Notizen zur Vorlesung

Lineare Algebra für PhysikerInnen

von Günther Hörmann[1]
(nach einem Manuskript von Michael Grosser)

im Wintersemester 2012/13
an der Fakultät für Physik
Universität Wien

[1]Fakultät für Mathematik der Universität Wien, E-Mail: guenther.hoermann@univie.ac.at

Vorwort

Dieses Vorlesungsskriptum entsteht beziehungsweise entstand im Laufe des Wintersemesters 2012/13 begleitend zur dreistündigen Vorlesung *Lineare Algebra für PhysikerInnen* an der Universität Wien. Ich bedanke mich sehr herzlich bei meinem früheren Lehrer und nunmehrigen Kollegen Michael Grosser dafür, dass er mir sein Manuskript zur Verfügung gestellt hat, das er zu einer fast gleichnamigen Vorlesung vor zehn Jahren ausgearbeitet hatte. Ich entschuldige mich bei ihm und bei den aktuell betroffenen Studierenden für etwaige inhaltliche und stilistische Verschlechterungen, die ich dem ursprünglichen Text im Zuge meiner Neugestaltung vielleicht antun werde beziehungsweise angetan habe: Fehler und Ungereimtheiten in der aktuellen Fassung sind daher zur Gänze mir anzulasten, alle gut gelungenen Aspekte hingegen sind fast sicher Michael Grossers Verdienst.

Eine Warnung an die Studierenden: Das Vorlesungsskriptum ist nicht dafür geeignet, den Besuch der Vorlesung zu ersetzen. Das gilt selbst dann, wenn „eh alles im Skriptum steht, was der in der VO auf die Tafel schreibt“. Weiters ist dieses Skriptum auch nicht so entworfen oder formuliert worden, dass es wie ein Buch lesbar wäre und die Beschäftigung mit Fachbüchern überflüssig oder uninteressant machen wollte. Das gilt gerade dann, wenn „der in der VO sich mit dem Stoff eh an kein Buch hält“. Lesen Sie dazu zum Beispiel den (sehr kurzen) Abschnitt 1.4 im unten angeführten Buch von Jänich.[2]

Günther Hörmann, Wien, Juli 2012 bis Jänner 2013

Hier noch eine sehr kleine Auswahl an einführenden Büchern mit Kürzestkommentaren dazu: Jänich ist ein (kleines, feines und) vergnügliches Begleit-Lesebuch zu den wesentlichen Begriffen (mit einigen spezifischen Bemerkungen auch zur Physik). Fischers Bücher sind etwas umfangreicher und mehr mit Beispielen durchsetzt. Bröcker ist ein wenig abstrakter angelegt und gibt auch kleine Einführungen in weiterführende Theorien. Muthsam hat eben diese Vorlesung mehrmals gehalten und bringt eine Fülle von konkreten Anwendungen u.a. auch mit numerischen Aspekten.

Th. Bröcker, *Lineare Algebra und analytische Geometrie*, Birkhäuser Verlag.

G. Fischer: früher in zwei Bänden *Lineare Algebra* und *Analytische Geometrie* und neuerdings *Lernbuch Lineare Algebra und Analytische Geometrie: Das Wichtigste ausführlich für das Lehramts- und Bachelorstudium*, Vieweg+Teubner.

K. Jänich, *Lineare Algebra*, Springer-Verlag.

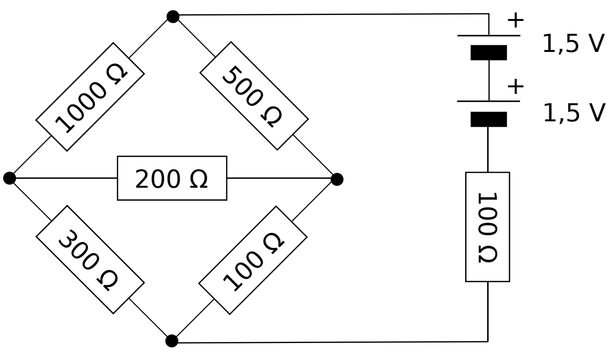
H. Muthsam, *Lineare Algebra und ihre Anwendungen*, Spektrum Akademischer Verlag.

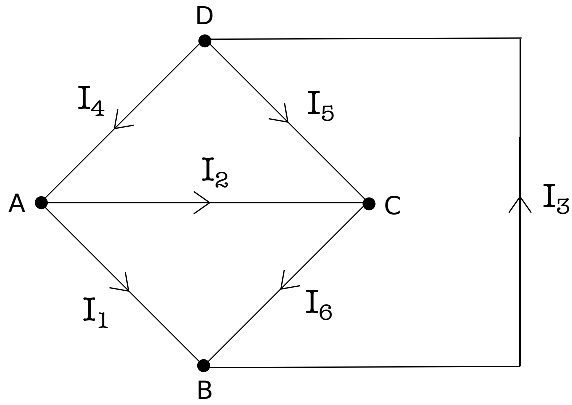
[2] Eine Kopie der betreffenden zwei Seiten lasse ich Studierenden gerne auf Anfrage zukommen.

Inhaltsverzeichnis

§1. LINEARE GLEICHUNGSSYSTEME I

1.1. Beispiel einer Brückenschaltung: In dem unten skizzierten elektrischen Netzwerk sollen die Ströme $I_1, \dots, I_6$ bestimmt werden.





Mit Hilfe der *Kirchhoffschen Gesetze* können wir Gleichungen für die Ströme in jedem Knoten und für die Spannungen in jeder Masche aufstellen.

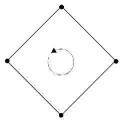
Knotengleichungen: Die Regel, dass die Summe der hereinfließenden Ströme in jedem der vier Knoten Null sein muss,

ergibt

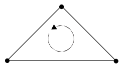
$$\begin{array}{ll} \text{bei A} & I_4 - I_1 - I_2 = 0, \\ \text{bei B} & I_1 + I_6 - I_3 = 0, \\ \text{bei C} & I_2 + I_5 - I_6 = 0, \\ \text{bei D} & I_3 - I_4 - I_5 = 0. \end{array}$$

Maschengleichungen: Die Summe der Spannungen beim orientierten Umlauf jeder Masche[3] muss Null sein. Die Spannungen an den einzelnen Widerständen erhalten wir jeweils als Produkt des elektrischen Widerstandes mit der Stromstärke. Für jene Maschen, die auch die beiden Spannungsquellen beinhalten, müssen wir zusätzlich eine Konstante entsprechend 3 V Spannung in die Bilanz aufnehmen. Wir finden sieben Maschen und erhalten

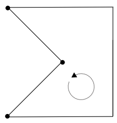
[3] ein geschlossener Weg im Netzwerk, der keine Verbindung zwischen zwei Knoten ein zweites Mal passiert



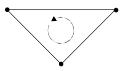
$$300I_1 - 100I_6 - 500I_5 + 1000I_4 = 0,$$



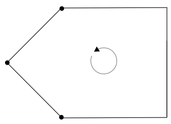
$$200I_2 - 500I_5 + 1000I_4 = 0,$$



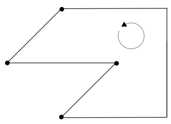
$$500I_5 + 100I_6 + 100I_3 = 3,$$



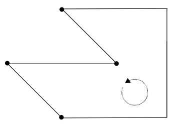
$$300I_1 - 100I_6 - 200I_2 = 0,$$



$$1000I_4 + 300I_1 + 100I_3 = 3,$$



$$1000I_4 + 200I_2 + 100I_6 + 100I_3 = 3,$$



$$500I_5 - 200I_2 + 300I_1 + 100I_3 = 3.$$

Insgesamt ergeben sich also elf Gleichungen für die sechs Unbekannten $I_1, \ldots, I_6$.

1.2. Fragen und Probleme im Zusammenhang mit solchen Beispielen:

1. Wie lautet ein zuverlässiges und halbwegs „sparsames“ Lösungsverfahren, das für jedes derartige Gleichungssystem anwendbar ist?
2. Elf Gleichungen und sechs Unbekannte – kann man nicht von Anfang an fünf Gleichungen weglassen und mit sechs Gleichungen in sechs Unbekannten rechnen?
3. Wenn ja: Wie findet man fünf „überflüssige“ Gleichungen heraus?
4. Gibt es (auch bei sehr komplizierten Netzwerken) immer genau eine Lösung?
5. Falls es möglich ist, dass ein lineares Gleichungssystem keine Lösung oder mehrere Lösungen hat: Wie kann man das bereits *vor* dem Lösen erkennen?

6. Falls es mehrere Lösungen gibt: Welche Gestalt hat die Menge aller Lösungen eines derartigen Systems und auf welche Art kann man sie möglichst übersichtlich darstellen?

Punkt 1. lässt sich ohne lange Vorbereitungen beantworten, indem wir gleich unten den sogenannten Gauß-Algorithmus besprechen. Die Klärung der Punkte 2.-6. erfordert jedoch eine *Theorie linearer Gleichungssysteme* und keine – auch noch so ausführliche – Betrachtung einzelner Gleichungssysteme kann dies leisten. Kapitel 5 wird, mit Hilfe von Resultaten aus Kapiteln 3 und 4, Antworten auf die Fragen unter 2.-6. geben. In Kapitel 2 werden wir zunächst besprechen, wie mit den „Lösungstupeln“ $(I_1, \dots, I_6)$ und dem „Koeffizientenhaufen“ $\begin{pmatrix} -1 & -1 & \dots\dots & & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 300 & -200 & \dots\dots & & 500 & 0 \end{pmatrix}$ aus den obigen Gleichungen sachgemäß umzugehen ist.

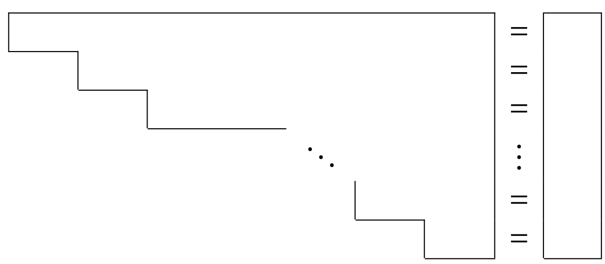
1.3. Das Eliminationsverfahren von Gauß ist ein Algorithmus zum Lösen linearer Gleichungssysteme

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

mit den n Unbekannten $x_1, \dots, x_n$ und den $m \cdot n$ Koeffizienten $a_{11}, \dots, a_{mn}$ aus $\mathbb{R}$ (oder aus $\mathbb{C}$ oder aus einem beliebigen Körper $\mathbb{K}$).

Wenn $b_i = 0$ für alle $i \in \{1, \dots, m\}$ gilt, dann sprechen wir von einem *homogenen* linearen Gleichungssystem – ein solches besitzt natürlich immer zumindest die Lösung $x_1 = \ldots = x_n = 0$. Wenn $b_i \neq 0$ für mindestens ein $i \in \{1, \dots, m\}$ gilt, dann handelt es sich um ein *inhomogenes* lineares Gleichungssystem.

Idee des Verfahrens: Wir versuchen, das Gleichungssystem durch passende (Äquivalenz-) Umformungen auf die Gestalt



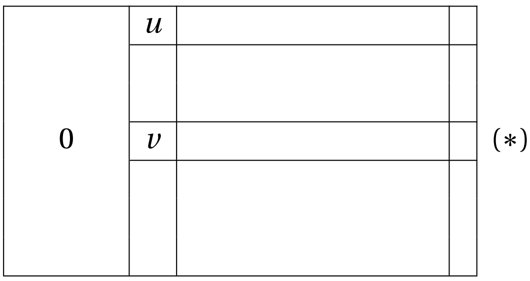
zu bringen und dann von unten nach oben aufzurollen („Vorwärtseliminieren“ und „Rückwärtssubstituieren“).

Für den Beginn des Verfahrens schreiben wir die Unbekannten nicht mehr zu den Koeffizienten dazu, dafür aber alles in Spalten nach den Nummern der Unbekannten sortiert und fein säuberlich untereinander, wobei wir „fehlende“ Glieder durch Nullen als Koeffizienten auffüllen. Im Fall des obigen Netzwerkbeispiels haben wir $m = 11$ und $n = 6$ und das Schema sieht so aus:

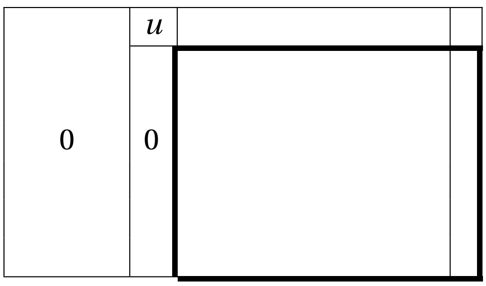
$$\begin{array}{cccccc|c}
-1 & -1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & -1 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 1 & -1 & -1 & 0 & 0 \\
300 & 0 & 0 & 1000 & -500 & -100 & 0 \\
0 & 200 & 0 & 1000 & -500 & 0 & 0 \\
0 & 0 & 100 & 0 & 500 & 100 & 3 \\
300 & -200 & 0 & 0 & 0 & -100 & 0 \\
300 & 0 & 100 & 1000 & 0 & 0 & 3 \\
0 & 200 & 100 & 1000 & 0 & 100 & 3 \\
300 & -200 & 100 & 0 & 500 & 0 & 3
\end{array}$$

Genauere Beschreibung des Verfahrens:

1.a. Suche links beginnend die erste Spalte, in der mindestens ein Koeffizient $\neq 0$ ist. Setze eine Gleichung (d.h. Zeile) an die erste Stelle, die in dieser Spalte einen Koeffizienten $u \neq 0$ hat. Dann ergibt sich folgende Situation:



1.b. Stelle nun unterhalb von u immer Nullen dadurch her, dass z.B. von Gleichung (∗) das $\frac{v}{u}$-fache der ersten Gleichung abgezogen wird. Dann ergibt sich folgendes Bild:



2. Betrachte das fett umrandete Teil-Gleichungssystem und verfahre dort wie bei 1.a und 1.b. Außerhalb des fetten Randes ändert sich dabei nichts mehr. Wiederhole diese Umformungen in den immer kleiner werdenden Teil-Systemen, bis das Endresultat dieses Prozesses folgendermaßen aussieht:

$$
\begin{array}{l|cccccc|c|l}
 & * * * * * & & & & & & & \\
\text{Zeile } 1 & 0 & c_{1j_1} * * * * * & & & & & d_1 & \\
\text{Zeile } 2 & & 0 & c_{2j_2} * * * * & & & & d_2 & \\
\vdots & & & 0 & & & & \vdots & \\
\vdots & & & & \ddots & & & \vdots & \text{alle } c_{i\,j_i} \neq 0 \\
\vdots & & & & & \ddots & & \vdots & \\
\text{Zeile } r & & & & & 0 & c_{r\,j_r} * * * * * & d_r & \\
\vdots & & & & & & 0 & d_{r+1} & \\
\vdots & & & & & & 0 & \vdots & \\
\text{Zeile } m & & & & & & 0 & d_m & \\
 & & \uparrow \text{Spalte } j_1 & \uparrow \text{Spalte } j_2 & \cdots & & \uparrow \text{Spalte } j_r & &
\end{array}
$$

3. Falls eines der $d_{r+1},\ldots,d_m$ ungleich 0 ist, so hat das Gleichungssystem keine Lösung. (In den Zeilen $r+1$ bis m stehen ja links nur Nullen!)

   Sind alle $d_{r+1},\ldots,d_m$ gleich 0, so ist das System lösbar.

4. Im letzteren Fall sind die Unbekannten entsprechend der Spalten mit Sternen $*$ frei wählbar. Dies ergibt $n-r$ viele „freie Parameter“, zum Beispiel $t_1,\ldots,t_{n-r}$.

5. Die restlichen r Unbekannten $x_{j_1},\ldots,x_{j_r}$ werden mit Zeile r beginnend von unten nach oben durch die $d_1,\ldots,d_r$ und die $t_1,\ldots,t_{n-r}$ ausgedrückt. Hier brauchen wir, dass alle $c_{i\,j_i} \neq 0$ sind.

An dieser Stelle empfehle ich, am besten gleich mal zur Gauß-Elimination in einem übersichtlichen Zahlenbeispiel in Punkt 1.4 weiter unten zu springen und sich bei Interesse danach auch an der konkreten Lösung für unser obiges Schema aus der Brückenschaltung zu versuchen.

„Außer Konkurrenz“ – Lösung des Netzwerkbeispiels: Übermütige können das eben beschriebene Verfahren natürlich sofort auf das obige Netzwerkbeispiel anwenden, aber ich empfehle eher, dies allenfalls erst nach dem nächsten kleineren Zahlenbeispiel wirklich zu tun. Im Schema des Endresultats hängen die konkreten Zahlen im rechten oberen Bereich übrigens von der jeweiligen Zeilenwahl ab, die immer in Schritt 1.a getroffen wird, aber die Grundstruktur ist für alle Variationen dieselbe (was aber erst mit ein bisschen mehr Theorie leicht einzusehen ist). Unterwegs ist es (vor allem bei händischer Rechnung) auch ratsam, gelegentlich bei einzelnen Gleichungen zur Vermeidung großer Zahlen durch gemeinsame Faktoren zu dividieren oder zur Vermeidung von Brüchen passend zu multiplizieren. In dieser Art können wir ausgehend von dem obigen Beginnschema diese Folge von Schemata produzieren:

−1	−1	0	1	0	0	0
0	−1	−1	1	0	1	0
0	1	0	0	1	−1	0
0	0	1	−1	−1	0	0
0	−300	0	1300	−500	−100	0
0	200	0	1000	−500	0	0
0	0	100	0	500	100	3
0	−500	0	300	0	−100	0
0	−300	100	1300	0	0	3
0	200	100	1000	0	100	3
0	−500	100	300	500	0	3

−1	−1	0	1	0	0	0
0	−1	−1	1	0	1	0
0	0	−1	1	1	0	0
0	0	1	−1	−1	0	0
0	0	300	1000	−500	−400	0
0	0	−200	1200	−500	200	0
0	0	100	0	500	100	3
0	0	500	−200	0	−600	0
0	0	400	1000	0	−300	3
0	0	−100	1200	0	300	3
0	0	600	−200	500	−500	3

Bevor wir weiterrechnen, dividieren wir hier die Gleichungen der Zeilen 5, 6 und 8 allesamt durch 10 und kommen nach Nullenproduktion in der dritten Spalte auf

−1	−1	0	1	0	0	0
0	−1	−1	1	0	1	0
0	0	−1	1	1	0	0
0	0	0	0	0	0	0
0	0	0	13	−2	−4	0
0	0	0	10	−7	2	0
0	0	0	100	600	100	3
0	0	0	3	5	−6	0
0	0	0	1400	400	−300	3
0	0	0	1100	−100	300	3
0	0	0	400	1100	−500	3

−1	−1	0	1	0	0	0
0	−1	−1	1	0	1	0
0	0	−1	1	1	0	0
0	0	0	3	5	−6	0
0	0	0	13	−2	−4	0
0	0	0	10	−7	2	0
0	0	0	100	600	100	3
0	0	0	1400	400	−300	3
0	0	0	1100	−100	300	3
0	0	0	400	1100	−500	3
0	0	0	0	0	0	0

Wir haben rechts einfach nur die 8. Zeile mit dem führenden 3er hinauf geholt und die 4. Zeile (mit lauter Nullen) gleich ganz nach unten geschoben. Die nächstkleineren Teilsysteme sehen nach „kosmetischen Operationen zur Erhaltung der Ganzzahligkeit“ (vorher Multiplikation mit 3, dann nach Reduktion wieder Division durch 3) der betroffenen Zeilen dann so aus:

−1	−1	0	1	0	0	0
0	−1	−1	1	0	1	0
0	0	−1	1	1	0	0
0	0	0	3	5	−6	0
0	0	0	0	−71	66	0
0	0	0	0	−71	66	0
0	0	0	0	1300	900	9
0	0	0	0	−5800	7500	9
0	0	0	0	−5800	7500	9
0	0	0	0	1300	900	9
0	0	0	0	0	0	0

−1	−1	0	1	0	0	0
0	−1	−1	1	0	1	0
0	0	−1	1	1	0	0
0	0	0	3	5	−6	0
0	0	0	0	−71	66	0
0	0	0	0	0	0	0
0	0	0	0	0	49900	213
0	0	0	0	0	49900	213
0	0	0	0	0	49900	213
0	0	0	0	0	49900	213
0	0	0	0	0	0	0

Und unser Verfahren endet nun schließlich mit folgendem Schema:

−1	−1	0	1	0	0	0
0	−1	−1	1	0	1	0
0	0	−1	1	1	0	0
0	0	0	3	5	−6	0
0	0	0	0	−71	66	0
0	0	0	0	0	49900	213
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

Es ist also $r = 6$ und wir sehen unmittelbar, dass das Gleichungssystem lösbar ist, weil in den Zeilen mit Nummern 7-11 auch rechts nur Nullen stehen. Weiters ist hier $r = 6 = n$, also haben wir keine freien Parameter und die Unbekannten ergeben sich sukzessive aus den

Gleichungen in Zeile 6, dann Zeile 5, … und schließlich Zeile 1 wie folgt:

$$
\begin{aligned}
I_6 &= \frac{213}{49900} && \text{(entspricht} \approx 4{,}27\ \text{mA),}\\
I_5 &= \frac{66}{71} I_6 = \frac{99}{24950} && \text{(entspricht} \approx 3{,}97\ \text{mA),}\\
I_4 &= 2I_6 - \frac{5}{3} I_5 = \frac{24}{12475} && \text{(entspricht} \approx 1{,}92\ \text{mA),}\\
I_3 &= I_4 + I_5 = \frac{147}{24950} && \text{(entspricht} \approx 5{,}89\ \text{mA),}\\
I_2 &= -I_3 + I_4 = \frac{3}{9980} && \text{(entspricht} \approx 0{,}30\ \text{mA),}\\
I_1 &= -I_2 + I_4 = \frac{81}{49900} && \text{(entspricht} \approx 1{,}62\ \text{mA).}
\end{aligned}
$$

1.4. Ein Zahlenbeispiel mit 5 Gleichungen in den 7 Unbekannten $x_1, \dots, x_7$ (d.h. $m = 5$ und $n = 7$):

$$
\begin{array}{rrrrrrl}
x_2 + & & 2x_4 & - \ x_5 & - \ 4x_6 & & = \ 4\\
& x_3 - & x_4 & - \ x_5 & + \ 2x_6 & + x_7 & = -4\\
x_2 + & & 2x_4 & + \ x_5 & - \ 2x_6 & & = \ 4\\
& x_3 - & x_4 & + & 2x_6 & - x_7 & = -2\\
x_2 + & x_3 + & x_4 & + & & x_7 & = \ 0
\end{array}
$$

Zunächst machen wir daraus das Anfangsschema und bemerken, dass die Konstellation für Schritt 1.a bereits mit Spalte 2 erfüllt ist:

$$
\begin{array}{ccccccc|c}
0 & 1 & 0 & 2 & -1 & -4 & 0 & 4\\
0 & 0 & 1 & -1 & -1 & 2 & 1 & -4\\
0 & 1 & 0 & 2 & 1 & -2 & 0 & 4\\
0 & 0 & 1 & -1 & 0 & 2 & -1 & -2\\
0 & 1 & 1 & 1 & 0 & 0 & 1 & 0
\end{array}
$$

Nach Anwendung von Schritt 1.b kommen wir auf folgendes Schema mit dem entsprechend umrandeten Teil-Gleichungssystem bestehend aus 4 Gleichungen für 5 Unbekannte:

$$
\begin{array}{cc|ccccc|c|}
0 & 1 & 0 & 2 & -1 & -4 & 0 & 4\\
\cline{3-8}
0 & 0 & 1 & -1 & -1 & 2 & 1 & -4\\
0 & 0 & 0 & 0 & 2 & 2 & 0 & 0\\
0 & 0 & 1 & -1 & 0 & 2 & -1 & -2\\
0 & 0 & 1 & -1 & 1 & 4 & 1 & -4\\
\cline{3-8}
\end{array}
$$

Gemäß Verfahrensschritt 2 wenden wir nun 1.a und 1.b auf das Teilsystem an und kommen auf

$$
\begin{array}{ccc|cccc|c|}
0 & 1 & 0 & 2 & -1 & -4 & 0 & 4\\
0 & 0 & 1 & -1 & -1 & 2 & 1 & -4\\
\cline{4-8}
0 & 0 & 0 & 0 & 2 & 2 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0 & -2 & 2\\
0 & 0 & 0 & 0 & 2 & 2 & 0 & 0\\
\cline{4-8}
\end{array}
$$

und dann weiter auf

$$
\begin{array}{ccccccc|c}
0 & 1 & 0 & 2 & -1 & -4 & 0 & 4 \\
0 & 0 & 1 & -1 & -1 & 2 & 1 & -4 \\
0 & 0 & 0 & 0 & 2 & 2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & -2 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}
$$

und haben somit ohne weitere Rechnung als Endschema

$$
\begin{array}{ccccccc|c}
0 & 1 & 0 & 2 & -1 & -4 & 0 & 4 \\
0 & 0 & 1 & -1 & -1 & 2 & 1 & -4 \\
0 & 0 & 0 & 0 & 2 & 2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & -2 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}
$$

mit $r = 4$ und $j_1 = 2$, $j_2 = 3$, $j_3 = 5$ und $j_4 = 6$.

