineare Algebra und analytische Geometrie*. Springer Lehrbuch, 4., erg. u. aktualisierte Aufl., Korr. Nachdruck. Springer-Verlag, Berlin, Heidelberg, ..., 1997. ISBN 978-3-540-62903-0.

Falko Lorenz. *Lineare Algebra II*. Hochschultaschenbuch. Spektrum Akademischer Verlag; Auflage: 3., überarb. A.; 4., korr. Nachdr. 2005., Heidelberg, Berlin, 2005.

Index