Im 3. Schritt stellen wir mit Blick auf Zeile 5 fest, dass das Gleichungssystem lösbar ist.

Im 4. Schritt stellen wir die $n - r = 7 - 4 = 3$ freien Parameter t_1, t_2, t_3 auf und setzen

$$x_1 = t_1, \quad x_4 = t_2 \quad \text{und} \quad x_7 = t_3.$$

(Erinnere: hier kommen jene x_i dran, für die i verschieden von j_1, j_2 und j_3 ist!)

Im 5. Schritt kommt das „Aufrollen von unten“:

4. Gleichung: $-x_6 - 2x_7 = 2 \quad \Rightarrow \quad x_6 = -2 - 2t_3$.

3. Gleichung: $2x_5 + 2x_6 = 0 \quad \Rightarrow \quad x_5 = -x_6 = 2 + 2t_3$.

2. Gleichung: $x_3 - x_4 - x_5 + 2x_6 + x_7 = -4 \quad \Rightarrow \quad x_3 = -4 + t_2 + 2 + 2t_3 + 4 + 4t_3 - t_3 = 2 + t_2 + 5t_3$.

1. Gleichung: $x_2 + 2x_4 - x_5 - 4x_6 = 4 \quad \Rightarrow \quad x_2 = 4 - 2t_2 + 2 + 2t_3 - 8 - 8t_3 = -2 - 2t_2 - 6t_3$.

Zusammenfassend können wir die Lösungen so darstellen: Es gilt mit t_1, t_2, t_3 beliebig (aus $\mathbb{R}$)

$$
\begin{array}{rrrrr}
x_1 = & & t_1 & & \\
x_2 = & -2 & & -2t_2 & -6t_3 \\
x_3 = & 2 & & +t_2 & +5t_3 \\
x_4 = & & & t_2 & \\
x_5 = & 2 & & & 2t_3 \\
x_6 = & -2 & & & -2t_3 \\
x_7 = & & & & t_3
\end{array}
$$

beziehungsweise in Tupelschreibweise (d.h. Vektoren im $\mathbb{R}^7$)

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 2 \\ 0 \\ 2 \\ -2 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_3 \begin{pmatrix} 0 \\ -6 \\ 5 \\ 0 \\ 2 \\ -2 \\ 1 \end{pmatrix}$$

1.5. Diskussion des Verfahrens:

1. Betrachtet man die vorgeschriebenen Schritte sorgfältig, so erkennt man, dass das Verfahren „immer funktioniert", d.h. es ist auf jedes lineare Gleichungssystem anwendbar und liefert immer ein Ergebnis.

2. Eigentlich wird in den Schritten 4 und 5 nicht das *gegebene* System gelöst, sondern das *umgeformte*. Wir können uns aber leicht davon überzeugen, dass im Laufe des Umformungsprozesses weder Lösungen verloren gegangen noch welche dazugekommen sind: Jeder Schritt war eine „Äquivalenzumformung", d.h. ist ein 7-Tupel $(x_1, \ldots, x_7)$ *vorher* eine Lösung des Systems, dann ist es auch *nachher* eine; und ist es *nachher* Lösung, dann war es schon *vorher* eine (die Umformungsschritte waren Tausch von Gleichungen, Subtraktion eines Vielfachen einer Gleichung von einer anderen Gleichung).

3. Das Verfahren beinhaltet im Schritt 1.a (und damit auch in der Folge in den Schritten 2) gegebenenfalls die Wahlfreiheit, welche von mehreren in Frage kommenden Zeilen man an die erste Stelle setzt. Das Verfahren ist daher nicht vollständig deterministisch, und verschiedene Auswahlen dieser Zeilen könnten unterschiedlich aussehende Ergebnisse liefern. Aber die sollten dann doch „in Wirklichkeit" gleich sein!? [Hier kann uns nur die Theorie weiterhelfen.] Dasselbe gilt für einen „Platztausch" der Unbekannten, der doch die Lösung eigentlich nicht verändern sollte!

1.6. Beispiel:

$$\begin{array}{llll} x_1 & +x_2 & +x_3 & = 5 \\ x_1 & -x_2 & & = -1 \end{array}$$

(Schnitt zweier Ebenen im Raum)

$$\begin{array}{rrr|r} 1 & 1 & 1 & 5 \\ 1 & -1 & 0 & -1 \end{array} \quad \text{ergibt} \quad \begin{array}{rrr|r} 1 & 1 & 1 & 5 \\ 0 & -2 & -1 & -6 \end{array} \quad \text{und somit} \quad \begin{array}{rcll} x_3 & = & & t \\ x_2 & = & 3 & -t/2 \\ x_1 & = & 2 & -t/2 \end{array} \quad (t \in \mathbb{R})$$

oder in Vektorschreibweise $$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} + t \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} \qquad (1).$$

Noch weniger Rechenarbeit haben wir aber so:
$$\begin{aligned} x_3 \;\; +x_1 \;\; +x_2 &= 5 \\ x_1 \;\; -x_2 &= -1 \end{aligned}$$

$$\left.\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & 1 & -1 & -1 \end{array}\right. \quad \text{und somit} \quad \begin{aligned} x_2 &= \;\; u \\ x_1 &= -1 \;\; +u \qquad (u \in \mathbb{R}) \\ x_3 &= 6 \;\; -2u \end{aligned}$$

bzw. $$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 6 \end{pmatrix} + u \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \qquad (2).$$

Die Lösungsformeln (1) und (2) sollten doch „gleich“ sein!?
Sie sind es natürlich: Wähle in (1) einen bestimmten freien Parameter t, dann bekomme ich in (2) für $u = 3 - t/2$ dieselbe Lösung; wähle ich umgekehrt in (2) ein bestimmtes u, dann liefert mir (1) für $t = 6 - 2u$ dieselbe Lösung.

Wie steht es aber in unübersichtlicheren Fällen mit vielen Gleichungen, vielen Unbekannten und vielen freien Parametern? Hatten wir oben nur Glück oder geht es immer gut? Welche Freiheiten hat man im Allgemeinen beim Hinschreiben der Lösung?

1.7. Ausblick auf die nächsten Kapitel, ausgehend von linearen Gleichungssystemen:

Kurzschreibweise für $(x_1, \ldots, x_n)$: Vektoren im $\mathbb{R}^n$, Vektorräume	Kapitel 2, 3
Kurzschreibweise für $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$: Matrizen, lineare Abbildungen	Kapitel 2, 4
Kurzschreibweise $A \cdot x = b$ für lineare Gleichungssysteme: Matrixmultiplikation	Kapitel 2
Lösungsmenge eines linearen Gleichungssystems: Teilraum, affiner Teilraum	Kapitel 3
Überflüssige Gleichungen: Lineare Abhängigkeit	Kapitel 3
Verschiedene Darstellungen der Lösungsmenge: Erzeugendensystem, Basis	Kapitel 3
Anzahl der freien Parameter der Lösung: Dimension, Rang	Kapitel 3, 4
Anwendung der allgemeinen Theorie auf lineare Gleichungssysteme	Kapitel 5

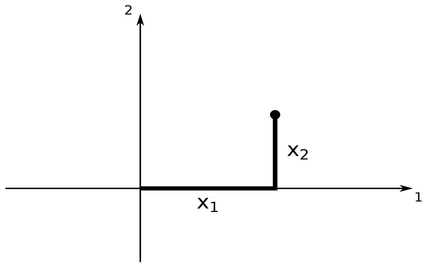
§2. VEKTOREN IM $\mathbb{R}^n$, MATRIZEN

2.1. Der Raum $\mathbb{R}^n$: Wir setzen $\mathbb{R}^n := \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ mal}} = \{\underbrace{(x_1, \ldots, x_n)}_{n\text{-Tupel}} \mid x_1, \ldots, x_n \in \mathbb{R}\}$ und bezeichnen jedes Element $(x_1, \ldots, x_n) \in \mathbb{R}^n$ als *Vektor im* $\mathbb{R}^n$. Hier ist $(x_1, \ldots, x_n)$ die Zeilenform und $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ die Spaltenform des Vektors – erstere Schreibweise ist zwar in Texten platzsparend, aber letztere ist im Zusammenhang mit Matrizenmultiplikation (s.u.) der Standard. Wir treffen die Vereinbarung der Kurzschreibweisen $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \mathbb{R}^n$ und somit

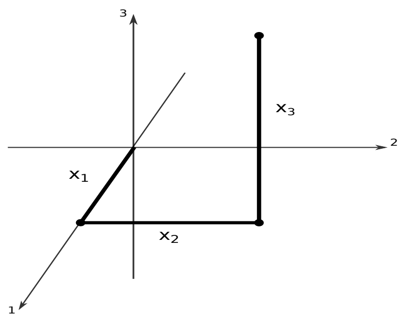
$$x = y \quad \Longleftrightarrow \quad x_i = y_i \quad (i = 1, \ldots, n).$$

2.2. Die Spezialfälle $n = 1, 2, 3$ und $n = 4$: $\mathbb{R}^1$ ist praktisch dasselbe wie $\mathbb{R}$, die Menge der reellen Zahlen.

$\mathbb{R}^2 = \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R}\}$ lässt sich mittels kartesischem Koordinatensystem als Ebene veranschaulichen:



$\mathbb{R}^3 = \{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1, x_2, x_3 \in \mathbb{R}\}$ ist in ähnlicher Art als 3-dimensionaler Raum auffassbar:



$\mathbb{R}^4 = \{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mid x_1, x_2, x_3, x_4 \in \mathbb{R}\}$ wird in der Relativitätstheorie auch als $\{\begin{pmatrix} t \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid t, x_1, x_2, x_3 \in \mathbb{R}\}$ oder $\{(t, x) \mid t \in \mathbb{R}, x \in \mathbb{R}^3\}$ interpretiert.

2.3. Rechenoperationen:

Grundlegend: 1. Addition: $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$

2. Vervielfachung: $\lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$ $(\lambda \in \mathbb{R})$; Multiplikation mit „Skalaren“

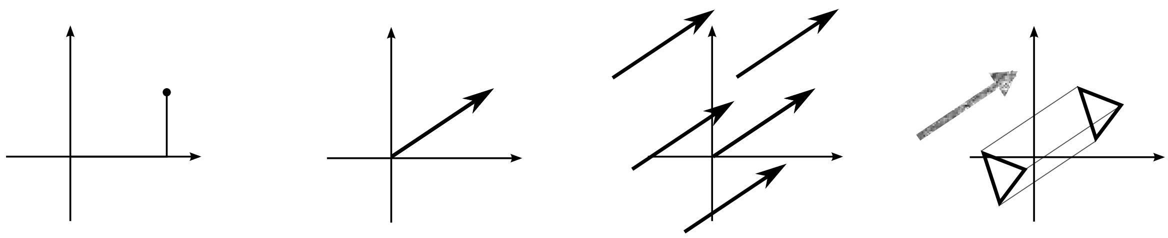
Weiters [vgl. Kapitel 7]:

3. Skalarprodukt: $\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle := x_1 y_1 + \ldots + x_n y_n$ $[\in \mathbb{R}!]$

Nur im $\mathbb{R}^3$ [siehe Anhang A.1]:

4. Vektorprodukt (äußeres Produkt): $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$

Zur Veranschaulichung von 1. und 2. im $\mathbb{R}^2$ bzw. im $\mathbb{R}^3$ müssen wir diese Räume in flexiblerer Weise geometrisch interpretieren als bisher bloß durch Punkte und Koordinaten. Zum Beispiel für $\begin{pmatrix} 3 \\ 2 \end{pmatrix} \in \mathbb{R}^2$:



Punkt

Ortsvektor

Menge aller ,,solchen'' Pfeile (d.h. gleichlang, gleiche Richtung, gleiche Orientierung)

Verschiebung (Translation) (3 nach rechts, 2 hinauf)

2.4. Matrizen

Definition: Eine *$m \times n$-Matrix über* $\mathbb{R}$ $(m, n \in \mathbb{N})$ ist ein rechteckiges Schema $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ reeller Zahlen a_{ij}, genannt die *Koeffizienten (oder Komponenten) von A*, der Form

$$A := \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_{\substack{n\text{ Spalten,} \\ \text{2. Index = Spaltenindex}}} \left.\vphantom{\begin{pmatrix} a \\ a \\ a \end{pmatrix}}\right\} \substack{m\text{ Zeilen,} \\ \text{1. Index = Zeilenindex,}}$$

Kurznotationen $(A)_{ij} = a_{ij}$ und $A = (a_{ij})$.

Analog sind Matrizen über einem anderen Körper $\mathbb{K}$ definiert: Koeffizienten $a_{ij} \in \mathbb{K}$; $A = B \iff A$ und B haben gleiches Format und $a_{ij} = b_{ij}$ für alle i, j. Die Menge aller $m \times n$-Matrizen mit Koeffizienten aus $\mathbb{K}$ bezeichnen wir mit $M(m, n; \mathbb{K})$, im Falle $\mathbb{K} = \mathbb{R}$ schreiben wir $M(m, n)$ kurz für $M(m, n; \mathbb{R})$. Spezialfälle sind $M(m, 1) = \mathbb{R}^m$ (als Spaltenvektoren), $M(1, n) = \mathbb{R}^n$ (als Zeilenvektoren) und $M(1, 1) = \mathbb{R}$.

2.5. Rechenoperationen mit Matrizen

1. Addition: Für Matrizen mit gleichem Format (und nur für solche) ist diese komponentenweise definiert

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} := \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

2. Vervielfachung: $$\lambda \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} := \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix} \qquad (\lambda \in \mathbb{R})$$

3. Multiplikation $A \cdot B$: Zunächst der Versuch einer zweistufigen Motivation.

A) Für $A \in M(m, n)$ und einen Spaltenvektor $x \in \mathbb{R}^n = M(n, 1)$

soll $$A \cdot x = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{gerade} \quad \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} \in \mathbb{R}^m$$ ergeben,

d.h. für $1 \le i \le m$ soll die i-te Komponente von $A \cdot x$ gerade das Skalarprodukt der i-ten Zeile von A mit x sein, also gleich $s_i := a_{i1}x_1 + \ldots + a_{in}x_n$; schematisch sieht das so aus:

$$\begin{pmatrix} \vdots \\ \boxed{a_{i1} \cdots\cdots a_{in}} \\ \vdots \\ \vdots \end{pmatrix} \cdot \begin{pmatrix} \boxed{\begin{matrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{matrix}} \end{pmatrix} = \begin{pmatrix} \vdots \\ \boxed{s_i} \\ \vdots \\ \vdots \end{pmatrix}.$$

B) Steht anstelle von $x \in M(n, 1)$ nun eine $n \times p$-Matrix $B \in M(n, p)$, so zerlegen wir diese in p Spalten, behandeln jede Spalte einzeln ($k = 1, \ldots, p$) wie in A) den Vektor x und schreiben die Ergebnisvektoren nebeneinander:

$$A \cdot B = \begin{pmatrix} \vdots \\ \boxed{a_{i1} \cdots\cdots a_{in}} \\ \vdots \\ \vdots \end{pmatrix} \cdot \begin{pmatrix} \cdots\cdots & \boxed{\begin{matrix} b_{1k} \\ \vdots \\ \vdots \\ b_{nk} \end{matrix}} & \cdots \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots\cdots & \boxed{s_{ik}} & \cdots \\ & \vdots & \\ & \vdots & \end{pmatrix}, \quad s_{ik} := a_{i1}b_{1k} + \ldots + a_{in}b_{nk}.$$

Zusammenfassend kommen wir so zur „amtlichen“

Definition: Für $A \in M(m,n)$ und $B \in M(n,p)$ [beachte die Formatbedingung: die Spaltenanzahl von A muss also gleich der Zeilenanzahl von B sein!] definieren wir das Matrixprodukt $A \cdot B \in M(m,p)$ durch

$$(A\cdot B)_{ik} := a_{i1}b_{1k} + \ldots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk} \qquad (i=1,\ldots,m;\ k=1,\ldots,p).$$

Bemerkung: $A \cdot A = A^2$, $A^3, \ldots$ ist also sinnvoll nur für $A \in M(n,n)$.

Beispiele: 1. $\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 2 & -2 \\ 2 & -1 & 0 & -3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 \\ 0 & -1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -5 \\ 0 & 1 \\ -1 & -3 \end{pmatrix} \qquad (m=3, n=4, p=2)$

2. $\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + \ldots + a_nb_n \in \mathbb{R} = M(1,1)$

3. $\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \cdot \begin{pmatrix} b_1 & \cdots & b_p \end{pmatrix} = \begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_p \\ \vdots & \vdots & & \vdots \\ a_mb_1 & a_mb_2 & \cdots & a_mb_p \end{pmatrix} \in M(m,p)$

2.6. Rechenregeln

a) für Addition und Vervielfachung von Matrizen: Wie für Vektoren im $\mathbb{R}^n$; siehe Kapitel 3;

Nullmatrix $\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} = 0 \in M(m,n)$

b) für Matrixmultiplikation:

1. Assoziativität $(A\cdot B)\cdot C = A\cdot(B\cdot C) \quad (A \in M(m,n), B \in M(n,p), C \in M(p,q))$
2. Einheitsmatrix $I_m \cdot A = A = A \cdot I_n \quad (A \in M(m,n))$, wobei

$$I_m := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in M(m,m); \quad I_m = (\delta_{ij}),\ \delta_{ij} := \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \quad \textit{Kronecker-Delta.}$$

3. Im Allgemeinen gilt NICHT $A\cdot B = B\cdot A$ [Matrixmultiplikation ist NICHT kommutativ].

c) für Matrixmultiplikation und Addition:

4. $(A+B)\cdot C = A\cdot C + B\cdot C \quad (A,B\in M(m,n), C\in M(n,p))$

5. $A\cdot(B+C) = A\cdot B + A\cdot C \quad (A\in M(m,n), B,C\in M(n,p))$

d) für Matrixmultiplikation und Vervielfachung:

6. $(\lambda A)\cdot B = \lambda(A\cdot B) \quad (A\in M(m,n), B\in M(n,p), \lambda\in\mathbb{R})$

7. $A\cdot(\lambda B) = \lambda(A\cdot B)$ ——— - " ——— -

Beweise:

ad 1. Für alle $i = 1,\dots,m$ und $l = 1,\dots,q$ gilt jeweils

$$\left.\begin{array}{l} \big((A\cdot B)\cdot C\big)_{il} = \sum\limits_{k=1}^{p} \big(A\cdot B\big)_{ik} c_{kl} = \sum\limits_{k=1}^{p} \big(\sum\limits_{j=1}^{n} a_{ij}b_{jk}\big)c_{kl} = \sum\limits_{k=1}^{p}\sum\limits_{j=1}^{n} a_{ij}b_{jk}c_{kl} \\ \big(A\cdot(B\cdot C)\big)_{il} = \sum\limits_{j=1}^{n} a_{ij}\big(B\cdot C\big)_{jl} = \sum\limits_{j=1}^{n} a_{ij}\sum\limits_{k=1}^{p} b_{jk}c_{kl} = \sum\limits_{j=1}^{n}\sum\limits_{k=1}^{p} a_{ij}b_{jk}c_{kl} \end{array}\right\} =$$

und somit $(A\cdot B)\cdot C = A\cdot(B\cdot C)$.

ad 2. $$\begin{pmatrix} \boxed{1\,0\cdots\cdots 0} \\ \vdots \\ \boxed{0\cdots\cdots 0\,1} \end{pmatrix}\cdot\begin{pmatrix} \boxed{\begin{matrix} a_{11} \\ \vdots \\ a_{m1}\end{matrix}} & \cdots & \boxed{\begin{matrix} a_{1n} \\ \vdots \\ a_{mn}\end{matrix}} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn}\end{pmatrix} \text{ und}$$

$$\begin{pmatrix} \boxed{a_{11}\ \cdots\cdots\ a_{1n}} \\ \vdots \\ \boxed{a_{m1}\ \cdots\cdots\ a_{mn}} \end{pmatrix}\cdot\begin{pmatrix} \boxed{\begin{matrix} 1 \\ 0 \\ \vdots \\ 0\end{matrix}} & \cdots & \boxed{\begin{matrix} 0 \\ \vdots \\ 0 \\ 1\end{matrix}} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn}\end{pmatrix};$$

alternativ $\big(I_m\cdot A\big)_{ik} = \sum\limits_{j=1}^{m}\delta_{ij}a_{jk} = \delta_{ii}a_{ik} = a_{ik}$ für alle i und k, daher $I_m\cdot A = A$; analog $A\cdot I_n = A$ (Übung).

ad 3. Hier ist *kein* „allgemeiner Beweis" mit Buchstaben a_{ij} etc. gefragt, sondern ein konkretes Zahlenbeispiel für Matrizen A und B, bei denen sich $A\cdot B \neq B\cdot A$ ergibt! Denn:

„$A\cdot B = B\cdot A$ gilt allgemein" heißt $\forall A\forall B:\ A\cdot B = B\cdot A$.

„$A\cdot B = B\cdot A$ gilt NICHT allgemein" heißt $\neg\Big(\forall A\forall B:\ A\cdot B = B\cdot A\Big)$,

das heißt $\exists A\exists B:\ A\cdot B \neq B\cdot A$.

Konkret sehen wir zum Beispiel für die 2×2-Matrizen $A = \begin{pmatrix} 0 & 1 \\ 0 & 0\end{pmatrix}$ und $B = \begin{pmatrix} 0 & 0 \\ 1 & 0\end{pmatrix}$, dass

$$A\cdot B = \begin{pmatrix} 0 & 1 \\ 0 & 0\end{pmatrix}\cdot\begin{pmatrix} 0 & 0 \\ 1 & 0\end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0\end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1\end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0\end{pmatrix}\cdot\begin{pmatrix} 0 & 1 \\ 0 & 0\end{pmatrix} = B\cdot A.$$

(Natürlich ist $A\cdot B = B\cdot A$ nicht für alle A, B falsch! Z.B. gilt ja $A\cdot I = A = I\cdot A$ oder $A\cdot 0 = 0 = 0\cdot A$.)

ad 4. $\big((A+B)\cdot C\big)_{ik} = \sum\limits_j \big(A+B\big)_{ij}\, c_{jk} = \sum\limits_j (a_{ij} + b_{ij})\, c_{jk} = \sum\limits_j (a_{ij}\, c_{jk} + b_{ij}\, c_{jk}) =$

$$= \sum_j a_{ij}\, c_{jk} + \sum_j b_{ij}\, c_{jk} = \big(A\cdot C\big)_{ik} + \big(B\cdot C\big)_{ik} = \big(A\cdot C + B\cdot C\big)_{ik}$$

und, weil dies für alle i und k gilt, folgt die Behauptung. ($\sum\limits_j$ heißt natürlich ungefähr: „Wir summieren über j und ich bin zu faul, hier immer die[selben] Grenzen dazuzuschreiben".)

ad 5. analog zu 4. (Übung)

ad 6. $\forall\, i,k$: $\big((\lambda A)\cdot B\big)_{ik} = \sum\limits_j (\lambda a_{ij}) b_{jk} = \sum\limits_j \lambda\,(a_{ij}\, b_{jk}) = \lambda \sum\limits_j a_{ij} b_{jk} = \lambda\big(A\cdot B\big)_{ik} = \big(\lambda(A\cdot B)\big)_{ik}$

ad 7. analog zu 6. (Übung)

Bemerkungen: 1. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, also: $A\cdot B = 0 \quad \not\Longrightarrow \quad A = 0$ oder $B = 0$.

D.h.: Es gibt „Nullteiler" (Elemente $\neq 0$, deren Produkt 0 ist) unter den Matrizen.

2. $(M(n,n;\mathbb{K}),+,\cdot)$ bildet einen *Ring*. (Aber keinen Körper für $n \geq 2$, denn ein Körper hat keine Nullteiler!)

§3. VEKTORRÄUME

3.1. Die grundlegenden Rechengesetze: Vektoren im $\mathbb{R}^n$ erfüllen zusammen mit den Skalaren in $\mathbb{R}$ offensichtlich die folgenden Rechengesetze (V1)-(V8), die wir aber nun allgemeiner für eine Menge V statt $\mathbb{R}^n$ und für einen Skalarenkörper $\mathbb{K}$ statt $\mathbb{R}$ aufschreiben.

Definition: Sei $\mathbb{K}$ ein Körper. Ein *Vektorraum* über $\mathbb{K}$ (oder $\mathbb{K}$-Vektorraum; abgekürzt: $\mathbb{K}$-VR) ist eine Menge V zusammen mit den Operationen

$$\begin{aligned} +\colon V \times V \to V, &\qquad (v,w) \mapsto v+w, \\ \cdot\colon \mathbb{K} \times V \to V, &\qquad (\lambda, v) \mapsto \lambda v, \end{aligned}$$

sodass für alle Elemente $u, v, w \in V$ und $\lambda, \mu \in \mathbb{K}$ die folgenden Eigenschaften erfüllt sind:

(V1) $(u+v)+w = u+(v+w)$.

(V2) Es gibt ein Element $0 \in V$ [„Nullvektor“] mit $0+v = v+0 = v$.

(V3) Zu v existiert ein Element $-v \in V$ mit $v+(-v) = 0$.

(V4) $v+w = w+v$.

(V5) $\lambda(v+w) = \lambda v + \lambda w$.

(V6) $(\lambda+\mu)v = \lambda v + \mu v$.

(V7) $(\lambda\mu)v = \lambda(\mu v)$.

(V8) $1v = v$.

Aus Regel (V2) folgt insbesondere, dass $V \neq \emptyset$ gilt (weil V eben zumindest das Element 0 enthalten muss). Die Regeln (V1)-(V4) besagen gerade, dass V bzgl. „+“ eine kommutative Gruppe ist. Im $\mathbb{R}^n$ gelten die Eigenschaften (V1)-(V8) offensichtlich mit $0 = (0,\dots,0)$ und $-x = (-x_1,\dots,-x_n)$ (additiv) invers zu $x = (x_1,\dots,x_n)$.

Ein „Vektor“ ist also bloß ein „Element eines Vektorraumes“!

3.2. Beispiele: 1. Mit der üblichen komponentenweisen Addition und Vervielfachung ist $\mathbb{R}^n$ ein VR über $\mathbb{R}$, $\mathbb{C}^n$ einer über $\mathbb{C}$; allgemeiner ist für einen Körper $\mathbb{K}$ stets $\mathbb{K}^n$ ein $\mathbb{K}$-VR.

2. $M(m,n;\mathbb{K})$ ist mit den Matrizenoperationen aus 1. und 2. in 2.5 ein $\mathbb{K}$-VR.

Weitere Beispiele für $\mathbb{R}$-VR sind:

3. Die Menge der reellen Polynome $P(\mathbb{R}) := \{a_0 + a_1 x + \ldots + a_k x^k \mid \underbrace{k \in \mathbb{N}}_{\text{beliebig}}, a_i \in \mathbb{R}\ (i = 1, \ldots, k)\}$

und für $n \in \mathbb{N}$ fix die Menge $P_n(\mathbb{R}) := \{a_0 + a_1 x + \ldots + a_n x^n \mid a_i \in \mathbb{R}\ (i = 1, \ldots, n)\}$ der reellen Polynome vom Grad kleiner gleich n mit den bekannten Operationen.

4. Mit komponentenweisen Operationen analog zu $\mathbb{R}^n$:
   $c_0 := \{(x_1, x_2, \ldots) \mid x_n \in \mathbb{R}\,(n \in \mathbb{N}),\ \lim_{n\to\infty} x_n = 0\}$ Raum der Nullfolgen,
   $l^\infty := \{(x_1, x_2, \ldots) \mid x_n \in \mathbb{R}\,(n \in \mathbb{N}),\ (x_n) \text{ beschränkt}\}$ Raum der beschränkten Folgen,
   $\mathbb{R}^{\mathbb{N}} := \{(x_1, x_2, \ldots) \mid x_n \in \mathbb{R}\,(n \in \mathbb{N})\}$ Raum aller Folgen.

5. Mit den punktweisen Operationen $(f+g)(x) := f(x) + g(x)$ und $(\lambda f)(x) := \lambda \cdot f(x)$:
   $F[a,b] := \{f\colon [a,b] \to \mathbb{R}\}$,
   $C[a,b] := \{f\colon [a,b] \to \mathbb{R} \mid f \text{ stetig}\}$,
   $D[a,b] := \{f\colon [a,b] \to \mathbb{R} \mid f \text{ differenzierbar}\}$,
   $\{\underbrace{\lambda \sin(\omega t) + \mu \cos(\omega t)}_{f(t)} \mid \lambda, \mu \in \mathbb{R}\}$.

   In der Signalverarbeitung und für die Fourier-Analyse ist speziell der folgende Raum der „trigonometrischen Polynome“ von Bedeutung
   $TP(\mathbb{R}) := \{\frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) + \sum_{k=1}^{n} b_k \sin(kx) \mid \underbrace{n \in \mathbb{N}}_{\text{beliebig}} \text{ und } a_k, b_k \in \mathbb{R}\ (k = 1, \ldots, n)\}$.

Bei allen obigen Beispielen haben wir ordnungsgemäß nicht bloß die Menge der „Vektoren“ hingeschrieben, sondern auch die Operationen angegeben. An und für sich müssten wir für jedes der Beispiele beweisen:

1) dass die angegebenen Operationen wieder ein Element dieser Menge liefern;

und

2) dass (V1)-(V8) erfüllt sind.

1) ist in den meisten Fällen offensichtlich, außer:

Summe und Vielfache von Nullfolgen sind Nullfolgen: in der Analysis-VO wird gezeigt
$\lim x_n = 0 \text{ und } \lim y_n = 0 \implies \lim(x_n + y_n) = 0.$

Summe und Vielfache von beschränkten Folgen sind beschränkte Folgen: aus Analysis-VO
$|x_n| \le C_1 \text{ und } |y_n| \le C_2 \implies |x_n + y_n| \le C_1 + C_2.$

Summe und Vielfache von stetigen [bzw. differenzierbaren] Funktionen sind stetig [bzw. differenzierbar]: folgt aus entsprechenden Sätzen in der Analysis-VO.

2) ist schrecklich langweilig … Nullelement ist jeweils Nullmatrix, das Nullpolynom, die Folge $(0, 0, \ldots)$, die Funktion identisch 0, … additiv inverses Element ist jeweils $-A = (-a_{ij})$, $(-x_n)$, $(-f)(x) := -f(x)$, …

3.3. Proposition: Ist V ein $\mathbb{K}$-VR, so gelten für alle $v \in V$ und $\lambda \in \mathbb{K}$ die Aussagen:

(i) $0v = 0$ [links $0 \in \mathbb{K}$, rechts ist 0 der Nullvektor]

(ii) $\lambda 0 = 0$

(iii) $\lambda v = 0 \implies \lambda = 0$ oder $v = 0$

(iv) $(-1)v = -v$

Ich schreibe hier mal der Vollständigkeit halber einen relativ knapp formulierten Beweis auf und will Sie damit ermuntern, den einmal in Ruhe durchzudenken oder mit StudienkollegInnen durchzugehen — am besten selbst aufschreiben und bei den einzelnen Schritten zusätzlich heraussuchen, auf welche der Axiome (V1)-(V8) wir uns jeweils stützen. Dann können Sie zumindest einmal im Kleinen einen sehr formalen Aspekt der Mathematik nachvollziehen und sehen, wie diese Maschine prinzipiell läuft. Ich erwarte aber nicht von Ihnen, dass Sie diesen konkreten Beweis für die Prüfung auswendig lernen.

Beweis: (i) $0v = (0+0)v = 0v + 0v$, daher $0v = 0$.

(ii) $\lambda 0 = \lambda(0+0) = \lambda 0 + \lambda 0$, daher $\lambda 0 = 0$.

(iii) Sei $\lambda v = 0$. Falls $\lambda = 0$, dann fertig. Falls $\lambda \neq 0$, dann $v = 1v = (\lambda^{-1}\lambda)v = \lambda^{-1}(\lambda v) = \lambda^{-1}0 = 0$.

(iv) $v + (-1)v = 1v + (-1)v = (1-1)v = 0$, daher $(-1)v = -v$. □

3.4. Definition: Sei V ein $\mathbb{K}$-VR und $W \subseteq V$ mit $W \neq \emptyset$. Dann heißt W *Teilraum* (kurz: TR) von V, falls W mit den von V „geerbten“ Operationen selbst ein $\mathbb{K}$-VR ist, d.h. es gilt:

(i) $w_1 + w_2 \in W \quad \forall w_1, w_2 \in W$

(ii) $\lambda w \in W \quad \forall \lambda \in \mathbb{K}, w \in W$

(iii) (V1)-(V8) gelten für W (statt V) und $\mathbb{K}$.

Bemerkung: *Wenn (i) und (ii) erfüllt sind, gilt (iii) automatisch.*
Das ist für (V1) und (V4)-(V8) klar, weil die sogar für alle $v \in V$ gelten.
(V2): Wir zeigen, dass $0 \in W$ gilt — die Gleichungen $0 + w = w + 0 = w$ gelten dann ohnehin. Sei $w \in W$ beliebig [W ist nicht leer!], dann ist nach 3.3 und (ii) auch $-w = (-1)w \in W$ und nach (i) somit $0 = w - w = w + (-1)w \in W$.
(V3): Wir haben gerade schon gesehen, dass $-w = (-1)w \in W$ gilt. Weil wir auch $0 \in W$ gezeigt haben, dient somit $-w$ auch für die Vektoraddition in W als inverses Element zu $w \in W$.

3.5. Beispiele:

1. $P_n(\mathbb{R})$ ist ein TR von $P(\mathbb{R})$.
   c_0 ist ein TR von l^∞ und l^∞ ist ein TR von $\mathbb{R}^\mathbb{N}$.
   $D[a,b]$ ist TR von $C[a,b]$ und $C[a,b]$ ist TR von $F[a,b]$.
   Die letzten beiden Räume unter 3.2.5 und $P(\mathbb{R})$ sind Teilräume von $D(\mathbb{R})$.
   [$D(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ differenzierbar}\}$]

2. $W := \{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1 + 2x_2 + 3x_3 = 0\}$ ist ein TR von $\mathbb{R}^3$: Es ist $W \neq \emptyset$, weil offensichtlich $0 \in W$; seien $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in W$ [d.h. $x_1 + 2x_2 + 3x_3 = 0$], $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in W$ [d.h. $y_1 + 2y_2 + 3y_3 = 0$] und $\lambda \in \mathbb{R}$, dann gilt

$$\begin{aligned} (x_1 + y_1) + 2(x_2 + y_2) + 3(x_3 + y_3) &= x_1 + 2x_2 + 3x_3 + y_1 + 2y_2 + 3y_3 = 0 + 0 = 0, \\ (\lambda x_1) + 2(\lambda x_2) + 3(\lambda x_3) &= \lambda(x_1 + 2x_2 + 3x_3) = \lambda \cdot 0 = 0 \end{aligned}$$

und daher $x + y \in W$ und $\lambda x \in W$.

3. Die Lösungsmenge jedes homogenen linearen Gleichungssystems mit m Gleichungen in n Unbekannten ist ein Teilraum des $\mathbb{R}^n$ [bzw. $\mathbb{K}^n$]. Der Beweis dafür verläuft im wesentlichen wie in 2., nur dass eben jeweils m Bedingungen nachzurechnen sind.

4. Wir betrachten die Lösungsmenge L der *Differentialgleichung* (DGL)

$$f'' + af' + bf = 0,$$

wobei $a, b \in \mathbb{R}$ gegebene Konstanten sind und alle zeimal differenzierbaren Funktionen $f : \mathbb{R} \to \mathbb{R}$ gesucht werden, die obige Gleichung erfüllen. (Ein Beispiel für so eine DGL mit $a = r/m$ und $b = k/m$ ergibt sich beim gedämpften harmonischen Oszillator mit Masse m, Reibungskonstante r und Federkonstante k.)

Die Lösungsmenge L bildet einen Teilraum von $D(\mathbb{R})$: $L \neq \emptyset$, weil die Nullfunktion sicher eine Lösung ist; sind $f \in L$ [d.h. $f'' + af' + bf = 0$], $g \in L$ [d.h. $g'' + ag' + bg = 0$] und $\lambda \in \mathbb{R}$, dann gilt

$$\begin{aligned} (f+g)'' + a(f+g)' + b(f+g) &= f'' + af' + bf + g'' + ag' + bg = 0 + 0 = 0, \\ (\lambda f)'' + a(\lambda f)' + b(\lambda f) &= \lambda(f'' + af' + bf) = \lambda \cdot 0 = 0 \end{aligned}$$

und daher $f + g \in L$ und $\lambda f \in L$.

5. Die sogenannten trivialen Teilräume eines VR V sind $\{0\}$ und V selbst.

3.6. Konstruktion neuer Vektorräume aus gegebenen:

a) Seien V ein Vektorraum, W und U Teilräume von V, dann sind die folgenden beiden Behauptungen sehr leicht nachzuweisen:

Der *Durchschnitt* $U \cap W$ ist stets ein Teilraum von V.

$U + W := \{u + w \mid u \in U, w \in W\}$ heißt die *Summe* von U und W und ist ebenfalls TR von V.

Achtung: $U \cup W$ ist im Allgemeinen kein TR!

b) Seien V und W Vektorräume über K, dann wird das *Produkt* $V \times W$ mit den Operationen $(v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_1)$ und $\lambda \cdot (v, w) := (\lambda v, \lambda w)$ auch zum $\mathbb{K}$-VR.

[Beweis als Übungsaufgabe.]

Bemerkung: Für einen VR V mit TR W gibt es auch noch den sogenannten *Quotienten* oder *Faktorraum* V/W. (Hier wird V/W als Menge der Äquivalenzklassen bzgl. der Äquivalenzrelation $u \sim v :\Leftrightarrow u - v \in W$ definiert. Die VR-Operationen im Faktorraum werden dann auf die Operationen in V für Repräsentanten der Klassen zurückgeführt.)

3.7. Definition: Für Teilräume U, W eines Vektorraumes V mit $U \cap W = \{0\}$ wird $U + W$ auch als $U \oplus W$ geschrieben und als *direkte Summe* von U und W bezeichnet.

Proposition: Für Teilräume U, W des Vektorraumes V sind folgende Aussagen äquivalent:

(i) $U \cap W = \{0\}$

(ii) Jeder Vektor $v \in U + W$ hat nur *eine* Darstellung der Form $v = u + w$ mit $u \in U$ und $w \in W$.
(D.h. falls $u_1, u_2 \in U$ und $w_1, w_2 \in W$ die Gleichung $u_1 + w_1 = u_2 + w_2$ erfüllen, dann muss $u_1 = u_2$ und $w_1 = w_2$ gelten.)

Beweis als Übungsaufgabe.

Wir werden im weiteren Verlauf den Begriff der „Dimension" eines Vektorraumes entwickeln. Damit kann in gewisser Hinsicht die „Größe" von Vektorräumen und Teilräumen beschrieben werden.[4]

Dazu suchen wir im betreffenden Vektorraum ein $\underbrace{\text{„möglichst kleines}}_{\substack{\text{keine überflüssigen Elemente;}\\ \text{Bausteine voneinander } \textit{unabhängig}}}$ $\underbrace{\text{Bausteinsystem“}}_{\substack{\text{ausreichend groß, um durch VR-Ope-}\\ \text{rationen alle Elemente des VR zu } \textit{erzeugen}}}$ (bezeichnet als „Basis").

3.8. Definition: Eine *Linearkombination* der Vektoren $v_1, \ldots, v_k$ ist ein Ausdruck der Gestalt

$$\lambda_1 v_1 + \ldots + \lambda_k v_k \quad \left(= \sum_{i=1}^{k} \lambda_i v_i \right), \qquad \text{wobei } \lambda_i \in \mathbb{K} \ (i = 1, \ldots, k).$$

3.9. Definition: Ein System von Vektoren $v_1, \ldots, v_k$ $(k \geq 2)$ heißt *linear unabhängig*, wenn es unmöglich ist, einen dieser Vektoren als Linearkombination der übrigen auszudrücken.
[Andernfalls wäre dann dieser eine Vektor bei der Bildung aller Linearkombinationen entbehrlich, also „überflüssig".]

[4] Der Begriff der Anzahl an Elementen ist dazu ungeeignet, weil z.B. jeder $\mathbb{R}$-Vektorraum $V \neq \{0\}$ unendlich viele Elemente enthält.

3.10. Definition: Ein System von Vektoren $v_1, \ldots, v_k$ in einem VR V heißt *Erzeugendensystem von V*, wenn sich jeder Vektor $v \in V$ als Linearkombination $v = \lambda_1 v_1 + \ldots + \lambda_k v_k$ mit geeigneten $\lambda_i \in \mathbb{K}$ darstellen lässt.

3.11. Definition: Eine *Basis* eines Vektorraumes V ist ein linear unabhängiges Erzeugendensystem von V.

Beim Lösen von homogenen linearen Gleichungssystemen [ebenso: hom. lin. DGL] möchte man gern eine Basis des Lösungsraumes bestimmen.

3.12. Beispiele: 1. In $V = \mathbb{R}^n$ (analog für $\mathbb{C}^n$ und $\mathbb{K}^n$) bildet das System

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \; \ldots, \; e_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

die sogenannte *Standardbasis*. Das System ist linear unabhängig, denn kein e_i lässt sich durch die restlichen e_j $(j \neq i)$ ausdrücken (letztere haben in der i-ten Komponente ja alle nur 0 stehen). Außerdem ist es ein Erzeugendensystem, weil für jedes $x \in \mathbb{R}^n$ gilt

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{pmatrix} + \ldots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x_n \end{pmatrix} = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n.$$

2. $M(2,2)$ hat $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ als Basis.

3. $P_n(\mathbb{R})$ hat $1, x, x^2, x^3, \ldots, x^n$ als Basis. $P(\mathbb{R})$ hat $1, x, x^2, x^3, \ldots$ als Basis.

4. $\mathbb{R}^3$ hat $b_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, b_2 = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}, b_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ als Basis: Für die lineare Unabhängigkeit ist zu zeigen, dass jede der Annahmen $b_1 = \lambda_2 b_2 + \lambda_3 b_3$ oder $b_2 = \mu_1 b_1 + \mu_3 b_3$ oder $b_3 = \nu_1 b_1 + \nu_2 b_2$ auf einen Widerspruch führt.[5] Weiters ist zu zeigen, dass für einen beliebig gegeben Vektor $a \in \mathbb{R}^3$ das Gleichungssystem $\sigma_1 b_1 + \sigma_2 b_2 + \sigma_3 b_3 = a$ für $\sigma_1, \sigma_2, \sigma_3$ stets lösbar ist (z.B. mittels Gauß-Elimination). [Details als zusätzliche Übung empfohlen!]

Wir werfen hier zwei Fragen auf: Hat jeder Vektorraum eine Basis?
Haben je zwei Basen eines VR gleich viele Elemente?

[5] Z.B. durch Betrachtung der einzelnen Zeilen. Später werden wir hierfür eine einfachere Methode haben.

Sobald geklärt ist, dass die Antworten jeweils 'ja' lauten, können wir festsetzen:
„Dimension eines Vektorraumes" = Anzahl der Basiselemente.

$P(\mathbb{R})$ im Beispiel 3 zeigt, dass die Begriffe „linear unabhängig" und „Erzeugendensystem" auch für unendliche Systeme sinnvoll sind: Lies in 3.9 und 3.10 jeweils: „als *endliche* Linearkombination".

Definition 3.9 hat Nachteile: sie ist umständlich bei Beweisen und Rechnungen; und sie ist nicht anwendbar auf einen einzelnen Vektor v_1. Daher ist folgende Umformulierung handelsüblich.

3.13. Definition: Ein System von Vektoren $v_1, \ldots, v_k$ $(k \geq 1)$ heißt *linear unabhängig*, wenn der Nullvektor nur auf die „triviale" Weise als Linearkombination von $v_1, \ldots, v_k$ dargestellt werden kann, das heißt:

Falls $\lambda_1 v_1 + \ldots + \lambda_k v_k = 0$ gilt, folgt notwendig $\lambda_1 = \ldots = \lambda_k = 0$.

Andernfalls heißt $v_1, \ldots, v_k$ *linear abhängig.*

3.14. Proposition: Für $k \geq 2$ ist 3.9 äquivalent zu 3.13.

Beweis: 3.9 $\Rightarrow$ 3.13: $v_1, \ldots, v_k$ möge 3.9 erfüllen.

Indirekt: Angenommen, 3.13 ist NICHT erfüllt, d.h.

$$\exists \lambda_1, \ldots, \lambda_k \in \mathbb{K}: \text{ mindestens ein } \lambda_j \neq 0 \text{ und } \lambda_1 v_1 + \ldots \lambda_k v_k = 0.$$

Dann folgt

$$v_j = -\frac{1}{\lambda_j} \underbrace{(\lambda_1 v_1 + \ldots + \lambda_k v_k)}_{\text{ohne } v_j}$$

und daher wäre 3.9 verletzt. Das ist ein WIDERSPRUCH! (Also muss doch 3.13 gelten.)

3.13 $\Rightarrow$ 3.9: $v_1, \ldots, v_k$ möge 3.13 erfüllen.

Indirekt: Angenommen, 3.9 ist NICHT erfüllt, d.h.

$$\exists j: \ v_j = \underbrace{\mu_1 v_1 + \ldots + \mu_k v_k}_{\text{ohne } v_j} \qquad \text{mit geeigneten } \mu_i \in \mathbb{K}.$$

Dann folgt

$$\mu_1 v_1 + \ldots + (-1) v_j + \ldots + \mu_k v_k = 0$$

und somit wäre also 0 als nichttriviale Linearkombination von $v_1, \ldots, v_k$ darstellbar — ein WIDERSPRUCH! (Also muss doch 3.9 gelten.) $\square$

Ab jetzt verwenden wir nur mehr 3.13 als Definition der linearen Unabhängigkeit!

Zum Beispiel können wir nun die lineare Unabhängigkeit der Vektoren b_1, b_2, b_3 aus 3.12.4 gemäß 3.13 mittels Gauß-Elimination beweisen: Die Gleichung $\lambda_1 b_1 + \lambda_2 b_2 + \lambda_3 b_3 = 0$ aufgefasst als Gleichungssystem für $\lambda_1, \lambda_2, \lambda_3$ ergibt die Folge von Schemata

$$\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & -2 & -1 & 0 \end{array} \quad \Rightarrow \quad \begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & -1 & 0 \end{array} \quad \Rightarrow \quad \begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{array} \quad \Rightarrow \quad \lambda_1 = \lambda_2 = \lambda_3.$$

Übrigens ist hier die Koeffizientenmatrix natürlich identisch mit jener, die man in demselben Beispiel beim Nachweis, dass b_1, b_2, b_3 ein Erzeugendensystem ist, vewendet:

$$\begin{array}{ccc|c} 1 & 0 & 0 & a_1 \\ 0 & 1 & 1 & a_2 \\ 1 & -2 & -1 & a_3 \end{array} \quad \Rightarrow \quad \begin{array}{ccc|c} 1 & 0 & 0 & a_1 \\ 0 & 1 & 1 & a_2 \\ 0 & -2 & -1 & a_3 - a_1 \end{array} \quad \Rightarrow \quad \begin{array}{ccc|c} 1 & 0 & 0 & a_1 \\ 0 & 1 & 1 & a_2 \\ 0 & 0 & 1 & -a_1 + 2a_2 + a_3 \end{array} \quad \text{lösbar!}$$

3.15. Proposition: v ist linear unabhängig $\iff$ $v \neq 0$.

Beweis: $\Rightarrow$: Sei v linear unabhängig. Indirekt: $v = 0 \Rightarrow 1 \cdot v = 0$ ist nichttriviale Linearkombination $\Rightarrow$ v linear abhängig — WIDERSPRUCH! Also $v \neq 0$.

$\Leftarrow$: Sei $\lambda \cdot v = 0$. Falls $\lambda \neq 0$ gelte, dann auch $v = 1 \cdot v = \frac{1}{\lambda}\lambda v = 0$ – WIDERSPRUCH, also $\lambda = 0$. Somit v linear unabhängig. □

3.16. Wenn wir die homogene Variante des Gleichungssystems in 1.4 betrachten, dann können die Vektoren des Lösungsraumes $U \subseteq \mathbb{R}^7$ wie folgt dargestellt werden:

$$x = t_1 \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{v_1} + t_2 \underbrace{\begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{v_2} + t_3 \underbrace{\begin{pmatrix} 0 \\ -6 \\ 5 \\ 0 \\ 2 \\ -2 \\ 1 \end{pmatrix}}_{v_3} \qquad t_1, t_2, t_3 \in \mathbb{R}.$$

Offenbar ist v_1, v_2, v_3 ein Erzeugendensystem von U. Mit etwas Scharfblick (z.B. auf die Zeilen 1, 4 und 7) sehen wir unmittelbar, dass v_1, v_2, v_3 auch linear unabhängig sind, also eine Basis von U bilden. Wie können wir aber in Situationen mit viel größeren Erzeugendensystemen vorgehen, um eine Basis zu finden?

Wir machen dazu eine Art Gedankenexperiment: Es sei $v_1, \ldots\ldots\ldots\ldots, v_N$ ein (möglicher Weise sehr laaaaaaanges) Erzeugendensystem eines Vektorraumes V. Ist ein passendes und möglichst kleines Teilsystem $v_1, \ldots, v_k$ $(k \leq N)$ eine Basis von V?

Wir untersuchen alle Teilsysteme $v_1, \dots, v_k$ für $k = 1, \dots, N$ und notieren jeweils:

falls $v_1, \dots, v_k$ linear unabhängig ist, das Zeichen $\star$ bei k,

falls $v_1, \dots, v_k$ ein Erzeugendensystem ist, das Zeichen $\circ$ bei k.

Wir erhalten etwa folgendes Bild: $\dfrac{\begin{matrix}\star & \star & \star & \dots & \dots & \dots & \circ & \circ\end{matrix}}{\begin{matrix}1 & 2 & 3 & \dots & n & \dots & N-1 & N\end{matrix}}$

$\boxed{?}$ $\exists n$ mit $\star$ und $\circ$?

Plausibel ist: 1. Die Zeichen $\star$ bilden von links weg eine lückenlose Folge. (Schlimmstenfalls leer.)
2. Die Zeichen $\circ$ bilden von rechts weg eine lückenlose Folge. (Mindestens ein Zeichen)
3. Die Folgen von $\star$ und $\circ$ können sich höchstens bei einer Zahl überlappen.

Daher bleiben also nur zwei Möglichkeiten:

Entweder $\dfrac{\begin{matrix}\star & \dots & \star & \text{leer} & \circ & \dots & \circ\end{matrix}}{\begin{matrix}1 & \dots & \dots & \dots & \dots & \dots & N\end{matrix}}$ d.h. keines der $v_1, \dots, v_k$ ist eine Basis

oder $\dfrac{\begin{matrix}\star & \dots & \star & \overset{\star}{\circ} & \circ & \dots & \circ\end{matrix}}{\begin{matrix}1 & \dots & \dots & n & \dots & \dots & N\end{matrix}}$ d.h. $v_1, \dots, v_n$ ist eine Basis.

4. Beim Vergleich von zwei Systemen $v_1, \dots, v_N$ und $w_1, \dots, w_M$ kann nie ein $\circ$ links von einem $\star$ stehen: Wenn z.B. $v_1, \dots, v_k$ ein Erzeugendensystem bildet, dann können die $k+1$ Vektoren $w_1, \dots, w_{k+1}$ sicher nicht mehr linear unabhängig sein.
5. Je zwei Basen haben gleichviele Elemente.

Die Diskussion in 3.16 soll eine Motivation für den folgenden, wieder mehr formalen Abschnitt darstellen, in dessen Verlauf wir in den Sätzen 3.17-3.20 auch die obigen Punkte 1.-5. rechtfertigen werden.

3.17. Proposition: Sei V ein Vektorraum und $S_1 \subseteq S_2 \subseteq V$, dann gilt:

(i) S_2 linear unabhängig $\Longrightarrow$ S_1 linear unabhängig,

(ii) S_1 Erzeugendensystem $\Longrightarrow$ S_2 Erzeugendensystem.

Beweis folgt direkt aus den Definitionen. □

3.18. Proposition: Sei V ein Vektorraum und $S_1 \subseteq S_2 \subseteq V$ mit $S_1 \neq S_2$, dann gilt:

(i) Ist S_2 (und damit auch S_1) linear unabhängig, dann ist S_1 kein Erzeugendensystem.

(ii) Ist S_1 (und damit auch S_2) ein Erzeugendensystem, dann ist S_2 nicht linear unabhängig.

Beweis: (i) Angenommen, S_1 wäre ein Erzeugendensystem. Wähle $v \in S_2 \setminus S_1$, dann gibt es $v_1, \dots, v_k \in S_1$, sodass mit geeigneten $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ gilt $v = \sum_{i=1}^{k} \lambda_i v_i$. Daher ist $S_1 \cup \{v\}$ linear abhängig, also S_2 linear abhängig — WIDERSPRUCH!

(ii) Lies den Beweis von (i) als direkten Beweis. □

3.19. Satz (Austauschsatz von Steinitz): „Linear unabhängige Systeme lassen sich in Erzeugendensysteme einbauen", d.h.:

Sei $w_1, \ldots, w_k$ linear unabhängig und $v_1, \ldots, v_n$ ein Erzeugendensystem. Dann ist $k \leq n$ und nach geschickter Umnummerierung der v_i ist $w_1, \ldots, w_k, v_{k+1}, \ldots, v_n$ ebenfalls ein Erzeugendensystem.

Beweis: Vollständige Induktion nach k (bei fixem n).

$k = 1:$ w_1 ist linear unabhängig, daher $w_1 \neq 0$. Daher muss das Erzeugendensystem mindestens einen Vektor enthalten, also ist $n \geq 1$.

Einbau von w_1: Mit geeigneten Skalaren λ_i ist $w_1 = \sum_{i=1}^{n} \lambda_i v_i$. Hier ist mindestens ein $\lambda_i \neq 0$, da $w_1 \neq 0$. OBdA[6] ist $i = 1$ (sonst Umnummerieren). Somit folgt

$$v_1 = \frac{1}{\lambda_1}(w_1 - \lambda_2 v_2 - \ldots - \lambda_n v_n).$$

Da jedes $v \in V$ durch $v_1, \ldots, v_n$ darstellbar ist, ist es nach obiger Gleichung auch durch $w_1, v_2, \ldots, v_n$ darstellbar, letzteres also ein Erzeugendensystem.

$(k-1) \mapsto k:$ Die Behauptung sei richtig für $k-1$, d.h. es gilt $k - 1 \leq n$ (somit $k \leq n+1$) und nach Umnummerierung: $(\star)$ $w_1, \ldots, w_{k-1}, v_k, \ldots, v_n$ ist ein Erzeugendensystem.

Wir wissen also schon, dass $k \leq n+1$ gilt. Ist Gleichheit möglich? Angenommen, es wäre $k = n+1$, d.h. $k-1 = n$, dann wäre also bereits $w_1, \ldots, w_{k-1}$ ein Erzeugendensystem. Das ist aber gemäß 3.18(i) ein WIDERSPRUCH zur Tatsache, dass $w_1, \ldots, w_k$ linear unabhängig ist. Somit ist $k < n+1$, daher $k \leq n$.

Einbau von w_k: Es gilt $w_k = \lambda_1 w_1 + \ldots + \lambda_{k-1} w_{k-1} + \lambda_k v_k + \ldots + \lambda_n v_n$. Hier muss mindestens eines der $\lambda_k, \ldots, \lambda_n$ ungleich 0 sein, weil andernfalls $w_1, \ldots, w_k$ linear abhängig wäre. OBdA ist $\lambda_k \neq 0$. Dann folgt

$$v_k = \frac{1}{\lambda_k}(w_k - \lambda_1 w_1 - \ldots - \lambda_{k-1} w_{k-1} - \lambda_{k+1} v_{k+1} - \ldots - \lambda_n v_n)$$

Da jedes $v \in V$ durch das Erzeugendensystem $(\star)$ darstellbar ist, ist es nach obiger Gleichung auch durch $w_1, \ldots, w_k, v_{k+1}, \ldots, v_n$ darstellbar, letzteres also auch ein Erzeugendensystem. □

3.20. Korollar: Je zwei Basen eines Vektorraumes haben gleichviele Elemente.

Beweis: Sind $v_1, \ldots, v_n$ und $w_1, \ldots, w_m$ Basen, dann ist nach 3.19 sowohl $n \leq m$ als auch $m \leq n$, also $n = m$. □

[6] OBdA = „Ohne Beschränkung der Allgemeinheit".

Im Folgenden sagen wir „Länge eines Systems“ statt „Anzahl der Elemente eines Systems“.

3.21. Ein Vektorraum V heißt *endlichdimensional*, wenn es eine Obergrenze für die Länge linear unabhängiger Systeme gibt; ansonsten ist V *unendlichdimensional*.

Ist also V endlichdimensional und n diese Obergrenze, dann gibt es ein linear unabhängiges System mit n Elementen $v_1, \ldots, v_n$, aber jedes System mit mehr als n Vektoren ist linear abhängig.

Der Vektorraum $P(\mathbb{R})$ ist offenbar unendlichdimensional (vgl. 3.12.3).

3.22. Satz: Jeder endlichdimensionale Vektorraum hat eine Basis.
[Dies gilt auch für unendlichdimensionale Vektorräume; das beweisen wir hier aber nicht.]

Beweis: Sei n die Obergrenze gemäß 3.21 und $v_1, \ldots, v_n$ ein linear unabhängiges System in V. Es ist noch zu zeigen, dass letzteres auch ein Erzeugendensystem von V ist.

Sei $w \in V$ beliebig, dann ist $v_1, \ldots, v_n, w$ sicher linear abhängig (weil $n+1$ Vektoren). Daher gibt es $\lambda_1, \ldots, \lambda_n, \mu \in \mathbb{K}$, nicht alle gleich 0, sodass gilt

$$\lambda_1 v_1 + \ldots + \lambda_n v_n + \mu w = 0.$$

Es kann hier unmöglich $\mu = 0$ sein, weil sonst $v_1, \ldots, v_n$ linear abhängig wären. Also erhalten wir

$$w = \left(\frac{-\lambda_1}{\mu}\right) \cdot v_1 + \ldots + \left(\frac{-\lambda_n}{\mu}\right) \cdot v_n$$

und, weil $w \in V$ beliebig war, weiters, dass $v_1, \ldots, v_n$ ein Erzeugendensystem von V ist. □

Sei V endlichdimensional; dann existiert mindestens eine Basis (3.22); jede Basis ist endlich (3.21); je zwei Basen haben gleichviele Elemene (3.20). Daher ist die folgende Definition einwandfrei:

3.23. Definition: Ist V ein endlichdimensionaler Vektorraum und $v_1, \ldots, v_n$ eine Basis. Dann setzen wir $\dim V := n$ und nennen diese Zahl die *Dimension* von V. Für einen unendlichdimensionalen Vektorraum V setzen wir $\dim V := \infty$.

Aus dem Beweis von 3.22 ist ersichtlich: $\dim V$ ist die maximale Länge linear unabhängiger Systeme in V.

3.24. Beispiele: Mittels Standardbasen (vgl. 3.12.1) sehen wir, dass $\dim \mathbb{R}^n = n$, beziehungsweise allgemeiner $\dim \mathbb{K}^n = n$, gilt.

$\dim M(m, n; \mathbb{K}) = m \cdot n$ sieht man mittels einer Basis aus Matrizen ähnlich wie in 3.12.2

$\dim P(\mathbb{R}) = \infty$ wurde oben schon besprochen.

$\dim P_n(\mathbb{R}) = n+1$ wie wir aus der Basis in 3.12.3 sehen.

Offensichtlich unendlichdimensional sind die Folgenräume c_0, l^∞ und $\mathbb{R}^\mathbb{N}$, weil hier jeweils das unendliche System $e_1, e_2, e_3 \ldots$ linear unabhängig ist, wobei e_j die Folge bezeichnet, deren j-tes Glied 1 ist und alle anderen Glieder 0 sind.

Weitere Beispiele für unendlichdimensionale Vektorräume sind $F[a,b]$, $C[a,b]$, $D[a,b]$ und $TP(\mathbb{R})$ (ohne Beweis).

Für $\omega \neq 0$ fix und $V = \{f : \mathbb{R} \to \mathbb{R}, f(t) = \lambda \sin(\omega t) + \mu \cos(\omega t) \mid \lambda, \mu \in \mathbb{R}\}$ (vgl. 3.2.5) ist $\dim V = 2$.
(Versuchen Sie selbst einen Beweis!)

Sei $TP_n(\mathbb{R})$ definiert ähnlich wie $TP(\mathbb{R})$, aber Summation jeweils nur bis n.
Dann gilt $\dim TP_n(\mathbb{R}) = 2n+1$. (Warum?)

3.25. Warnung: Die Begriffe „linear unabhängig“, „Erzeugendensystem“, „Basis“ und „Dimension“ sollten eigentlich stets explizit mit Hinweis auf den zu Grunde gelegten Skalarenkörper $\mathbb{K}$ beschrieben werden; es sollte also z.B. heißen $\dim_\mathbb{K} V = n$, wenn V ein endlichdimensionaler $\mathbb{K}$-VR ist und es eine $\mathbb{K}$-Basis aus n Elementen gibt; letztere ist also ein $\mathbb{K}$-Erzeugendensystem (d.h. alle $\mathbb{K}$-Linearkombinationen ergeben ganz V), das zusätzlich *linear unabhängig über* $\mathbb{K}$ ist (d.h. es gibt keine nichttriviale $\mathbb{K}$-Linearkombination des Nullvektors).

Zum Beispiel ist natürlich $\underline{\dim_\mathbb{C} \mathbb{C} = 1}$, denn ist $\mathbb{C} = \mathbb{C}^1$ hat als Basis $1 \in \mathbb{C}$: jedes $z \in \mathbb{C}$ erfüllt ja $z = z \cdot 1$, wobei hier z als Skalar und 1 als Basisvektor fungiert.
Nun können wir aber $\mathbb{C}$ auch als Vektorraum über $\mathbb{R}$ auffassen, weil die Multiplikation $\mathbb{R} \times \mathbb{C} \to \mathbb{C}$ auch die erforderten VR-Axiome erfüllt. Es gilt $\underline{\dim_\mathbb{R} \mathbb{C} = 2}$, weil das System 1, i eine Basis bildet: Zunächst ist $1, i$ linear unabhängig über $\mathbb{R}$, weil die Gleichung $\lambda 1 + \mu i = 0$ mit $\lambda, \mu \in \mathbb{R}$ sofort $\lambda = \mu = 0$ impliziert (einfach Real- und Imaginärteil nehmen!); außerdem ist $1, i$ ein $\mathbb{R}$-Erzeugendensystem von $\mathbb{C}$, weil wir jedes $z \in \mathbb{C}$ in der Form $z = \lambda 1 + \mu i$ mit $\lambda = \operatorname{Re} z \in \mathbb{R}$ und $\mu = \operatorname{Im} z \in \mathbb{R}$ schreiben können. Somit ist $1, i$ eine $\mathbb{R}$-Basis von $\mathbb{C}$.
(Selbstverständlich ist $1, i$ linear abhängig über $\mathbb{C}$: $i \cdot 1 + (-1) \cdot i = 0$.)

3.26. Satz: Sei V ein endlichdimensionaler $\mathbb{K}$-VR. Dann gilt:

(i) Jedes linear unabhängige System kann zu einer Basis verlängert werden.

(ii) Jedes endliche Erzeugendensystem kann zu einer Basis verkürzt werden.

Beweis: (i) Nach 3.22 gibt es eine Basis, in die das linear unabhängige System gemäß 3.19 eingebaut werden kann.
(ii) Falls das Erzeugendensystem selbst bereits linear unabhängig ist, sind wir fertig. Andernfalls gibt es einen Vektor v in dem System, der sich als Linearkombination der restlichen Vektoren darstellen lässt. Wir können v also weglassen und haben immer noch ein Erzeugendensystem. Diesen Prozess setzen wir so lange fort, bis der Rest ein linear unabhängiges Erzeugendensystem ergibt. □

3.27. Satz: Sei V ein $\mathbb{K}$-VR und S ein System von Elementen aus V. Dann sind folgende Aussagen äquivalent:

(i) S ist linear unabhängig und ein Erzeugendensystem (d.h. S ist eine Basis).

(ii) S ist linear unabhängig und nicht linear unabhängig verlängerbar.

(iii) S ist ein Erzeugendensystem und nicht als Erzeugendensystem verkürzbar.

(iv) S ist ein Erzeugendensystem und jedes $v \in V$ hat nur eine *einzige* Darstellung durch Elemente aus S.

Ist überdies $n = \dim V < \infty$, dann sind weiters äquivalent:

(v) S ist linear unabhängig und hat Länge n.

(vi) S ist ein Erzeugendensystem und hat Länge n.

Beweis: (i) $\Rightarrow$ (ii): Sei S_1 die Verlängerung von S mit $v \in V$. Der Vektor v ist aber schon durch Elemene aus S darstellbar, daher muss S_1 linear abhängig sein.

(ii) $\Rightarrow$ (iii): Für beliebiges $v \in V$ ist S, verlängert um v, nach (ii) sicher linear abhängig. Daher gibt es eine nichttriviale Linearkombination

$$\sum_{i=1}^{k} \lambda_i v_i + \lambda v = 0$$

mit $v_i \in S$. Es gilt sicherlich $\lambda \neq 0$, weil andernfalls S linear abhängig wäre. Somit erhalten wir

$$v = -\frac{1}{\lambda} \sum_{i=1}^{k} \lambda_i v_i = \sum_{i=1}^{k} \left(-\frac{\lambda_i}{\lambda}\right) v_i$$

und S ist als Erzeugendensystem erkannt. Sowohl S als auch jede seiner Verkürzungen ist linear unabhängig. Nach 3.18(i) kann keine der Verkürzungen ein Erzeugendensystem sein.

(iii) $\Rightarrow$ (iv): Hätte $v \in V$ zwei Darstellungen $\sum \lambda_i v_i = v = \sum \mu_i v_i$, wobei $v_i \in S$ und nicht immer $\lambda_i = \mu_i$ gilt, dann wäre wegen $\sum(\lambda_i - \mu_i) v_i = 0$ eines der v_i durch die restlichen ausdrückbar. Somit wäre S als Erzeugendensystem verkürzbar — ein WIDERSPRUCH!

(iv) $\Rightarrow$ (i): Speziell hat 0 nur eine Darstellung; diese muss also die triviale Darstellung sein. Daher ist S linear unabhängig.

Nun sei zusätzlich $n = \dim V < \infty$.

(i) $\Rightarrow$ (v) und (vi): Folgt nach Definition der Dimension.

(v) $\Rightarrow$ (i): S hat Länge n und ist nach 3.26(i) zu einer Basis (der Länge $n = \dim V$) verlängerbar, ist daher selbst schon eine Basis.

(vi) $\Rightarrow$ (i): S hat Länge n und ist nach 3.26(ii) zu einer Basis (der Länge $n = \dim V$) verkürzbar, ist daher selbst schon eine Basis. $\square$

3.28. Satz (Dimension von Teilräumen): Sei V ein Vektorraum und W ein Teilraum von V. Dann gilt:

(i) $\dim W \le \dim V$

(ii) $\dim W = \dim V < \infty \implies W = V$

Beweis: (i) ist sicher richtig, falls $\dim V = \infty$. Im Falle $\dim V < \infty$ folgt die Behauptung aus dem Zusatz unmittelbar nach Definition 3.23.

(ii): W besitzt eine Basis mit $\dim W = \dim V$ Elementen und diese ist somit ein linear unabhängiges System in V mit $\dim V$ Elementen, also eine Basis von V. □

3.29. Beispiele: Eine vollständige Liste von Teilräumen W im $\mathbb{R}^1$, $\mathbb{R}^2$ und $\mathbb{R}^3$:

$\mathbb{R}^1$: $\dim W = 0$: $W = \{0\}$,
$\dim W = 1$: $W = \mathbb{R}^1$.

$\mathbb{R}^2$: $\dim W = 0$: $W = \{0\}$,
$\dim W = 1$: W hat Basis aus einem l.u. Vektor v, d.h. $W = \{\lambda v \mid \lambda \in \mathbb{R}\}$ .. Gerade durch 0,
$\dim W = 2$: $W = \mathbb{R}^2$.

$\mathbb{R}^3$: $\dim W = 0$: $W = \{0\}$,
$\dim W = 1$: W hat Basis aus einem l.u. Vektor v, d.h. $W = \{\lambda v \mid \lambda \in \mathbb{R}\}$ .. Gerade durch 0,
$\dim W = 2$: W hat Basis aus zwei l.u. Vektoren v, w,
d.h. $W = \{\lambda v + \mu w \mid \lambda, \mu \in \mathbb{R}\}$ ... Ebene durch 0,
$\dim W = 3$: $W = \mathbb{R}^3$.

Frage: Wie passt hier das Beispiel 3.5.2 hinein?

3.30. Satz: Für Teilräume U und W eines Vektorraumes V gilt

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W.$$

Beweis: Falls $\dim U = \infty$ oder $\dim W = \infty$ ist, dann folgt auch $\dim(U + W) = \infty$ und beide Seiten der behaupteten Gleichung ergeben ∞.

Wir betrachten nun den Fall, dass $\dim U$ und $\dim W$ endlich sind: Sei $v_1, \dots, v_r$ eine Basis von $U \cap W$; wir ergänzen diese jeweils

($\star$) zu einer Basis $v_1, \dots, v_r, u_1, \dots, u_s$ von U

und

($\star\star$) zu einer Basis $v_1, \dots, v_r, w_1, \dots, w_t$ von W.

Behauptung: Das System S: $v_1, \dots, v_r, u_1, \dots, u_s, w_1, \dots, w_t$ ist eine Basis von $U + W$.

Sobald die Behauptung gezeigt ist, sind wir fertig, weil die entsprechenden Dimensionen sich ja wie folgt ergeben: $\dim(U+W)+\dim(U\cap W) = (r+s+t)+r = (r+s)+(r+t) = \dim U+\dim W$.

S ist linear unabhängig: Sei $\lambda_1 v_1 + \ldots + \lambda_r v_r + \mu_1 u_1 + \ldots + \mu_s u_s + \nu_1 w_1 + \ldots + \nu_t w_t = 0 \quad (+)$, dann ist

$$\underbrace{\sum \lambda_i v_i + \sum \mu_j u_j}_{\in U} = \underbrace{-\sum \nu_k w_k}_{\in W} \ \in U \cap W$$

und somit auch darstellbar als $\sum \rho_i v_i$. Wir erhalten $\sum \rho_i v_i = -\sum \nu_k w_k$, was aber wegen $(\star\star)$ erzwingt, dass alle $\rho_i = 0$ und alle $\nu_k = 0$. Weiters sind nun wegen $(+)$ und $(\star)$ aber auch alle $\lambda_i = 0$ und alle $\mu_j = 0$.

S ist ein Erzeugendensystem: Sei $v = u + w$ mit $u \in U$, $w \in W$. Mittels Basisdarstellungen $u = \sum \lambda_i v_i + \sum \mu_j u_j$ und $w = \sum \lambda'_i v_i + \sum \nu_k w_k$ erhalten wir dann eine Darstellung

$$v = u + w = \sum (\lambda_i + \lambda'_i) v_i + \sum \mu_j u_j + \sum \nu_k w_k$$

von v durch Elemente von S. □

Beispiel: Für $V = \mathbb{R}^3$, U eine Gerade und W eine Ebene mit $U \not\subseteq W$ erhalten wir in der obigen Dimensionsgleichung $3+0 = 1+2$; wenn U und W verschiedene Ebenen sind, $3+1 = 2+2$.

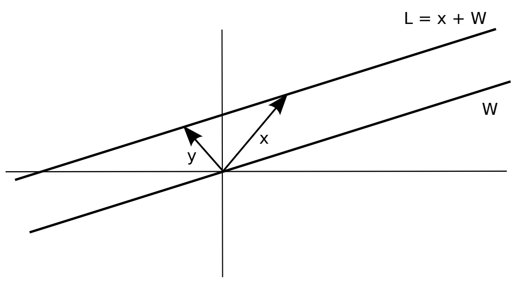
3.31. Korollar: $\dim(U \oplus W) = \dim U + \dim W$.

Für die Behandlung des Lösungsraumes eines inhomogenen linearen Gleichungssystems benötigen wir noch den Begriff des affinen Teilraumes eines Vektorraumes.

3.32. Definition: Sei V ein $\mathbb{K}$-VR. Eine Teilmenge $L \subseteq V$ heißt *affiner Teilraum* von V, wenn L von der Gestalt

$$L = x + W = \{x + w \mid w \in W\}$$

ist, wo $x \in V$ und W ein Teilraum von V ist. Wir setzen $\dim L = \dim W$.



Achtung! Mit $y \neq x$ kann trotzdem $x + W = y + W$ gelten; aber verschiedene W's ergeben auf jeden Fall verschiedene affine Teilräume. Falls $x \in W$ ist, ist $L = x + W = W$, also L selbst ein Teilraum. Im Allgemeinen ist ein affiner Teilraum aber kein Teilraum, weil ersterer ja oft an Null vorbei geht! [Das ist eigentlich ein terminologisches Unglück, weil nun *Teilräume* eben spezielle *affine Teilräume* sind, dies aber in der Namensgebung gerade andersrum klingt; es liegt hier also nicht so vernünftig wie etwa beim Oberbegriff *Auge* mit den Unterbegriffen *linkes Auge* und *rechtes Auge*.]

§4. LINEARE ABBILDUNGEN

4.1. Definition: Seien V und W jeweils $\mathbb{K}$-Vektorräume. Eine Abbildung $\varphi\colon V \to W$ heißt ($\mathbb{K}$-)*linear*, wenn gilt

$$\begin{array}{rcll} \varphi(v_1 + v_2) &=& \varphi(v_1) + \varphi(v_2) & (v_1, v_2 \in V), \\ \varphi(\lambda v) &=& \lambda\varphi(v) & (\lambda \in \mathbb{K}, v \in V). \end{array}$$

Strukturmathematisch gesprochen: „φ respektiert die VR-Operationen“.

4.2. Beispiele:

1. $\varphi\colon D[a,b] \to F[a,b]$, $\varphi(f) := f'$ ist linear: $\varphi(f+g) = (f+g)' = f' + g' = \varphi(f) + \varphi(g)$, und $\varphi(\lambda f) = (\lambda f)' = \lambda f' = \lambda\varphi(f)$, wobei die mittleren Gleichheitszeichen jeweils aus entsprechenden Sätzen in der Analysis folgen.
2. $\varphi\colon C[a,b] \to \mathbb{R}$, $\varphi(f) := \int_a^b f(t)\,dt$ ist linear: folgt analog aus passenden Formeln der Art $\int(f+g) = \int f + \int g$ und $\int\lambda f = \lambda\int f$ in der Analysis.
3. Sei $\varphi\colon \mathbb{R}^n \to \mathbb{R}^m$ linear, dann erhalten wir mittels Standardbasisentwicklung

$$\varphi(x) = \varphi(x_1 e_1 + \ldots + x_n e_n) = x_1\varphi(e_1) + \ldots + x_n\varphi(e_n).$$

Schreiben wir nun die Bilder $\varphi(e_i) \in \mathbb{R}^m$ der Basisvektoren $e_i \in \mathbb{R}^n$ jeweils als Spaltenvektoren

$$\varphi(e_1) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \ldots, \varphi(e_i) = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}, \ldots, \varphi(e_n) = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix},$$

so kommen wir zu einer Matrixdarstellung

$$\varphi(x) = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = A \cdot x \qquad \text{mit} \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

4.3. Satz: (i) Jede lineare Abbildung $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ ist von der Gestalt

$$\varphi(x) = A \cdot x,$$

wobei die Spalten von $A \in M(m,n;\mathbb{K})$ die Bilder der Vektoren der Standardbasis sind.

(ii) Für jede Matrix $A \in M(m,n;\mathbb{K})$ ist die Abbildung $\varphi_A\colon x \mapsto A \cdot x$ linear $\mathbb{K}^n \to \mathbb{K}^m$.

Beweis: (i): analog zu 4.2.3.

(ii): $\varphi_A(x+y) = A\cdot(x+y) = A\cdot x + A\cdot y = \varphi_A(x)+\varphi_A(y)$; $\varphi_A(\lambda x) = A\cdot(\lambda x) = \lambda(A\cdot x) = \lambda\varphi_A(x)$. $\square$

Wir fassen im Folgenden einige der grundlegenden Eigenschaften linearer Abbildungen zusammen.

4.4. Proposition: (i) φ linear $\implies$ $\varphi(0) = 0$

(ii) φ linear $\implies$ $\varphi(v_1 - v_2) = \varphi(v_1) - \varphi(v_2)$

(iii) Die Verknüpfung linearer Abbildungen ist wieder linear.

Beweis: (i): $\varphi(0) = \varphi(0+0) = \varphi(0) + \varphi(0)$, daher $\varphi(0) = 0$.

(ii): $\varphi(v_1) = \varphi((v_1 - v_2) + v_2) = \varphi(v_1 - v_2) + \varphi(v_2)$, daher $\varphi(v_1) - \varphi(v_2) = \varphi(v_1 - v_2)$.

(iii): Seien U, V, W jeweils $\mathbb{K}$-VR und $\psi\colon U \to V$ sowie $\varphi\colon V \to W$ linear, dann berechnen wir für die Verknüpfungsabbildung $\varphi\circ\psi\colon U \to W$ quasi mechanisch

$$\begin{aligned}(\varphi\circ\psi)(u_1+u_2) &= \varphi(\psi(u_1+u_2)) = \varphi(\psi(u_1)+\psi(u_2)) \\ &= \varphi(\psi(u_1)) + \varphi(\psi(u_2)) = (\varphi\circ\psi)(u_1) + (\varphi\circ\psi)(u_2)\end{aligned}$$

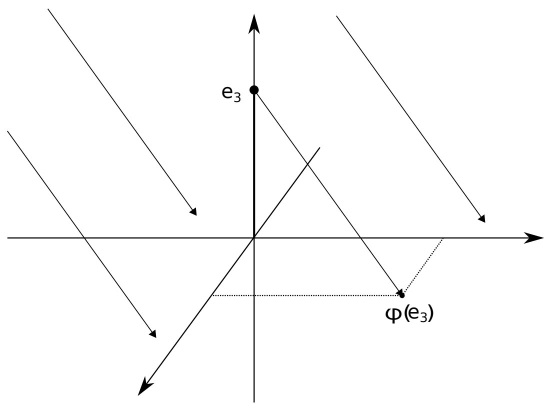
und $(\varphi\circ\psi)(\lambda u) = \varphi(\psi(\lambda u)) = \varphi(\lambda\psi(u)) = \lambda\varphi(\psi(u)) = \lambda(\varphi\circ\psi)(u)$. $\square$

Jede *Nullabbildung* $0\colon V \to W$, $v \mapsto 0$, ist linear (V, W beliebige $\mathbb{K}$-VR); jede *identische Abbildung* $\mathrm{id}_V\colon V \to V$, $v \mapsto v$, ist linear (V beliebiger $\mathbb{K}$-VR).

4.5. Einige Beispiele für das Auftreten bzw. die Verwendung linearer Abbildungen:

1. Als linke Seite von linearen Gleichungssystemen, linearen Differentialgleichungen oder linearen Integralgleichungen.

2. Um Bilder von geometrischen Objekten auf lineare Art zu erzeugen:

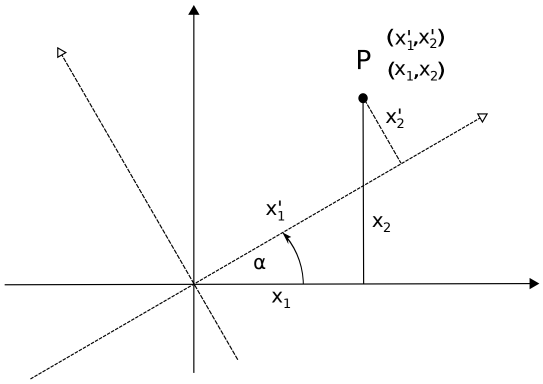


z.B. bei der links skizzierten Parallelprojektion φ, die 2-dimensionale Bilder von 3-dimensionalen Körpern erzeugt: es ist

$$\varphi(e_1) = \begin{pmatrix}1\\0\end{pmatrix},\ \varphi(e_2) = \begin{pmatrix}0\\1\end{pmatrix},\ \varphi(e_3) = \begin{pmatrix}1\\1\end{pmatrix}, \text{ daher}$$

$$\varphi\colon \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} \mapsto \begin{pmatrix}1&0&1\\0&1&1\end{pmatrix}\cdot\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

3. Um eine Situation auf eine andere Weise zu beschreiben:



z.B. in um den Winkel α gedrehten Koordinaten:

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

(vgl. 4.38 später).

Unser **Programm** ist nun die *Entlarvung aller n-dimensionalen Vektorräume über $\mathbb{K}$ als Kopien von $\mathbb{K}^n$* sowie die *Entlarvung aller linearen Abbildungen $\varphi\colon V \to W$ zwischen endlichdimensionalen $\mathbb{K}$-Vektorräumen V und W als Kopien von durch Matrizen gegebene lineare Abbildungen $\varphi_A\colon x \mapsto A\cdot x$.*

4.6. Basisdarstellung von Vektoren: Es sei V ein endlichdimensionaler $\mathbb{K}$-VR mit Basis $B = (b_1,\dots,b_n)$. Für jeden Vektor $v \in V$ haben wir die gemäß 3.27(iv) *eindeutige* Basisdarstellung $v = \sum \lambda_i b_i$ mit Skalaren $\lambda_1,\dots,\lambda_n$. Dies definiert eine Abbildung $\Phi_B\colon V \to \mathbb{K}^n$, gegeben durch

$$V \ni v = \sum_{i=1}^n \lambda_i b_i \quad \xmapsto{\Phi_B} \quad \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{K}^n; \quad [v]_B := \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \dots \text{„Koordinatenvektor von } v \text{ bzgl. } B\text{“.}$$

Wir erhalten ebenso eine Abbildung $\Psi_B\colon \mathbb{K}^n \to V$ durch $\mathbb{K}^n \ni \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \xmapsto{\Psi_B} \sum_{i=1}^n \lambda_i b_i.$

Offenbar gilt $\Psi_B \circ \Phi_B = \mathrm{id}_V$ und $\Phi_B \circ \Psi_B = \mathrm{id}_{\mathbb{K}^n}$, d.h. beide Abbildungen sind bijektiv und $\Psi_B = \Phi_B^{-1}$.

Beide Abbildungen sind auch linear; z.B. für die Additivität von Φ_B mühelos nachgerechnet:

$$\Phi_B(v+w) = \Phi_B\left(\sum \lambda_i b_i + \sum \mu_i b_i\right) = \Phi_B\left(\sum(\lambda_i+\mu_i)b_i\right) = \begin{pmatrix} \lambda_1+\mu_1 \\ \vdots \\ \lambda_n+\mu_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \Phi_B(v) + \Phi_B(w).$$

Analog für $\Phi_B(\lambda v) = \dots$ und die Linearität von Ψ_B wäre sogar noch leichter nachzuprüfen. Wegen $\Psi_B = \Phi_B^{-1}$ folgt letztere aber auch aus allgemeineren Prinzipien, wie wir nun sehen.

4.7. Proposition: $\varphi\colon V \to W$ linear und bijektiv $\implies$ φ^{-1} ist linear

Beweis: Zu $w_1, w_2 \in W$ existieren (eindeutige) $v_1, v_2 \in V$ mit $w_i = \varphi(v_i)$ $(i = 1,2)$, daher $\varphi^{-1}(w_1 + w_2) = \varphi^{-1}(\varphi(v_1) + \varphi(v_2)) = \varphi^{-1}(\varphi(v_1 + v_2)) = v_1 + v_2 = \varphi^{-1}(w_1) + \varphi^{-1}(w_2)$ und weiter $\varphi^{-1}(\lambda w_1) = \varphi^{-1}(\lambda\varphi(v_1)) = \varphi^{-1}(\varphi(\lambda v_1)) = \lambda v_1 = \lambda\varphi^{-1}(w_1)$. □

4.8. Definition: Seien V, W VR über $\mathbb{K}$. Eine lineare Abbildung $\varphi\colon V \to W$ heißt *Isomorphismus*, wenn sie bijektiv ist. In diesem Fall heißen V und W „isomorph“ (griech. gleichgestaltig) via φ und wir schreiben $V \cong W$.

Nun können wir die Erkenntnisse aus 4.6 so formulieren:

4.9. Satz: Jeder n-dimensionale $\mathbb{K}$-VR V ist isomorph zu $\mathbb{K}^n$, also $V \cong \mathbb{K}^n$; bei Vorgabe einer Basis $B = (b_1, \dots, b_n)$ ist ein derartiger Isomorphismus gegeben durch

$$\Phi_B\colon V \to \mathbb{K}^n, \quad v = \sum_{i=1}^n \lambda_i b_i \longmapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = [v]_B.$$

4.10. Beispiele: 1. $V = P_3(\mathbb{R})$ mit der Basis $1, x, x^2, x^3$ ergibt den Isomorphismus

$$a_3x^3 + a_2x^2 + a_1x + a_0 \longmapsto \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad P_3(\mathbb{R}) \to \mathbb{R}^4.$$

2. $V = \mathbb{K}^n$ mit der Standardbasis $E = (e_1, \dots, e_n)$: wir erhalten als Isomorphismus

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n x_i e_i \longmapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = [x]_E,$$

d.h. bzgl. der Standardbasis in $\mathbb{K}^n$ ist jeder Vektor sein eigener Koordinatenvektor ($\Phi_E = \mathrm{id}_{\mathbb{K}^n}$).

3. Vgl. Beispiel 3.12.4: $b_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, b_2 = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}, b_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ ist eine Basis B des $\mathbb{R}^3$.

Zu $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ führt die Berechnung der Koeffizienten für die Darstellung $a = \sigma_1 b_1 + \sigma_2 b_2 + \sigma_3 b_3$ mittels Gauß-Elimination (vgl. auch den Absatz vor 3.15) auf das Schema

$$\begin{array}{ccc|c} 1 & 0 & 0 & a_1 \\ 0 & 1 & 1 & a_2 \\ 0 & 0 & 1 & -a_1 + 2a_2 + a_3 \end{array} \qquad \text{daher} \qquad \begin{array}{l} \sigma_1 = a_1, \\ \sigma_2 = a_1 - a_2 - a_3, \\ \sigma_3 = -a_1 + 2a_2 + a_3, \end{array}$$

$$\text{und somit} \qquad [a]_B = \Phi_B(a) = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_1 - a_2 - a_3 \\ -a_1 + 2a_2 + a_3 \end{pmatrix}.$$

Dieses Beispiel ist im folgenden Sinne „schwierig“: Wir haben im $\mathbb{R}^3$ ohnehin die Standardbasis E, und $[a]_E = a$; wir wählen aber bewusst eine andere Basis B (wie es auch in den Anwendungen oft nötig ist) und behandeln $\mathbb{R}^3$ mit B wie einen abstrakten 3-dimensionalen $\mathbb{R}$-VR, den wir als isomorph zu $\mathbb{R}^3$ „entlarven“! Dieser Isomorphismus von $\mathbb{R}^3$ mit sich selbst ist explizit gegeben durch

$$\mathbb{R}^3 \ni a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \longmapsto \quad [a]_B = \begin{pmatrix} a_1 \\ a_1 - a_2 - a_3 \\ -a_1 + 2a_2 + a_3 \end{pmatrix} \in \mathbb{R}^3.$$

Wir können nun auch mit der Entlarvung der linearen Abbildungen beginnen.

4.11. Seien V, W endlichdimensionale $\mathbb{K}$-VR, $\dim V = n$, $\dim W = m$ und $\varphi\colon V \to W$ linear. Weiters sei $B = (b_1, \ldots, b_n)$ eine Basis von V und $C = (c_1, \ldots, c_m)$ eine Basis von W. Dann erhalten wir das folgende „kommutative Diagramm“ von linearen Abbildungen:

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "\Phi_B"'] & W \arrow[d, "\Phi_C"] \\ \mathbb{K}^n \arrow[r, dashed, "\underbrace{\Phi_C \circ \varphi \circ \Phi_B^{-1}}_{\varphi_A}"'] & \mathbb{K}^m \end{tikzcd}$$

Die Abbildung $\Phi_C \circ \varphi \circ \Phi_B^{-1}$ ist gemäß 4.4(iii) und 4.7 linear $\mathbb{K}^n \to \mathbb{K}^m$, daher gibt es nach 4.3 eine Matrix $A \in M(m, n; \mathbb{K})$, sodass $\Phi_C \circ \varphi \circ \Phi_B^{-1} = \varphi_A$ gilt.

Wir nennen $[\varphi]_{B,C} := A$ die *Matrix von φ bezüglich der Basen B und C*. Wie sieht A explizit aus? Gemäß 4.3 ist die

$$\begin{aligned} \text{1. Spalte von } A &= \text{Bild von } e_1 \text{ unter } \varphi_A = \Phi_C \circ \varphi \circ \Phi_B^{-1}(e_1) = (\Phi_C \circ \varphi)(1 \cdot b_1 + 0 \cdot b_2 + \ldots + 0 \cdot b_n) = \\ &= \Phi_C(\varphi(b_1)) = [\varphi(b_1)]_C \ \ldots \ \text{Koordinatenvektor von } \varphi(b_1); \end{aligned}$$

und analog für die anderen Spaltenvektoren; also ergibt sich

$$[\varphi]_{B,C} = A = \begin{pmatrix} \underbrace{\boxed{}}_{[\varphi(b_1)]_C} & \cdots\cdots & \underbrace{\boxed{}}_{[\varphi(b_n)]_C} \end{pmatrix}$$

$$\text{und} \qquad \underline{[\varphi(v)]_C} = \Phi_C(\varphi(v)) = \underbrace{\Phi_C(\varphi(\Phi_B^{-1}}_{\varphi_A}\underbrace{(\Phi_B(v)}_{[v]_B}))) = \varphi_A([v]_B) = A \cdot [v]_B = \underline{[\varphi]_{B,C} \cdot [v]_B}.$$

4.12. Zusammenfassung (und Umkehrung): Es seien V, W endlichdimensionale $\mathbb{K}$-VR, $\dim V = n$, $\dim W = m$ und $\varphi\colon V \to W$ linear. Weiters sei $B = (b_1, \dots, b_n)$ eine Basis von V, $C = (c_1, \dots, c_m)$ eine Basis von W und

$$\varphi(b_j) = \sum_{i=1}^{m} a_{ij} c_i, \quad \text{d.h.} \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = [\varphi(b_j)]_C \text{ ist der Koordinatenvektor von } \varphi(b_j) \text{ bzgl. } C.$$

Dann ist die *Matrix von φ bzgl. B und C* definiert durch

$$A := [\varphi]_{B,C} = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

und es gilt für alle $v \in V$

$$[\varphi(v)]_C = [\varphi]_{B,C} \cdot [v]_B = A \cdot [v]_B$$

beziehungsweise

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "\Phi_B"'] & W \arrow[d, "\Phi_C"] \\ \mathbb{K}^n \arrow[r, "A"'] & \mathbb{K}^m \end{tikzcd}$$

[wir schreiben ab jetzt einfach A statt φ_A in den Diagrammen].

Umgekehrt liefert jede gegebene Matrix $A \in M(m, n; \mathbb{K})$ eine lineare Abbildung $\varphi\colon V \to W$ durch $\varphi := \Phi_C^{-1} \circ \varphi_A \circ \Phi_B$.

Bezeichnen wir mit $L(V, W)$ die *Menge aller linearen Abbildungen von V nach W*, dann erhalten wir also nach Vorgabe der Basen B und C eine bijektive Abbildung

$$L(V, W) \to M(m, n; \mathbb{K}).$$

4.13. $L(V, W)$ ist ein $\mathbb{K}$-VR bzgl. der Operationen

$$\begin{aligned} (\varphi + \psi)(v) &:= \varphi(v) + \psi(v), \\ (\lambda\varphi)(v) &:= \lambda\varphi(v) \end{aligned}$$

(der Beweis ist straightforward und fad).

$M(m, n; \mathbb{K})$ ist ebenfalls ein $\mathbb{K}$-VR und aus 4.12 folgt unter Beachtung von $(\varphi + \psi)(b_j) = \varphi(b_j) + \psi(b_j)$ usw. nach einfacher Rechnung

$$\begin{aligned} [\varphi + \psi]_{B,C} &= [\varphi]_{B,C} + [\psi]_{B,C}, \\ [\lambda\varphi]_{B,C} &= \lambda[\varphi]_{B,C}, \end{aligned}$$

d.h. die Zuordnung $L(V, W) \to M(m, n; \mathbb{K})$ ist auch linear, also wegen der Bijektivität sogar ein *Isomorphismus.*

Wir zeigen nun, dass der Isomorphismus aus 4.13 darüberhinaus die Eigenschaft hat, Verknüpfungen linearer Abbildungen in Produkte von Matrizen überzuführen.

4.14. Satz: Seien B, C, D Basen in den endlichdimensionalen $\mathbb{K}$-Vektorräumen V, W, U und $\varphi\colon V \to W$ sowie $\psi\colon W \to U$ linear. Dann gilt:

$$[\psi \circ \varphi]_{B,D} = [\psi]_{C,D} \cdot [\varphi]_{B,C}.$$

Beweis: Wir setzen $T = [\varphi]_{B,C}$ und $S = [\psi]_{C,D}$ und betrachten das kommutative Diagramm

$$\begin{tikzcd}
V \arrow[r, "\varphi"] \arrow[d, "\Phi_B"'] \arrow[rr, bend left, "\psi\circ\varphi"] & W \arrow[r, "\psi"] \arrow[d, "\Phi_C"'] & U \arrow[d, "\Phi_D"] \\
\mathbb{K}^n \arrow[r, "T"'] \arrow[rr, bend right, "S\cdot T"', "x\mapsto S\cdot(T\cdot x)=(S\cdot T)\cdot x"' below] & \mathbb{K}^m \arrow[r, "S"'] & \mathbb{K}^p
\end{tikzcd}$$

dessen äußere Teile zeigen, dass die Matrix von $\psi \circ \varphi$ bzgl. B und D gerade $S \cdot T$ ist. $\square$

4.15. Beispiel: Für $V = W = P_3(\mathbb{R})$ wollen wir die lineare Abbildung $\varphi(p) := p - p'$ bzgl. der Basen $B = C = \{1, x, x^2, x^3\}$ als Matrix darstellen. (Linearität von φ folgt aus kurzer Rechnung.)

Wir berechnen die Bilder der Basiselemente und deren Koordinatenvektoren bgl. C:

$$\varphi(1) = 1 - 0 = 1 \quad \text{mit Koordinatenvektor} \quad \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \varphi(x) = x - 1 = -1 + x \quad \text{mit} \quad \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

$$\varphi(x^2) = x^2 - 2x = -2x + x^2 \quad \text{mit} \quad \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \quad \varphi(x^3) = x^3 - 3x^2 = -3x^2 + x^3 \quad \text{mit} \quad \begin{pmatrix} 0 \\ 0 \\ -3 \\ 1 \end{pmatrix},$$

daher erhalten wir

$$A = [\varphi]_{B,C} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Vorteile dieser Darstellung: Probleme über φ sind auf Matrizenprobleme zurückgeführt, z.B.: Sei $r \in P_3(\mathbb{R})$ gegeben. Wir betrachten die Gleichung $\varphi(p) = r$ für $p \in P_3(\mathbb{R})$, d.h. $p - p' = r$.

Wir fragen uns: Existiert immer eine Lösung? (D.h.: Ist φ surjektiv?)
Wenn ja: Ist sie eindeutig? (D.h.: Ist φ injektiv?)

Aus späteren Sätzen werden wir sofort sehen, dass die Abbildung $x \mapsto A \cdot x$ bijektiv ist, somit ist auch φ bijektiv und die Antwort heißt beide Male JA! (Siehe 4.52.2.)

Wir verwandeln auf diese Art also abstrakte endlichdimensionale Probleme in konkrete $\mathbb{K}^n$-Probleme! Nochmals für obiges Beispiel in Diagrammform:

$$\begin{tikzcd} P_3(\mathbb{R}) \arrow[r, "\varphi"] \arrow[d, "\Phi_B"'] & P_3(\mathbb{R}) \arrow[d, "\Phi_B"] \\ \mathbb{R}^4 \arrow[r, "A"'] & \mathbb{R}^4 \end{tikzcd}$$

Abstrakt ist $\varphi(p) = p - p'$ für jedes $p \in P_3(\mathbb{R})$. In der üblichen Polynomschreibweise $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$, die hier der Basisentwicklung bzgl. $B = C$ entspricht, erhalten wir

$$\begin{aligned} \varphi(a_0 + a_1 x + a_2 x^2 + a_3 x^3) &= (a_0 + a_1 x + a_2 x^2 + a_3 x^3) - (a_1 + 2a_2 x + 3a_3 x^2) \\ &= (a_0 - a_1) + (a_1 - 2a_2)x + (a_2 - 3a_3)x^2 + a_3 x^3. \end{aligned}$$

An den Koeffizienten lässt sich hier schon die Wirkung der Matrix $A = [\varphi]_{B,C}$ ablesen:

$$[\varphi(p)]_C = [\varphi]_{B,C} \cdot [p]_B = A \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_0 - a_1 \\ a_1 - 2a_2 \\ a_2 - 3a_3 \\ a_3 \end{pmatrix}.$$

Kurz gefasst:

$$\begin{tikzcd} p \arrow[r, mapsto, "\varphi"] \arrow[d, mapsto, "\Phi_B"'] & p - p' \arrow[d, mapsto, "\Phi_C"] \\ \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} \arrow[r, mapsto, "A"'] & \begin{pmatrix} a_0 - a_1 \\ a_1 - 2a_2 \\ a_2 - 3a_3 \\ a_3 \end{pmatrix} \end{tikzcd}$$

4.16. Spezialfall: $V = \mathbb{K}^n$, $W = \mathbb{K}^m$, B und C jeweils die Standardbasen und $\varphi = \varphi_A \colon x \mapsto A \cdot x$, dann ist die Matrix von φ_A bzgl. der Standardbasen gerade A selbst, weil jeder Vektor sein eigener Koordinatenvektor ist (vgl. 4.10.2).

4.17. Beispiele: 1. Nullabbildung: unabhängig von Basen ist $\varphi(b_j) = 0$ für $\varphi = 0 \colon V \to W$ und daher $[0]_{B,C} = 0$ (Nullmatrix).

2. Identische Abbildung: mit $B = C$ ergibt sich $\varphi(b_j) = b_j = 0b_1 + \ldots + 1b_j + \ldots 0b_n$ für $\varphi = \mathrm{id}_V \colon V \to V$ und daher $[\mathrm{id}_V]_{B,B} = I$.

Kann eine lineare Abbildung die Dimension verkleinern oder vergrößern?

4.18. Definition: Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen $\mathbb{K}$-VR, dann nennen wir

$$\ker\varphi := \{v \in V \mid \varphi(v) = 0\} = \varphi^{-1}(\{0\}) \subseteq V$$

den *Kern von* φ und

$$\operatorname{im}\varphi := \{w \in W \mid \exists v \in V : w = \varphi(v)\} = \varphi(V) \subseteq W$$

das *Image oder Bild von* φ.

Die Menge $\ker\varphi$ ist also die *Lösungsmenge* der homogenen linearen Gleichung $\varphi(v) = 0$, während $\operatorname{im}\varphi$ aus allen Vektoren in W besteht, die als Bilder unter φ von Vektoren aus V auftreten können, d.h. die als rechte Seite in der linearen Gleichung $\varphi(v) = w$ zu einer *lösbaren* Gleichung führen.

4.19. Proposition: Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen $\mathbb{K}$-VR. Dann ist $\ker\varphi$ ein Teilraum von V und $\operatorname{im}\varphi$ ein Teilraum von W.

Beweis als Übung.

4.20. Beispiele: 1. $\varphi\colon P_3(\mathbb{R}) \to P_3(\mathbb{R})$, $\varphi(p) = p'$ hat

$\ker\varphi = \{p \mid p' = 0\} = \{p \mid p \text{ konstant}\} = \{p \mid p(x) = a_0\} = P_0(\mathbb{R})$ und

$$\begin{aligned}\operatorname{im}\varphi = \{p \mid p = q' \text{ mit } q \in P_3(\mathbb{R})\} &= \{a_1 + 2a_2x + 3a_3x^2 \mid a_1, a_2, a_3 \in \mathbb{R}\}\\ &= \{b_0 + b_1x + b_2x^2 \mid b_0, b_1, b_2 \in \mathbb{R}\} = P_2(\mathbb{R}).\end{aligned}$$

2. $\varphi\colon C[a,b] \to \mathbb{R}$, $\varphi(f) = \int\limits_a^b f(t)\,dt$ hat

$\ker\varphi = \{f \mid \int\limits_a^b f(t)\,dt = 0\} = \{f \mid \text{ Fläche über der } x\text{-Achse} = \text{Fläche unter der } x\text{-Achse}\}$ und

(im Falle $a < b$) $\operatorname{im}\varphi = \mathbb{R}$, weil z.B. für jede reelle Zahl λ gilt: $\int\limits_a^b \frac{\lambda}{b-a}\,dt = \lambda$.

3. Sei $\varphi\colon \mathbb{R}^3 \to \mathbb{R}$ linear, d.h. $\varphi = \varphi_A$ mit $A = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}$,

$$\varphi(x) = A \cdot x = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = a_1x_1 + a_2x_2 + a_3x_3.$$

Im Falle $\begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$ ist

$\ker\varphi = \{x \mid a_1x_1 + a_2x_2 + a_3x_3 = 0\}$, also eine *Ebene durch* 0 mit Normalvektor $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$.

4. Sei $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ linear, d.h. $\varphi = \varphi_A$ mit $A \in M(m,n;\mathbb{K})$, dann ist

$\ker\varphi = \{x \mid A\cdot x = 0\}$ = *Lösungsraum des homogenen linearen Gleichungssystems* $A\cdot x = 0$.

5. Sei φ wie in 4.; für welche $b \in \mathbb{K}^m$ ist $A\cdot x = b$ lösbar?

Genau für jene, für die es ein $x \in \mathbb{K}^n$ mit $\varphi_A(x) = A\cdot x = b$ gibt, d.h.

$\operatorname{im}\varphi = \{b \mid A\cdot x = b \text{ ist lösbar }\}$.

4.21. Proposition: Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen $\mathbb{K}$-VR, dann gilt:

(i) φ ist injektiv $\iff$ $\ker\varphi = \{0\}$

(ii) φ ist surjektiv $\iff$ $\operatorname{im}\varphi = W$

Beweis: (i) $\Rightarrow$: Aus $v \in \ker\varphi$ folgt $\varphi(v) = 0 = \varphi(0)$, daher wegen der Injektivität $v = 0$. (Somit $\ker\varphi \subseteq \{0\}$; aber $\ker\varphi \supseteq \{0\}$ ist ohnehin klar, weil $\ker\varphi$ ein Teilraum ist.)
$\Leftarrow$: $\varphi(v_1) = \varphi(v_2) \;\Rightarrow\; \varphi(v_1 - v_2) = \varphi(v_1) - \varphi(v_2) = 0 \;\Rightarrow\; v_1 - v_2 \in \ker\varphi = \{0\} \;\Rightarrow\; v_1 - v_2 = 0$.

(ii) ist klar nach Definition von $\operatorname{im}\varphi$. □

4.22. Proposition: Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen $\mathbb{K}$-VR, dann gilt:

(i) $\varphi(v_1),\dots,\varphi(v_k)$ linear unabhängig $\implies$ $v_1,\dots,v_k$ linear unabhängig

(ii) Falls φ injektiv ist, gilt auch
$v_1,\dots,v_k$ linear unabhängig $\implies$ $\varphi(v_1),\dots,\varphi(v_k)$ linear unabhängig.

Beweis: (i) Angenommen $v_1,\dots,v_k$ sind linear abhängig, d.h. es gibt eine nichttriviale Linearkombination $\sum\lambda_i v_i = 0$, dann können wir φ auf die Gleichung anwenden und erhalten $\sum\lambda_i\varphi(v_i) = \varphi(\sum\lambda_i v_i) = \varphi(0) = 0$. Somit wären aber $\varphi(v_1),\dots,\varphi(v_k)$ linear abhängig — ein WIDERSPRUCH!

(ii) Sei $\sum\lambda_i\varphi(v_i) = 0$, dann folgt $\varphi(\sum\lambda_i v_i) = \sum\lambda_i\varphi(v_i) = 0$, also $\sum\lambda_i v_i \in \ker\varphi$. Gemäß 4.21(i) erzwingt dies aber $\sum\lambda_i v_i = 0$. Da $v_1,\dots,v_k$ linear unabhängig sind, müssen weiters alle $\lambda_i = 0$ sein. Somit ist die lineare Unabhängigkeit von $\varphi(v_1),\dots,\varphi(v_k)$ gezeigt. □

4.23. Korollar: Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen $\mathbb{K}$-VR, dann gilt:

(i) $\dim\operatorname{im}\varphi \le \dim V$ [Bild hat höchstens die Dimension des Definitionsbereiches]

(ii) φ injektiv $\implies$ $\dim\operatorname{im}\varphi = \dim V$

Beweis: (i) Sind $w_1 = \varphi(v_1),\dots,w_k = \varphi(v_k)$ linear unabhängig in $\operatorname{im}\varphi$, so sind nach 4.22(i) auch $v_1,\dots,v_k$ linear unabhängig in V. Daher hat V mindestens so lange linear unabhängige Systeme wie $\operatorname{im}\varphi$.

(ii) Erinnere: Die Dimension eines VR ist die maximale Länge linear unabhängiger Systeme (Zusatz zu Definition 3.23). Nach (i) hat $\operatorname{im}\varphi$ immer höchstens so lange l.u. Systeme wie V, nach 4.22(ii) wegen der Injektivität aber auch mindestens so lange l.u. Systeme wie V; also muss Gleichheit der Dimensionen gelten. $\square$

Bemerkung: Für $\dim V < \infty$ folgt die Umkehrung von (ii)

$$\dim \operatorname{im}\varphi = \dim V \implies \varphi \text{ injektiv}$$

unmittelbar aus dem folgenden Satz.

4.24. Satz: Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen endlichdimensionalen $\mathbb{K}$-VR, dann gilt:

$$\dim\ker\varphi + \dim\operatorname{im}\varphi = \dim V.$$

Beweis: Sei $w_1 = \varphi(v_1), \dots, w_r = \varphi(v_r)$ eine Basis von $\operatorname{im}\varphi$ (und somit $\dim\operatorname{im}\varphi = r$). Nach 4.22(i) ist $v_1, \dots, v_r$ linear unabhängig.

Sei weiters $u_1, \dots, u_l$ eine Basis von $\ker\varphi$ (und somit $\dim\ker\varphi = l$).

Behauptung: Das System $S\colon v_1, \dots, v_r, u_1, \dots, u_l$ ist eine Basis von V (und somit $\dim V = r + l$, sodass dann der Beweis fertig ist).

Lineare Unabhängigkeit von S: Es gelte $\sum \lambda_i v_i + \sum \mu_j u_j = 0$. Wir wenden φ auf die Gleichung an und erhalten

$$\sum \lambda_i \underbrace{\varphi(v_i)}_{w_i} + \sum \mu_j \underbrace{\varphi(u_j)}_{0} = 0.$$

Wegen der linearen Unabhängigkeit der w_i müssen alle $\lambda_i = 0$ sein, und daraus in Folge auch alle $\mu_j = 0$, weil auch die u_j linear unabhängig sind.

S ist Erzeugendensystem: Sei $v \in V$, dann haben wir mit geeigneten Skalaren λ_i die Darstellung $\varphi(v) = \sum \lambda_i w_i = \sum \lambda_i \varphi(v_i) = \varphi(\sum \lambda_i v_i)$. Daraus folgt $\varphi(v - \sum \lambda_i v_i) = 0$, d.h. $v - \sum \lambda_i v_i \in \ker\varphi$. Somit gibt es Skalare μ_j, sodass $v - \sum \lambda_i v_i = \sum \mu_j u_j$ gilt. Dies ergibt die Darstellung $v = \sum \lambda_i v_i + \sum \mu_j u_j$. $\square$

4.25. Korollar: Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen endlichdimensionalen $\mathbb{K}$-VR mit $\dim V = \dim W$, dann sind folgende Aussagen äquivalent:

(i) φ ist bijektiv (d.h. ein Isomorphismus)

(ii) φ ist injektiv

(iii) φ ist surjektiv

Beweis: (i) $\Rightarrow$ (ii) ist klar.

(ii) $\Rightarrow$ (iii): φ injektiv $\overset{4.21}{\Rightarrow}$ $\dim\ker\varphi$ $\overset{4.24}{\Rightarrow}$ $\dim V = \dim\operatorname{im}\varphi \overset{3.28}{\leq} \dim W = \dim V \Rightarrow$
$\dim\operatorname{im}\varphi = \dim W \Rightarrow \operatorname{im}\varphi = W$, d.h. φ surjektiv.

(iii) $\Rightarrow$ (i): φ surjektiv $\Rightarrow$ $\dim \operatorname{im} \varphi = \dim W = \dim V \overset{4.24}{\Rightarrow} \dim \ker \varphi = 0 \Rightarrow$
$\ker \varphi = \{0\} \overset{4.21}{\Rightarrow} \varphi$ injektiv. $\square$

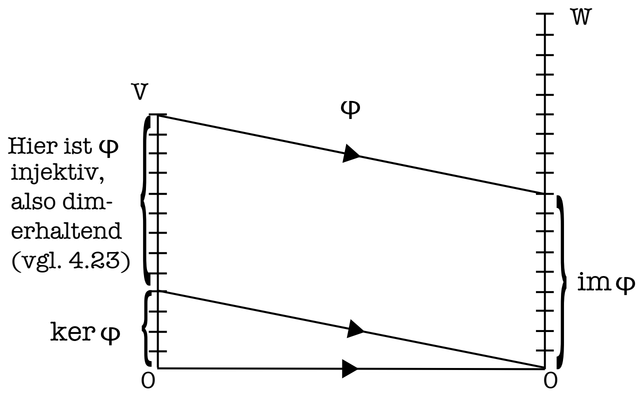
4.26. Satz: Es seien V und W endlichdimensionale $\mathbb{K}$-VR, dann gilt:

$$V \text{ und } W \text{ sind isomorph} \quad \Longleftrightarrow \quad \dim V = \dim W$$

Beweis: $\Rightarrow$: Sei $\varphi\colon V \to W$ ein Isomorphismus. Aus der Surjektivität von φ folgt $\dim W = \dim \operatorname{im} \varphi$ und aus der Injektivität auch $\dim \operatorname{im} \varphi = \dim V$, insgesamt daher $\dim W = \dim V$.

$\Leftarrow$: Seien B, C Basen von V, W und $\dim V = \dim W = n$. Betrachte $\varphi := \Phi_B \circ \Phi_C^{-1}\colon V \to \mathbb{K}^n \to W$. Sowohl Φ_B als auch Φ_C^{-1} ist linear und bijektiv, daher ist es auch deren Verknüpfung φ. Also ist φ ein Isomorphismus von V mit W. $\square$

4.27. Eine schematische Illustration zu 4.24 können wir versuchen, indem wir die Dimension als entsprechende Strecke auftragen und einfach pro Basisvektor einen Teilstrich machen:



4.28. Beispiel: Wir hatten in 4.5.2 die Projektion $\varphi\colon \mathbb{R}^3 \to \mathbb{R}^2$ auf die x-y-Ebene, die durch $\varphi = \varphi_A$ mit Matrix (bzgl. der Standardbasen)

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

gegeben ist. Wir berechnen Kern und Bild:

$\ker \varphi = \{x \in \mathbb{R}^3 \mid A \cdot x = 0\}$: das Gauß-Schema ist $\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}$, daher $x_3 = t,\ x_2 = -t,\ x_1 = -t$,

d.h. $\ker \varphi = \{t \cdot \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \mid t \in \mathbb{R}\}$ und hat Dimension 1 (Gerade durch 0).

$\operatorname{im} \varphi = \{A \cdot x \mid x \in \mathbb{R}^3\} = \mathbb{R}^2$, weil für beliebige $x_1, x_2 \in \mathbb{R}$ gilt, dass $A \cdot \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

(Die Projektion wirkt auf der x-y-Ebene wie die Identität.)

4.29. Invertierbare Matrizen - Motivation: Wenn wir ein lineares Gleichungssystem $A\cdot x = b$ vorliegen haben, dann würden wir im besten Fall und in Analogie zu Rechenregeln für reelle Zahlen (oder allgemeine Körper) erwarten, eine Lösung in der Form „$x = A^{-1}\cdot b$" zu erhalten.

Betrachten wir zwei durch Matrizen gegebene lineare Abbildungen $\varphi_A\colon x\mapsto A\cdot x$ und $\varphi_B\colon y\mapsto B\cdot y$, die invers zueinander sind, d.h.

$$B\cdot A\cdot x = x \quad \text{und} \quad A\cdot B\cdot y = y \quad \forall x,y,$$

d.h. $\boxed{B\cdot A = I_n}$ und $\boxed{A\cdot B = I_m}$, wobei $A\in M(m,n;\mathbb{K})$ und $B\in M(n,m;\mathbb{K})$. Dann ist $\varphi_A\colon \mathbb{K}^n\to\mathbb{K}^m$ linear und umkehrbar, also bijektiv, und somit gilt gemäß 4.26 die Gleichheit $\boxed{n=m}$. (Matrizen mit Zeilenanzahl = Spaltenanzahl nennen wir auch „quadratisch".)

4.30. Definition: Eine quadratische Matrix $A\in M(n,n;\mathbb{K})$ ist *invertierbar*, wenn es eine quadratische Matrix $B\in M(n,n;\mathbb{K})$ gibt, sodass gilt:

$$A\cdot B = I_n \quad \text{und} \quad B\cdot A = I_n.$$

In dem Fall wird $A^{-1} := B$ die *zu A inverse Matrix* genannt.

Die Inverse A^{-1} ist eindeutig bestimmt: Sind nämlich B und C beide invers zu A, dann folgt $B = B\cdot I_n = B\cdot(A\cdot C) = (B\cdot A)\cdot C = I_n\cdot C = C$.

4.31. Beispiele: 1. Die Nullmatrix hat nie eine Inverse, denn $0\cdot B = B\cdot 0 = 0\neq I$.

2. Die Einheitsmatrix ist zu sich selbst invers: $I\cdot I = I$.

3. $\begin{pmatrix} a_{11} & 0 & \cdots & 0\\ 0 & a_{22} & \cdots & 0\\ 0 & & \ddots & 0\\ 0 & \cdots & 0 & a_{nn}\end{pmatrix}$ hat $\begin{pmatrix} \frac{1}{a_{11}} & 0 & \cdots & 0\\ 0 & \frac{1}{a_{22}} & \cdots & 0\\ 0 & & \ddots & 0\\ 0 & \cdots & 0 & \frac{1}{a_{nn}}\end{pmatrix}$ als Inverse, falls alle $a_{ii}\neq 0$ sind. (Übung.)

4. $\begin{pmatrix} a & b\\ c & d\end{pmatrix}$ hat $\frac{1}{ad-bc}\begin{pmatrix} d & -b\\ -c & a\end{pmatrix}$ als Inverse, falls $ad-bc\neq 0$; andernfalls gibt es keine Inverse.
Ersteres kann natürlich einfach nachgerechnet werden, zweiteres ist mit ein bisschen mehr Theorie sehr einfach (siehe später), sonst ein wenig mühsam nachzuprüfen. Zum Beispiel ist $\left(\begin{smallmatrix} 1 & 2\\ -2 & -4\end{smallmatrix}\right)$ nicht invertierbar.

5. $\begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}\cdot\begin{pmatrix} 1 & 0\\ 0 & 1\\ 0 & 0\end{pmatrix} = \begin{pmatrix} 1 & 0\\ 0 & 1\end{pmatrix} = I$ und $\begin{pmatrix} 1 & 0\\ 0 & 1\\ 0 & 0\end{pmatrix}\cdot\begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix} = \begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}\neq I,$
d.h. aus $A\cdot B = I$ folgt nicht $B\cdot A = I$ bei rechteckigen Matrizen. Wie wir in 4.29 gesehen haben, müssen für die Gültigkeit von $A\cdot B = I$ und $B\cdot A = I$ ohnehin beide Matrizen quadratisch sein.

4.32. Proposition: Seien $A,B\in M(n,n;\mathbb{K})$, dann gilt:

$$A\cdot B = I \quad\Longrightarrow\quad B\cdot A = I.$$

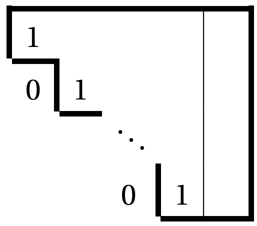
Beweis: $A \cdot B = I \;\Rightarrow\; \varphi_A \circ \varphi_B = \varphi_I = \mathrm{id}\colon \mathbb{K}^n \xrightarrow{\varphi_B} \mathbb{K}^n \xrightarrow{\varphi_A} \mathbb{K}^n \;\Rightarrow\; \varphi_A$ surjektiv $\overset{4.25}{\Rightarrow}$ φ_A bijektiv $\Rightarrow\; \varphi_B = (\varphi_A^{-1} \circ \varphi_A) \circ \varphi_B = \varphi_A^{-1} \circ (\varphi_A \circ \varphi_B) = \varphi_A^{-1} \;\Rightarrow\; \varphi_B \circ \varphi_A = \varphi_A^{-1} \circ \varphi_A = \mathrm{id} \;\Rightarrow\; B \cdot A = I$ □

4.33. Verfahren zur Berechnung der inversen Matrix: „Gauß verlängert“.

Gegeben sei $A = (a_{ij}) \in M(n,n;\mathbb{K})$ und gesucht sei $B = (b_{ij}) \in M(n,n;\mathbb{K})$, sodass $A \cdot B = I$ gilt. Wir teilen B und I in Spalten: $B = \left(\boxed{b_1} \;\cdots\; \boxed{b_n}\right)$, $I = \left(\boxed{e_1} \;\cdots\; \boxed{e_n}\right)$, dann bedeutet $A \cdot B = I$ gerade

$A \cdot b_1 = e_1, \dots, A \cdot b_n = e_n,$ das sind n lineare Gleichungssysteme für die gesuchten $b_1, \dots, b_n$.

Betrachten wir zunächst die Gauß-Elimination für das erste Gleichungssystem $A \cdot b_1 = e_1$: Ausgehend vom Schema $\boxed{A \mid e_1}$ erreichen wir wie üblich die Zeilenstufenform, wobei wir bei invertierbarem A eine obere Dreiecksform bekommen und nach eventueller Division einzelner Gleichungen entlang der Diagonalen jeweils 1 stehen haben:



Nun können wir aber auch oberhalb der Diagonale überall Nullen erzeugen, indem wir sukzessive passende Vielfache der untersten, zweituntersten, … Zeile zu den oberen Zeilen dazuzählen, bis wir auf die folgende Form kommen:

$$\begin{array}{cccc|c} 1 & 0 & \dots & 0 & \\ 0 & 1 & & \vdots & \\ \vdots & & \ddots & 0 & \\ 0 & \dots & 0 & 1 & \end{array}$$

Hier steht links die Identität I_n, daher muss rechts auch schon der Lösungsvektor b_1 stehen, weil die Variablenwerte sich nun unmittelbar aus den Einträgen der rechten Seite ergeben.

Alle Gleichungssysteme $A \cdot b_i$ $(i = 1, \dots, n)$ haben dieselbe Koeffizientenmatrix, daher sind auch dieselben Umformungen nötig; wir rechnen daher alles auf einen Schlag, indem wir alle n rechten Seiten zusammenfassen:

$\boxed{A \mid e_1\, e_2 \cdots e_n}$ führt mittels „verlängertem Gaußverfahren“ auf $\boxed{I \mid b_1\, b_2 \cdots b_n}$,

d.h. $\boxed{A \mid I}$ wird zu $\boxed{I \mid B}$ umgeformt, dann ist $A^{-1} = B$.

Beispiel: $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -4 \\ 1 & 1 & 2 \end{pmatrix}$, wir berechnen schrittweise

$$\left.\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -4 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right. \rightsquigarrow \left.\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -4 & 0 & 1 & 0 \\ 0 & -1 & 3 & -1 & 0 & 1 \end{array}\right. \rightsquigarrow \left.\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -4 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 & 1 & 1 \end{array}\right.$$

$$\rightsquigarrow \left.\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -4 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{array}\right. \rightsquigarrow \left.\begin{array}{ccc|ccc} 1 & 2 & 0 & 2 & -1 & -1 \\ 0 & 1 & 0 & 4 & -3 & -4 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{array}\right. \rightsquigarrow \left.\begin{array}{ccc|ccc} 1 & 0 & 0 & -6 & 5 & 7 \\ 0 & 1 & 0 & 4 & -3 & -4 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{array}\right.$$

und erhalten $A^{-1} = \begin{pmatrix} -6 & 5 & 7 \\ 4 & -3 & -4 \\ 1 & -1 & -1 \end{pmatrix}$.

Das System ist übrigens genau dann lösbar, wenn A eine Inverse besitzt.

4.34. Definition: Die Menge der invertierbaren $n \times n$-Matrizen wird mit $GL(n, \mathbb{K}) := \{A \in M(n, n; \mathbb{K}) \mid A \text{ invertierbar}\}$ bezeichnet und heißt auch „general linear group“.

4.35. $GL(n, \mathbb{K})$ bildet bzgl. der Matrizenmultiplikation eine *Gruppe*, d.h. es gilt:

1. $A, B \in GL(n, \mathbb{K}) \Rightarrow A \cdot B \in GL(n, \mathbb{K})$: Die Inverse zu $A \cdot B$ ist nämlich durch $B^{-1} \cdot A^{-1}$ gegeben, weil $(B^{-1} \cdot A^{-1}) \cdot A \cdot B = B^{-1} \cdot (A^{-1} \cdot A) \cdot B = B^{-1} \cdot I \cdot B^{-1} = I$ gilt; also ist $A \cdot B \in GL(n, \mathbb{K})$.

2. Assoziativität: Gilt gemäß 2.6.b)1.

3. Die Identität I gehört zu $GL(n, \mathbb{K})$ und ist neutrales Element: $I \cdot A = A \cdot I = A$.

4. Jedes Element $A \in GL(n, \mathbb{K})$ besitzt ein inverses Element in (!) $GL(n, \mathbb{K})$: A^{-1} ist invertierbar und hat als inverse Matrix A, weil $A \cdot A^{-1} = A^{-1} \cdot A = I$ gilt.

Für $n \geq 2$ ist die Gruppe $GL(n, \mathbb{K})$ nicht kommutativ: Zum Beispiel

$$A = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & 0 \\ 0 & & \cdots & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & 0 \\ 0 & & \cdots & 0 & 1 \end{pmatrix} \text{ sind invertierbar: } B^{-1} = \begin{pmatrix} 1 & -1 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & 0 \\ 0 & & \cdots & 0 & 1 \end{pmatrix}$$

und $A^{-1} = A$; also gilt $A, B \in GL(n, \mathbb{K})$.

Es ist $A\cdot B = \begin{pmatrix} 1 & 1 & \cdots & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & 0 \\ 0 & & \cdots & 0 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & -1 & \cdots & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & 0 \\ 0 & & \cdots & 0 & 1 \end{pmatrix} = B\cdot A.$

Bemerkung: $GL(1,\mathbb{K}) = \{A \in M(1,1,;\mathbb{K}) \mid A \text{ invertierbar}\} = \{\lambda \in \mathbb{K} \mid \lambda \neq 0\}$ ist kommutativ.

4.36. Satz: Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen endlichdimensionalen $\mathbb{K}$-VR, B eine Basis von V und C eine Basis von W. Dann gilt:

$$\varphi \text{ ist invertierbar} \iff [\varphi]_{B,C} \text{ ist invertierbar}$$

und in dem Fall gilt $\big([\varphi]_{B,C}\big)^{-1} = [\varphi^{-1}]_{B,C}$.

Beweis: $\Rightarrow$: $I = [\mathrm{id}_V]_{B,B} = [\varphi^{-1}\circ\varphi]_{B,B} \overset{4.14}{=} [\varphi^{-1}]_{C,B}\cdot[\varphi]_{B,C}$ und analog $I = [\varphi]_{B,C}\cdot[\varphi^{-1}]_{C,B}$.

$\Leftarrow$: Sei $A = [\varphi]_{B,C}$; wir definieren ähnlich wie in 4.12, Umkehrungsteil: $\psi := \Phi_B^{-1}\circ\varphi_{A^{-1}}\circ\Phi_C$. Dann ist $\psi\colon W \to V$ linear und es gilt

$$\psi\circ\varphi = \big(\Phi_B^{-1}\circ\varphi_{A^{-1}}\circ\Phi_C\big)\circ\big(\Phi_C^{-1}\circ\varphi_A\circ\Phi_B\big) = \Phi_B^{-1}\circ\varphi_{A^{-1}}\circ\varphi_A\circ\Phi_B = \mathrm{id}_V$$

und analog $\varphi\circ\psi = \mathrm{id}_W$. □

Bei den „Entlarvungen" von abstrakten Vektorräumen und abstrakten linearen Abbildungen hatten wir jeweils die Wahl von Basen frei. Wie verhalten sich die verschiedenen „Entlarvungen" zueinander, die man bei Verwendung verschiedener Basen enthält?

4.37. Basiswechsel in einem endlichdimensionalen Vektorraum V:

Seien $B = (b_1,\dots,b_n)$ und $B' = (b_1',\dots,b_n')$ Basen in V und sei $v \in V$.
Welcher Zusammenhang besteht zwischen $[v]_B$ und $[v]_{B'}$?

Sei $v = \sum\lambda_i b_i = \sum\mu_j b_j'$ und $b_1' = \sum_{i=1}^n a_{i1}b_i,\ \dots,\ b_n' = \sum_{i=1}^n a_{in}b_i$. Es ist

$$\sum_{i=1}^n \lambda_i b_i = v = \sum_{j=1}^n \mu_j b_j' = \sum_{j=1}^n \mu_j\left(\sum_{i=1}^n a_{ij}b_i\right) = \sum_{i=1}^n\left(\sum_{j=1}^n a_{ij}\mu_j\right)b_i,$$

daher wegen der Eindeutigkeit der Basisdarstellung $\lambda_i = \sum_{j=1}^n a_{ij}\mu_j$, d.h.

$$\underbrace{\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix}}_{[v]_B} = \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}}_{S}\cdot\underbrace{\begin{pmatrix}\mu_1\\ \vdots\\ \mu_n\end{pmatrix}}_{[v]_{B'}}, \qquad \underbrace{[v]_B}_{\text{„alt"}} = S\cdot\underbrace{[v]_{B'}}_{\text{„neu"}}.$$

Die Matrix S vermittelt also zwischen den Darstellungen von Vektoren bzgl. der Basen B und B'; wir nennen sie daher die *Matrix des Basiswechsels von B auf B'* und schreiben $S = [B \to B']$.

$$1.\ \text{Spalte von } S \;=\; \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} \;=\; \text{Koordinatenvektor von } b'_1 \text{ bzgl. } B \;=\; [b'_1]_B$$

usw. für die weiteren Spaltenvektoren von S: j. Spalte $= [b'_j]_B$.

Es werden also jeweils die Darstellungen der „neuen“ Basisvektoren durch die „alte“ Basis als Spaltenvektoren verwendet.

Ein Diagramm für den Basiswechsel sieht so aus

$$\begin{tikzcd}
\text{„alt“ } \mathbb{K}^n & \\
 & V \arrow[ul, "\Phi_B"'] \arrow[dl, "\Phi_{B'}"] \\
\text{„neu“ } \mathbb{K}^n \arrow[uu, "S"] &
\end{tikzcd}$$

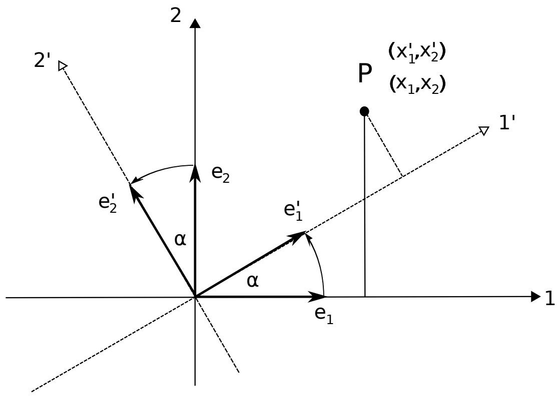
und eines für zwei Basiswechsel hintereinander $[B \to B''] = S \cdot T = [B \to B'] \cdot [B' \to B'']$ ist auch noch recht übersichtlich:

$$\begin{tikzcd}
\mathbb{K}^n & \\
\mathbb{K}^n \arrow[u, "S"] & V \arrow[ul, "\Phi_B"'] \arrow[l, "\Phi_{B'}"'] \arrow[dl, "\Phi_{B''}"] \\
\mathbb{K}^n \arrow[u, "T"] \arrow[uu, bend left, "S \cdot T"] &
\end{tikzcd}$$

Setzen wir speziell $B'' = B$ und beachten, dass $[B \to B]$ offensichtlich I ist, so folgt:
$I = [B \to B] = [B \to B'] \cdot [B' \to B]$, also ist S *stets invertierbar* und

$$S^{-1} = [B \to B']^{-1} = [B' \to B].$$

4.38. Beispiel: vgl. 4.5.3



Drehung um den Winkel α

„alte Basis“ $B = (e_1, e_2)$

„neue Basis“ $B' = (e'_1, e'_2)$

Einheitsvektoren („am Einheitskreis“)

$$e'_1 = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \quad e'_2 = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$$

Das heißt

$$e_1' = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} = \cos\alpha \cdot e_1 + \sin\alpha \cdot e_2, \quad \text{also} \quad [e_1']_B = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix},$$

$$e_2' = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix} = (-\sin\alpha) \cdot e_1 + \cos\alpha \cdot e_2, \quad \text{also} \quad [e_2']_B = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix},$$

und daher

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = S \cdot \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} \quad \text{mit} \quad S = \big([e_1']_B \;\; [e_2']_B\big) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Die inverse Matrix ist (gemäß 4.31.4)

$$S^{-1} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix},$$

daher $S^{-1} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = S^{-1} \cdot S \cdot \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}$ und somit

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

4.39. Basiswechsel bei linearen Abbildungen:

Seien V, W endlichdimensionale $\mathbb{K}$-VR mit $\dim V = n$, $\dim W = m$; seien B, B' Basen von V und C, C' Basen von W. Die zugehörigen Basiswechselmatrizen seien $S = [B \to B'] \in M(n, n; \mathbb{K})$ und $T = [C \to C'] \in M(m, m; \mathbb{K})$.

Für eine lineare Abbildung $\varphi\colon V \to W$ wollen wir eine Relation zwischen $A' := [\varphi]_{B',C'}$ und $A := [\varphi]_{B,C}$ herstellen. Dies erreichen wir sehr übersichtlich durch ein kommutatives Diagramm:

$$\begin{tikzcd}
\mathbb{K}^n \arrow[rrr, bend left, "A"] & & & \mathbb{K}^m \\
& V \arrow[ul, "\Phi_B"'] \arrow[dl, "\Phi_{B'}"] \arrow[r, "\varphi"'] & W \arrow[ur, "\Phi_C"] \arrow[dr, "\Phi_{C'}"'] & \\
\mathbb{K}^n \arrow[uu, "S"] \arrow[rrr, bend right, "A'"'] & & & \mathbb{K}^m \arrow[uu, "T"']
\end{tikzcd}$$

Wir lesen ab, dass $A' = T^{-1} \cdot A \cdot S$ gilt, d.h.

$$[\varphi]_{B',C'} = [C' \to C] \cdot [\varphi]_{B,C} \cdot [B \to B'].$$

Der Rest dieses Kapitels befasst sich mit dem sogenannten „Rang" einer Matrix bzw. einer linearen Abbildung. Aus dem Rang bestimmt sich die Zahl der freien Parameter eines linearen Gleichungssystems.

4.40. Definition: Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen den $\mathbb{K}$-VR V, W, dann definieren wir den *Rang von* φ durch

$$\operatorname{rg}\varphi := \dim \operatorname{im}\varphi.$$

4.41. Definition: Sei $A \in M(m, n; \mathbb{K})$, dann ist

$$\mathrm{ZR}(A) := \text{ Maximalzahl linear unabhängiger Zeilen von } A$$

der *Zeilenrang von A* und

$$\mathrm{SR}(A) := \text{ Maximalzahl linear unabhängiger Spalten von } A$$

der *Spaltenrang von A*.

In welchen Relationen stehen hier die obigen Rangbegriffe zueinander angesichts der Zusammenhänge zwischen linearen Abbildungen und Matrizen?

4.42. Definition: Sei V ein $\mathbb{K}$-VR und $S \subseteq V$, dann heißt

$$[S] := \{\sum_{i=1}^{n} \lambda_i v_i \mid n \in \mathbb{N}; 1 \le i \le n : \lambda_i \in \mathbb{K}, v_i \in S\}$$

das *lineare Erzeugnis von S* (auch „lineare Hülle“).

Bemerkungen: (i) $[S] \supseteq S$, denn $v \in S \Rightarrow v = 1 \cdot v \in [S]$.
(ii) $[S]$ ist ein Teilraum von V.
(iii) $[S]$ ist kleinstmöglich mit den Eigenschaften (i) und (ii), d.h. $[S]$ ist der kleinste Teilraum, der S umfasst: Ist nämlich W irgendein Teilraum, der S umfasst, dann enthält W alle Linearkombinationen wie in der Definition von $[S]$ angegeben und somit auch ganz $[S]$.
(iv) Nach Definition ist S ein Erzeugendensystem für $[S]$.
(v) Seien $v_1, \ldots, v_k \in V$; statt $[\{v_1, \ldots, v_k\}]$ schreiben wir oft auch einfach $[v_1, \ldots, v_k]$. Setzen wir $r := \dim[v_1, \ldots, v_k]$, dann ist r die Maximalzahl linear unabhängiger Vektoren unter den Vektoren $v_1, \ldots, v_k$: Verkürzt man nämlich das Erzeugendensystem $v_1, \ldots, v_k$ zu einer Basis von $[v_1, \ldots, v_k]$, dann verbleiben gerade r Vektoren; mehr als r linear unabhängige Vektoren kann es aber unter den $v_1, \ldots, v_k$ nach dem Zusatz zu 3.23 nicht geben.
(vi) Ist $A = \begin{pmatrix} \boxed{\quad z_1 \quad} \\ \vdots \\ \boxed{\quad z_m \quad} \end{pmatrix}$ mit Zeilenvektoren $z_1, \ldots, z_m$, so folgt $\mathrm{ZR}(A) = \dim[z_1, \ldots, z_m]$;

sind $s_1, \ldots, s_n$ Spaltenvektoren und $A = \begin{pmatrix} \boxed{s_1} & \cdots & \boxed{s_n} \end{pmatrix}$, dann ist $\mathrm{SR}(A) = \dim[s_1, \ldots, s_n]$.

4.43. Satz: (i) Für $A \in M(m,n;\mathbb{K})$ gilt $\operatorname{rg}\varphi_A = \operatorname{SR}(A)$.

(ii) Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen den endlichdimensionalen $\mathbb{K}$-VR V, W mit Basen B, C. Dann gilt $\operatorname{rg}\varphi = \operatorname{SR}([\varphi]_{B,C})$.

Beweis: (i): $\varphi_A(x) = A \cdot x = \begin{pmatrix} \boxed{s_1} & \cdots & \boxed{s_n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} = x_1 s_1 + \ldots + x_n s_n,$

daher ist $\operatorname{im}\varphi_A = \{\sum_{i=1}^n x_i s_i \mid x_i \in \mathbb{K}\} = [s_1, \ldots, s_n]$ und somit $\operatorname{rg}\varphi_A = \dim[s_1, \ldots, s_n] = \operatorname{SR}(A)$.

(ii): Sei $G := [\varphi]_{B,C}$, dann gilt:

$$\operatorname{rg}\varphi = \dim\operatorname{im}\varphi = \dim\operatorname{im}\Phi_C^{-1} \circ \varphi_G \circ \Phi_B = \dim\Phi_C^{-1}\big(\varphi_G\big(\overbrace{\Phi_B(V)}^{\mathbb{K}^n}\big)\big) \overset{[\Phi_C^{-1} \text{ injektiv; 4.23]}}{=}$$

$$= \dim\varphi_G(\mathbb{K}^n) = \dim\operatorname{im}\varphi_G = \operatorname{rg}\varphi_G \overset{(i)}{=} \operatorname{SR}(G) = \operatorname{SR}([\varphi]_{B,C}). \quad \square$$

Für die Anwendung auf lineare Gleichungssysteme benötigen wir noch die keineswegs selbstverständliche Tatsache, dass $\operatorname{ZR}(A) = \operatorname{SR}(A)$ gilt. Im Beweis verwenden wir die „Transposition“ einer Matrix, d.h. die Spiegelung an der „Hauptdiagonalen“, die aus Spalten Zeilen macht und umgekehrt.

4.44. Definition: Sei $A \in M(m,n;\mathbb{K})$, $A = (a_{ij})$. Die *Transponierte von* A ist die Matrix $A^t \in M(n,m;\mathbb{K})$, die durch

$$A^t := (b_{ij}), \quad b_{ij} := a_{ji} \quad (i = 1, \ldots, n; j = 1, \ldots, m)$$

gegeben ist.

4.45. Beispiele: 1. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

2. Die Transponierte eines Spaltenvektors ist der entsprechende Zeilenvektor und umgekehrt.

4.46. Proposition: (i) $(A^t)^t = A$,
(ii) $(A+B)^t = A^t + B^t$, $(\lambda A)^t = \lambda A^t$,
(iii) $(A \cdot B)^t = B^t \cdot A^t$,
(iv) Ist A invertierbar, so ist auch A^t invertierbar und es gilt: $(A^t)^{-1} = (A^{-1})^t$.

Beweis: (i) ist klar und (ii) ist sehr schnell klar gemacht.

(iii) $[(A \cdot B)^t]_{ij} = (A \cdot B)_{ji} = \sum_l A_{jl} B_{li} = \sum_l (A^t)_{lj} (B^t)_{il} = \sum_l (B^t)_{il} (A^t)_{lj} = (B^t \cdot A^t)_{ij}$.

(iv) $A \cdot A^{-1} = I \overset{\text{(ii)}}{\Rightarrow} (A^{-1})^t \cdot A^t = I^t = I$ und $A^{-1} \cdot A = I \overset{\text{(ii)}}{\Rightarrow} A^t \cdot (A^{-1})^t = I^t = I$,
daher ist A^t invertierbar und hat $(A^{-1})^t$ als Inverse. $\square$

4.47. Satz: Für jede Matrix $A \in M(m, n; \mathbb{K})$ gilt $\mathrm{ZR}(A) = \mathrm{SR}(A)$.

Beweis: Sei $A = (a_{ij}) = \begin{pmatrix} \boxed{\quad z_1 \quad} \\ \vdots \\ \boxed{\quad z_m \quad} \end{pmatrix} = \begin{pmatrix} \boxed{s_1} & \cdots & \boxed{s_n} \end{pmatrix}$ und $r := \mathrm{ZR}(A)$ sowie $W := [z_1, \dots, z_m]$.

Sei $b_1, \dots, b_r$ eine Basis von W (es ist ja $\dim W = r$). Wir definieren die $r \times n$-Matrix

$$(b_{kj}) := \begin{pmatrix} \boxed{\quad b_1 \quad} \\ \vdots \\ \boxed{\quad b_r \quad} \end{pmatrix}.$$

Jedes z_i ist Linearkombination der b_k, d.h. es gibt $c_{ik} \in \mathbb{K}$, sodass

$$(\star) \qquad z_i = \sum_{k=1}^{r} c_{ik} b_k \quad (i = 1, \dots, m).$$

Aus den Koeffizienten c_{ik} erhalten wir eine $m \times r$-Matrix, die wir in Spaltenvektoren organisieren

$$(c_{ik}) = \begin{pmatrix} \boxed{t_1} & \cdots & \boxed{t_r} \end{pmatrix}.$$

Wir hatten für die Zeilenvektoren $z_i = \begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix}$ und für die Spaltenvektoren $s_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$.

Daher heißt Gleichung $(\star)$ komponentenweise gelesen

$$a_{ij} = \sum_{k=1}^{r} c_{ik} b_{kj} \quad (i = 1, \dots, m; j = 1, \dots, n).$$

Wenn wir diese Gleichunge spaltenweise, d.h. jeweils für $i = 1, \dots, m$, zusammenfassen, erhalten wir die folgende Gleichung von Spaltenvektoren

$$s_j = \sum_{k=1}^{r} t_k b_{kj} \quad (j = 1, \dots, n).$$

Wir sehen somit, dass $t_1, \dots, t_r$ ein Erzeugendensystem für $[s_1, \dots, s_n]$ bildet, also ist

$$\mathrm{SR}(A) = \dim[s_1, \dots, s_n] \le r = \mathrm{ZR}(A).$$

Nun können wir analoge Überlegungen für A^t statt A anstellen und erhalten $\mathrm{SR}(A^t) \le \mathrm{ZR}(A^t)$. Klarerweise ist aber $\mathrm{SR}(A^t) = \mathrm{ZR}(A)$ und $\mathrm{ZR}(A^t) = \mathrm{SR}(A)$, sodass wir insgesamt die Ungleichungskette $\mathrm{ZR}(A) \le \mathrm{SR}(A) \le \mathrm{ZR}(A)$ und somit $\mathrm{SR}(A) = \mathrm{ZR}(A)$ gezeigt haben. □

4.48. Definition: Für $A \in M(m,n;\mathbb{K})$ definieren wir den *Rang von A* durch

$$\operatorname{rg} A := \operatorname{ZR}(A) = \operatorname{SR}(A).$$

Bemerkungen: 1. Es gilt natürlich stets $\operatorname{rg} A \leq \min(m,n)$.

2. Mit 4.47 und 4.48 lautet nun 4.43 so: (i) $\operatorname{rg} \varphi_A = \operatorname{rg} A$ und (ii) $\operatorname{rg} \varphi = \operatorname{rg}[\varphi]_{B,C}$.

Mit den in 2. erhaltenen Beziehung lassen sich nun die Aussagen in 4.21-4.25 in anderer Gestalt schreiben.

4.49. Seien V, W endlichdimensionale $\mathbb{K}$-VR, $\dim V = n$, $\dim W = m$ und $\varphi\colon V \to W$. Seien B, C Basen in V bzw. W und $A = [\varphi]_{B,C}$. Dann gilt:

(1) φ injektiv $\Leftrightarrow$ $\operatorname{rg} A = n$ (aus 4.23 Prop. und Bem.)

(2) φ surjektiv $\Leftrightarrow$ $\operatorname{rg} A = m$ (wegen $\operatorname{rg} A = \dim \operatorname{im} \varphi$)

(3) φ bijektiv $\Leftrightarrow$ $\operatorname{rg} A = n = m$ (Kombination der beiden vorhergehenden Aussagen)

(4) $\dim \ker \varphi + \operatorname{rg} A = n$ (aus 4.24)

(5) Ist $n = m$, dann sind folgende Aussagen äquivalent:

(i) φ bijektiv

(ii) φ injektiv

(iii) φ surjektiv

(iv) $\operatorname{rg} A = n$ (aus 4.25 und (3) oben)

4.50. Alle obigen Aussagen gelten natürlich erst recht im Spezialfall $V = \mathbb{K}^n$, $W = \mathbb{K}^m$, $B = C =$ Standardbasis, wobei wir statt φ in 4.49(1)-(5) jeweils φ_A lesen.

4.51. Satz: Sei $A \in M(n,n;\mathbb{K})$, dann gilt:

$$A \text{ ist invertierbar} \quad \Longleftrightarrow \quad \operatorname{rg} A = n \quad \text{(also größtmöglich)}.$$

Beweis: A invertierbar $\overset{4.36}{\Longleftrightarrow}$ φ_A invertierbar, d.h. bijektiv $\overset{4.49(5)}{\Longleftrightarrow}$ $\operatorname{rg} A = n$ □

4.52. Beispiele: 1. Für eine Matrix C, wie sie am Ende einer Gauß-Elimination für ein lineares Gleichungssystems entsteht, also von der folgenden Form (mit allen $c_{l j_l} \neq 0$)

$$C = \left(\begin{array}{cccccc} 0 & c_{1j_1} * * * * * & & & & \\ & 0 & c_{2j_2} * * * * & & & \\ & & 0 & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & 0 & c_{rj_r} * * * * * \\ & & & & & 0 \\ & & & & & 0 \\ & & & & & 0 \end{array}\right) \begin{array}{l} z_1 \\ z_2 \\ \vdots \\ \vdots \\ \vdots \\ z_r \\ z_{r+1} = 0 \\ \vdots \\ z_m = 0 \end{array}$$

sind die Zeilen $z_1, \dots, z_r$ linear unabhängig: Aus der Gleichung $\sum_{i=1}^{r} \lambda_i z_i = 0$ folgt nämlich der Reihe nach $\lambda_1 = 0$ (mit Blick auf die Komponente j_1), $\lambda_2 = 0$ (Komponente j_2) usw. Daher ist die Maximalzahl linear unabhängiger Zeilen gleich r und somit $\boxed{\operatorname{rg} C = r}$.

2. Für $\varphi\colon P_3(\mathbb{R}) \to P_3(\mathbb{R})$, $\varphi(p) = p' - p$ (vgl. 4.15) ist $\operatorname{rg}\varphi = 4$:

$$\operatorname{rg}\varphi = \operatorname{rg}[\varphi]_{B,B} = \operatorname{rg}\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 4$$ (Gauß-Schema!), daher ist φ bijektiv, was die beiden „Ja“ in 4.15 begründet.

4.53. Praktische Bestimmung des Ranges einer Matrix — wieder Gauß!

Die einzelnen Schritte im Zuge des Gauß-Eliminationsverfahrens lassen den Zeilenrang unverändert: Das ist klar beim Vertauschen zweier Zeilen. Wird nun etwa zur Zeile i das λ-fache der Zeile j addiert, also $z_i{}' = z_i + \lambda z_j$, so ist die lineare Hülle der neuen Zeilenvektoren sicherlich in der linearen Hülle der ursprünglichen Zeilenvektoren enthalten. Andererseits gilt wegen $z_i = z_i{}' - \lambda z_j$ auch die umgekehrte Inklusionsrelation der gerade genannten linearen Hüllen, sodass sie gleich sind und somit auch der Zeilenrang unverändert bleibt.

Wir erhalten also folgendes Rezept: Zur Bestimmung von $\operatorname{rg} A$ wendet man auf A das Verfahren der Gauß-Elimination an und liest am Ergebnis C den Rang wie oben in 4.52.1 ab. Zum Beispiel sehen wir nun mit einem Blick, dass die Koeffizientenmatrix für das Gleichungssystem in 1.4 den Rang 4 hat.

§5. LINEARE GLEICHUNGSSYSTEME II

Wir können nun mit Hilfe der bisher entwickelten Theorie die folgenden Themen kurz und schlagkräftig behandeln:

1. Lösbarkeit überhaupt
2. Form der Lösungsmenge:
   a) homogenes System
   b) inhomogenes System
3. Lösbarkeit für beliebige rechte Seite
4. Eindeutigkeit der Lösung
5. Eindeutige Lösung für beliebige rechte Seite

$\boxed{\text{ad 1.}}$ Sei $A = \left(\boxed{a_1} \cdots \boxed{a_n} \right)$, d.h. $A \cdot x = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} = x_1 a_1 + \ldots + x_n a_n.$

Dann haben wir folgende Kette von Äquivalenzen:

$$\begin{aligned} A \cdot x = b \text{ ist lösbar} \quad &\Longleftrightarrow \quad \exists x_1, \ldots, x_n : \quad x_1 a_1 + \ldots + x_n a_n = b \\ &\Longleftrightarrow \quad [a_1, \ldots, a_n] = [a_1, \ldots, a_n, b] \\ &\overset{3.28}{\Longleftrightarrow} \quad \dim[a_1, \ldots, a_n] = \dim[a_1, \ldots, a_n, b] \\ &\overset{4.48}{\Longleftrightarrow} \quad \operatorname{rg} A = \operatorname{rg} \underbrace{\left(\boxed{a_1} \cdots \boxed{a_n} \; \boxed{b} \right)}_{\substack{(A\ b)\text{...erweiterte Matrix} \\ \text{des Gleichungssystems}}} = \operatorname{rg}(A\ \ b) \end{aligned}$$

Somit erhalten wir die folgende Aussage:

5.1. Satz: $A \cdot x = b$ ist genau dann lösbar, wenn der Rang der erweiterten Matrix $(A\ \ b)$ mit dem Rang von A übereinstimmt.

$\boxed{\text{ad 2.a)}}$ Wir betrachten das Gleichungssystem $A \cdot x = 0$. Die zugehörige lineare Abbildung ist $\varphi_A \colon \mathbb{K}^n \to \mathbb{K}^m$, $x \mapsto A \cdot x$, d.h. der Lösungsraum des Gleichungssystems ist gerade $\ker \varphi_A$ und ist also ein Teilraum von $\mathbb{K}^n$, für den $\dim \ker \varphi_A = n - \dim \operatorname{im} \varphi_A = n - \operatorname{rg} A$ gilt. Wir haben daher die folgende Aussage gezeigt:

5.2. Satz: Der Lösungsraum eines homogenen linearen Gleichungssystems $A \cdot x = 0$ ist ein Teilraum von $\mathbb{K}^n$ der Dimension $n - \operatorname{rg} A$.

Der Lösungsraum kann daher z.B. durch Angabe von $n - \operatorname{rg} A$ linear unabhängigen Lösungsvektoren[7] beschrieben werden. Diese Basisvektoren werden dann mit $n - \operatorname{rg} A$ Koeffizienten — das sind die freien Parameter! — linear kombiniert. Vergleiche 4.52.1.: $\operatorname{rg} A = r$ und entsprechend in 1.3.4.: $n - r$ freie Parameter.

$\boxed{\text{ad 2.b)}}$ Das Gleichungssystem $A \cdot x = b$ sei lösbar. Wir nehmen irgendeine Lösung x_p (p für „partikulär") und halten sie fest. Wie sehen eventuelle andere Lösungen aus?

Ist x_a eine andere Lösung, d.h. $A \cdot x_\mathrm{a} = b$, dann folgt

$$A \cdot (x_\mathrm{a} - x_\mathrm{p}) = A \cdot x_\mathrm{a} - A \cdot x_\mathrm{p} = b - b = 0.$$

D.h. $y := x_\mathrm{a} - x_\mathrm{p}$ ist eine Lösung des homogenen Systems $A \cdot x = 0$.

Ist umgekehrt y eine Lösung von $A \cdot x = 0$ und $x_\mathrm{a} := x_\mathrm{p} + y$, dann folgt

$$A \cdot x_\mathrm{a} = A \cdot (x_\mathrm{p} + y) = A \cdot x_\mathrm{p} + A \cdot y = b + 0 = b.$$

D.h. x_a ist ebenfalls eine Lösung von $A \cdot x = b$. Wir fassen dies im folgenden Satz zusammen:

5.3. Satz: Sei $A \cdot x = b$ lösbar. Dann ist der Lösungsraum L dieses Systems ein affiner Teilraum des $\mathbb{K}^n$ von der Gestalt $L = x_\mathrm{p} + W$, wobei $W = \ker \varphi_A$ der Lösungsraum des homogenen Systems $A \cdot x = 0$ ist; weiters gilt $\dim L = \dim W = n - \operatorname{rg} A$.

$\boxed{\text{ad 3.}}$ $A \cdot x = b$ ist lösbar für jedes $b \in \mathbb{K}^m$ [„universell lösbar"] genau dann, wenn $\varphi_A \colon x \mapsto A \cdot x$ surjektiv $\mathbb{K}^n \to \mathbb{K}^m$ ist. Dies ist gleichbedeutend mit $\operatorname{im} \varphi_A = \mathbb{K}^m$, was wiederum (wegen 3.28) äquivalent zu $\dim \operatorname{im} \varphi_A = m$ ist. Nun lehrt uns 4.49(2), dass letzteres gerade $\operatorname{rg} A = m$ heißt, womit wir zur unten stehenden Aussage gelangen:

5.4. Satz: Für $A \in M(m, n; \mathbb{K})$ ist $A \cdot x = b$ genau dann universell lösbar, wenn $\operatorname{rg} A = m$ gilt.

$\boxed{\text{ad 4.}}$ $A \cdot x = b$ sei lösbar. Dann gilt:

$$\begin{aligned} A \cdot x = b \text{ ist eindeutig lösbar} \quad &\overset{5.3}{\iff} \quad L = x_\mathrm{p} + W \text{ besteht nur aus } \{x_\mathrm{p}\} \\ &\iff \quad W = \{0\} \\ &\iff \quad \dim W = 0 \\ &\overset{4.49(1)}{\iff} \quad \operatorname{rg} A = n. \end{aligned}$$

[7] Diese bilden dann automatisch eine Basis für den Lösungsraum.

Mit anderen Worten:

5.5. Satz: Für $A \in M(m,n;\mathbb{K})$ sei $A \cdot x = b$ lösbar. Die Lösung ist genau dann eindeutig, wenn $\operatorname{rg} A = n$ gilt.

$\boxed{\text{ad 5.}}$ $A \cdot x = b$ ist für jedes $b \in \mathbb{K}^m$ eindeutig lösbar

$$\begin{aligned}
&\overset{5.4,5.5}{\iff} && \operatorname{rg} A = n \text{ und } \operatorname{rg} A = m \\
&\iff && A \text{ ist } n \times n\text{-Matrix und } \operatorname{rg} A = n \\
&\overset{4.51}{\iff} && A \text{ ist invertierbar} \\
&\overset{6.11}{\iff} && \det A \neq 0.
\end{aligned}$$

ODER

$A \cdot x = b$ ist für jedes $b \in \mathbb{K}^m$ eindeutig lösbar

$$\begin{aligned}
&\iff && \varphi_A \colon x \mapsto A \cdot x \text{ ist bijektiv} \\
&\iff && \varphi_A \text{ ist invertierbar} \\
&\overset{4.29}{\iff} && A \text{ ist invertierbar} \\
&\overset{6.11}{\iff} && \det A \neq 0.
\end{aligned}$$

Somit gilt:

5.6. Satz: Sei $A \in M(m,n;\mathbb{K})$. Das Gleichungssystem $A \cdot x = b$ ist genau dann eindeutig lösbar für alle $b \in \mathbb{K}^m$, wenn A quadratisch ist (d.h. $m = n$) und eine der folgenden (für quadratische Matrizen) äquivalenten Bedingungen erfüllt:

(i) A hat größtmöglichen Rang, d.h. $\operatorname{rg} A = n$.

(ii) Die Zeilen von A sind linear unabhängig. [4.48]

(iii) Die Spalten von A sind linear unabhängig. [4.48]

(iv) Die Zeilen von A bilden eine Basis des $\mathbb{K}^n$. [3.27(v)]

(v) Die Spalten von A bilden eine Basis des $\mathbb{K}^n$. [3.27(v)]

(vi) A ist invertierbar.

(vii) $\det A \neq 0$.

(viii) $A \cdot x = 0$ hat nur die triviale Lösung $x = 0$. [5.2]

Es gilt in diesem Fall: $A \cdot x = b \iff x = A^{-1} \cdot b$.

Im Falle, dass A invertierbar ist ($\Leftrightarrow \det A \neq 0$), kann man anstelle des Verfahrens von Gauß auch die sogenannte *Cramersche Regel* zur Lösung von $A \cdot x = b$ verwenden:

5.7. Satz: Sei $A \in M(n,n;\mathbb{K})$ invertierbar und $b \in \mathbb{K}^n$. Dann hat die (eindeutig bestimmte) Lösung $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ von $A \cdot x = b$ die Gestalt

$$x_i = \frac{\det A_i}{\det A} \qquad \text{mit} \quad A_i := \underbrace{\begin{pmatrix} \boxed{a_1} & \cdots & \boxed{a_{i-1}} & \boxed{b} & \boxed{a_{i+1}} & \cdots & \boxed{a_n} \end{pmatrix}}_{i\text{-ter Spaltenvektor ersetzt durch } b} \qquad (i = 1, \ldots, n).$$

Beweise für die Cramersche Regel finden sich in den Standardlehrbüchern[8] und wir könnten im nächsten Kapitel ohne Mühe einen nachliefern, werden im Rahmen dieser Vorlesung aber darauf verzichten.

Die Cramersche Regel ist praktisch für $n = 2$, eventuell auch für $n = 3$. Für größere Systeme bzw. in numerischen Algorithmen ist sie für die praktische Berechnung der Lösung wenig geeignet, da sie viel mehr Rechenschritte erfordert als etwas die Gauß-Elimination. Allerdings hat sie eine wichtige theoretische Bedeutung: Als explizite Formel für x_i, in der rechts nur die mit den Grundrechnungsarten verknüpften Größen a_{ij}, b_j vorkommen, zeigt sie, dass eine kleine Änderung der a_{ij}, b_j auch nur eine kleine (d.h. nicht sprunghafte) Änderung der x_i bewirkt — „stetige Abhängigkeit der Lösung x von den Ausgangsdaten A und b“. Das ist wichtig, wenn A und b etwa gemessene Werte beinhalten und daher prinzipiell mit einem „Fehler“, also einer gewissen Unsicherheit behaftet sind.

DUALRÄUME

Für ein genaueres Studium linearer Gleichungssysteme ist es oft nötig, nicht nur das gesamte System (bzw. seine „linke Seite“) $\varphi_A\colon x \mapsto A \cdot x$ zu betrachten, sondern auch einzelne Gleichungen, d.h. in der Form $\varphi_a\colon x \mapsto a \cdot x$, wo a eine $(1 \times n)$-Matrix (also ein Zeilenvektor) ist:

$$a = (a_1, \ldots, a_n); \quad \varphi_a(x) := (a_1, \ldots, a_n) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + \ldots + a_n x_n.$$

Solche linearen Abbildungen $\varphi_a\colon \mathbb{K}^n \to \mathbb{K}$ bzw. allgemein $V \to \mathbb{K}$ erhalten eine eigene Bezeichnung, ebenso der Vektorraum $L(V,\mathbb{K})$ aller dieser linearen Abbildungen (siehe die folgende Definition). Unter anderem lässt sich mit der darauf aufbauenden Theorie die Aufgabe

[8] Z.B. in Fischers Büchern im Kapitel *Determinanten*, bei Muthsam im Kapitel *Lineare Abbildungen — Determinanten*, bei Jänich im Kapitel *Lineare Gleichungssysteme* und bei Bröcker im abstrakteren Kapitel *Ringe und Moduln*.

lösen, zu einem gegebenen Teilraum W des $\mathbb{K}^n$ lineare Gleichungssysteme $A \cdot x = 0$ zu konstruieren, deren Lösungsraum gerade W ist.

5.8. Definition: Sei V ein $\mathbb{K}$-Vektorraum. Ein *lineares Funktional* oder eine *Linearform auf* V ist eine lineare Abbildung $f\colon V \to \mathbb{K}$. Der Vektorraum $L(V,\mathbb{K})$ aller linearen Funktionale auf V wird mit V^* bezeichnet und *Dualraum von* V genannt, mit anderen Worten

$$V^* := \{f\colon V \to \mathbb{K} \mid f \text{ linear}\}.$$

5.9. Beispiele:

1. Der Dualraum von $\mathbb{K}^n$ ist $(\mathbb{K}^n)^* = L(\mathbb{K}^n,\mathbb{K}) \cong M(1,n;\mathbb{K})$, d.h. die Linearformen $f \in (\mathbb{K}^n)^*$ sind genau die Abbildungen der Gestalt

$$f_a(x) := (a_1,\ldots,a_n) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1x_1 + \ldots + a_nx_n \qquad (x \in \mathbb{K}^n).$$

2. Auf $V = P(\mathbb{R})$ sind die folgenden Abbildungen Beispiele für Linearformen $f, g \in (P(\mathbb{R}))^*$:

$$f\colon P(\mathbb{R}) \to \mathbb{R},\ f(p) := \frac{p''(0)}{2} \quad \text{und} \quad g\colon P(\mathbb{R}) \to \mathbb{R},\ g(p) := \int_0^1 p(t)\,dt$$

(Linearität ist leicht nachzurechnen.)

5.10. Duale Basis zu einer Basis

Sei V ein endlichdimensionaler $\mathbb{K}$-VR und $B = (b_1,\ldots,b_n)$ eine Basis von V.

(i) Wir definieren n lineare Funktionale $f_1,\ldots,f_n$ auf V wie folgt:

$$\begin{tikzcd} & \lambda_1 \in \mathbb{K} \\ V \ni v = \sum\limits_{i=1}^{n} \lambda_i b_i \arrow[ur, mapsto, "f_1"] \arrow[dr, mapsto, "f_n"'] & \vdots \\ & \lambda_n \in \mathbb{K} \end{tikzcd}$$

[d.h. $f_j = p_j \circ \Phi_B$, wobei $p_j\colon \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto \lambda_j$ die (lineare!) Projektion auf die j-te Achse ist].

(ii) Es gilt $f_j(b_i) = f_j(\underbrace{0 \cdot b_1 + \ldots + 1 \cdot b_i + \ldots + 0 \cdot b_n}_{\lambda_1=0,\ldots,\lambda_i=1,\ldots,\lambda_n=0}) = \left.\begin{cases} 1 & (j = i), \\ 0 & (j \neq i) \end{cases}\right\} = \delta_{ij}$.

(iii) Die Funktionale $f_1, \dots, f_n$ bilden eine Basis von V^*:

Lineare Unabhängigkeit: Seien $\mu_1, \dots, \mu_n \in \mathbb{K}$, sodass $\sum\limits_{j=1}^{n} \mu_j f_j = 0$ (links steht ein lineares Funktional und rechts die Nullabbildung: $V \to \mathbb{K}$);
d.h. $\forall v \in V$: $0 = \big(\sum \mu_j f_j\big)(v) = \sum \mu_j f_j(v)$.

Setzen wir speziell $v = b_i$ ein, so ergibt sich $0 = \sum\limits_{j=1}^{n} \mu_j \delta_{ij} = \mu_i$.

Erzeugendensystem: Sei $f \in V^*$ beliebig. Dann gilt für alle $v \in V$ mittels Basisdarstellung bzgl. B, dass

$$f(v) = f\Big(\sum_{i=1}^{n} \lambda_i b_i\Big) = \sum_{i=1}^{n} \lambda_i f(b_i) = \sum_{i=1}^{n} f_i(v) f(b_i) = \sum_{i=1}^{n} \underbrace{f(b_i)}_{\mu_j} f_i(v) = \left(\sum_{i=1}^{n} f(b_i) f_i\right)(v),$$

daher ist also $f = \sum\limits_{i=1}^{n} f(b_i) f_i \in [f_1, \dots, f_n]$.

Definition: $B^* := (f_1, \dots, f_n)$ heißt die zu $B = (b_1, \dots, b_n)$ *duale Basis;* sie ist durch die Gleichungen $f_j(b_i) = \delta_{ij}$ eindeutig bestimmt.

(iv) Die Situation ist im folgenden Sinne spiegelbildlich:

$$V \ni v = \sum_{i=1}^{n} \lambda_i b_i \;\begin{cases} \mapsto \lambda_1 = f_1(v) \\ \quad\vdots \\ \mapsto \lambda_n = f_n(v) \end{cases} \qquad V^* \ni f = \sum_{j=1}^{n} \mu_j f_j \;\begin{cases} \mapsto \mu_1 = f(b_1) \\ \quad\vdots \\ \mapsto \mu_n = f(b_n) \end{cases}$$

(v) Aus (iii) folgt $\dim V^* = n = \dim V$, daher gilt nach 4.26 auch $V \cong V^*$ (explizit via $\varphi\colon \sum \lambda_i b_i \mapsto \sum \lambda_i f_i$). Achtung! Letzteres ist falsch, wenn V unendlichdimensional ist.

5.11. Beispiele

1. Für $V = \mathbb{K}^n$ mit $B = (e_1, \dots, e_n)$ ist $B^* = (e_1^t, \dots, e_n^t)$ (mit $e_1^t = (1, 0, \dots, 0)$ usw.), denn

$$e_j^t \cdot e_i = \underbrace{(0, \dots, 1, \dots, 0)}_{1 \text{ an } j\text{-ter Stelle}} \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \delta_{ij}.$$

2. Sei $V = \mathbb{K}^n$ und $B = (b_1, \dots, b_n)$ eine beliebige Basis (von Spaltenvektoren); sei $B^* = (a_1, \dots, a_n)$ die zu B duale Basis (bestehend aus Zeilenvektoren), d.h. $a_i \cdot b_j = \delta_{ij}$. Dann folgt
$$\begin{pmatrix} \boxed{\quad a_1 \quad} \\ \vdots \\ \boxed{\quad a_n \quad} \end{pmatrix} \cdot \begin{pmatrix} \boxed{b_1} & \cdots & \boxed{b_n} \end{pmatrix} = (a_i \cdot b_j) = (\delta_{ij}) = I,$$
d.h. die a_i sind die Zeilen der inversen Matrix $\begin{pmatrix} \boxed{b_1} & \cdots & \boxed{b_n} \end{pmatrix}^{-1}$.

Die Berechnung der dualen Basis erfolgt hier also mittels Matrizenrechnung.

3. Für $V = P_3(\mathbb{R})$ mit Basis $B = (1, x, x^2, x^3)$ ist $B^* = (f_0, f_1, f_2, f_3)$, wobei für jedes $p \in P_3(\mathbb{R})$ gilt
$$f_0(p) := p(0), \quad f_1(p) := \frac{p'(0)}{1!}, \quad f_2(p) := \frac{p''(0)}{2!}, \quad f_3(p) := \frac{p'''(0)}{3!}.$$
(Rechnen Sie das nach!)

4. Aufgabe: Stelle $f \in (P_3(\mathbb{R}))^*$, gegeben durch $f(p) := \int\limits_0^1 p(t)\,dt$, als Linearkombination der Elemente der Dualbasis B^* aus 3. dar! (Verwende dafür 5.10(iii) bzw. (iv).)

Wir erhalten durch einfache Berechnung der Integrale $f(b_i)$ $(i = 0, 1, 2, 3)$
$$f = \sum_{i=0}^{3} f(b_i) f_i = 1 \cdot f_0 + \frac{1}{2} \cdot f_1 + \frac{1}{3} \cdot f_2 + \frac{1}{4} \cdot f_3,$$
d.h. bei Anwendung auf $p \in P_3(\mathbb{R})$
$$\int\limits_0^1 p(t)\,dt = \frac{p(0)}{1!} + \frac{p'(0)}{2!} + \frac{p''(0)}{3!} + \frac{p'''(0)}{4!}.$$
Achtung! Diese Formel gilt nur für Polynome vom Grad kleiner gleich 3.

5.12. Duale oder adjungierte lineare Abbildung zu einer gegebenen linearen Abbildung

Seien V und W Vektorräume über $\mathbb{K}$, $\varphi \colon V \to W$ linear und $g \in W^*$.

(i) $\begin{array}{ccc} V & \xrightarrow{\varphi} & W \\ & \underset{g\circ\varphi}{\searrow} & \downarrow g \\ & & \mathbb{K} \end{array}$ Die Verknüpfung $g \circ \varphi$ ist linear $V \to \mathbb{K}$, also ist $g \circ \varphi \in V^*$.

Die Abbildung $\varphi^*: g \mapsto g \circ \varphi$ ist linear $W^* \to V^*$: Es ist $\varphi^*(g_1+g_2) = (g_1+g_2)\circ\varphi \in V^*$, kann also auf $v \in V$ angewendet werden und ergibt

$$\big((g_1+g_2)\circ\varphi\big)(v) = (g_1+g_2)(\varphi(v)) = g_1(\varphi(v)) + g_2(\varphi(v)) = (g_1\circ\varphi)(v) + (g_2\circ\varphi)(v),$$

daher $\varphi^*(g_1+g_2) = (g_1+g_2)\circ\varphi = g_1\circ\varphi + g_2\circ\varphi = \varphi^*(g_1)+\varphi^*(g_2)$.
Analog zeigt man $\varphi^*(\lambda g) = \lambda\varphi^*(g)$.

Definition: Die lineare Abbildung $\varphi^*: W^* \to V^*$, $g \mapsto g\circ\varphi$ heißt die *zu φ adjungierte (oder duale) lineare Abbildung.*

(ii) **Matrix der adjungierten linearen Abbildung:** Seien V und W endlichdimensional, $\dim V = n$, $\dim W = m$ und $B = (b_1,\dots,b_n)$, $C = (c_1,\dots,c_m)$ Basen von V, W. Seien $B^* = (f_1,\dots,f_n)$ bzw. $C^* = (g_1,\dots,g_m)$ die entsprechenden Dualbasen in V^* bzw. W^*.

Es sei $A = (a_{ij}) = [\varphi]_{B,C} \in M(m,n;\mathbb{K})$ und $B = (b_{ij}) = [\varphi^*]_{C^*,B^*} \in M(n,m;\mathbb{K})$.

Welcher Zusammenhang besteht zwischen A und B?

$A = [\varphi]_{B,C}$ heißt ja $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj}\end{pmatrix} = [\varphi(b_j)]_C$, d.h. a_{ij} ist der i-te Koeffizient von $\varphi(b_j)$ bzgl. C,

d.h. $a_{ij} = g_i(\varphi(b_j))$ nach 5.9(iv).

$B = [\varphi^*]_{C^*,B^*}$ heißt $\begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj}\end{pmatrix} = [\varphi^*(g_j)]_{B^*}$, d.h. b_{ij} ist der i-te Koeffizient von $\varphi^*(g_j)$ bzgl. B^*,

d.h. gemäß 5.9(iv) also $b_{ij} = \varphi^*(g_j)(b_i) = g_j(\varphi(b_i))$.

Daher folgt $b_{ij} = a_{ji}$ $(i = 1,\dots,m;\ j = 1,\dots,n)$, d.h. $B = A^t$ oder

$$[\varphi^*]_{C^*,B^*} = [\varphi]^t_{B,C}.$$

(iii) $\operatorname{rg}\varphi^* = \operatorname{rg} B = \operatorname{rg} A^t = \operatorname{rg} A = \operatorname{rg}\varphi$

(iv) $(\varphi\circ\psi)^* = \psi^*\circ\varphi^*$ folgt aus den Diagrammen

$$\begin{tikzcd} U \arrow[r, "\psi"'] \arrow[rr, bend left, "\varphi\circ\psi"] & V \arrow[r, "\varphi"'] & W \end{tikzcd}$$

$$\begin{tikzcd} U^* & V^* \arrow[l, "\psi^*"'] & W^* \arrow[l, "\varphi^*"'] \arrow[ll, bend left, "(\varphi\circ\psi)^*"] \end{tikzcd}$$

und $(\mathrm{id}_V)^* = \mathrm{id}_{V^*}$ ist leicht zu zeigen.

§6. DETERMINANTEN

6.1. Übersicht

(i) Die Determinantenfunktion ordnet jeder quadratischen Matrix mit Eintragungen aus $\mathbb{K}$ eine Zahl $\det A \in \mathbb{K}$ zu.

(ii) Es gilt $\det(A \cdot B) = (\det A)(\det B)$ und $\det I = 1$.

(iii) Bedeutung der Determinante:

(α) Sie liefert ein Kriterium für die Invertierbarkeit einer Matrix.

(β) Sie gibt das Volumen von Parallelotopen[9] (oder Parallelepipeden) an.

(γ) Ihr Vorzeichen zeigt die Orientierung von Basen an.

(δ) Sie beschreibt das Verhalten von linearen Abbildungen bzgl. (β) und (γ).

(iv) Konkrete Formeln für $\det A$ lauten so: Für 2×2-Matrizen

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc,$$

für 3×3-Matrizen die sogenannte *Regel von Sarrus*

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \underbrace{aei}_{(\diagdown)} + \underbrace{bfg}_{(.\diagdown)} + \underbrace{dhc}_{(\diagdown^{\cdot})} - \underbrace{ceg}_{(\diagup)} - \underbrace{bdi}_{(\diagup.)} - \underbrace{fha}_{({}^{\cdot}\diagup)},$$

und allgemein für $n \times n$-Matrizen

$$\det(a_{ij}) = \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma)\, a_{1\sigma(1)}\, a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

wobei S_n die sogenannte *symmetrische Gruppe* bezeichnet, definiert als die Menge aller *Permutationen* der Menge $\{1,\dots,n\}$, das sind bijektive Abbildungen $\{1,\dots,n\} \to \{1,\dots,n\}$. Das *Signum* einer Permutation ist so erklärt: Es ist $\operatorname{sgn}\sigma = -1$, falls σ eine ungerade Anzahl an sogenannten Fehlständen aufweist, und $\operatorname{sgn}\sigma = +1$, falls die Anzahl der Fehlstände gerade ist. Ein *Fehlstand* von σ ist ein Paar $i < j$ mit $\sigma(i) > \sigma(j)$.

[9] Verallgemeinerung von Parallelogrammen im $\mathbb{R}^2$ oder Spaten, „schiefen Quadern", im $\mathbb{R}^3$.

ad (α): $\quad \det A \neq 0 \quad \Longleftrightarrow \quad A$ ist invertierbar, d.h. $\exists\, A^{-1}$

$\Leftarrow$: Wegen $A \cdot A^{-1} = I$ folgt mit (ii), dass $(\det A)(\det A^{-1}) = 1$ gilt; also muss $\det A \neq 0$ sein.

$\Rightarrow$: Wir werden in 6.15 sehen, dass es eine Formel für A^{-1} gibt, in der „$\det A$“ im Nenner vorkommt. (Vgl. auch die Cramersche Regel, die wir im Kapitel 5 erwähnt hatten.)

ad (β): Für n Vektoren $v_1, \ldots, v_n$ im $\mathbb{R}^n$ ist das davon aufgespannte Parallelotop P definiert durch $P := \{\lambda_1 v_1 + \ldots + \lambda_n v_n \mid 0 \leq \lambda_j \leq 1\,(j = 1, \ldots, n)\}$ und für das Volumen von P gilt $\mathrm{Vol}(P) = |\det(v_1 \; \cdots \; v_n)|$. Dies kann man für den Flächeninhalt eines Parallelogramms ($n = 2$) oder für das Volumens eines Spates ($n = 3$) elementargeometrisch nachrechnen und für den allgemeinen Fall entweder zur Definition erheben oder auf Basis der mehrdimensionalen Integralrechnung ermitteln.

ad (γ): Im $\mathbb{R}^3$ gilt

$$\det(b_1\; b_2\; b_3) > 0 \iff \begin{cases} b_1, b_2, b_3 \text{ sind wie Daumen, Zeigefinger, Mittelfinger} \\ \text{der rechten Hand,} \end{cases}$$

$$\det(b_1\; b_2\; b_3) < 0 \iff \begin{cases} b_1, b_2, b_3 \text{ sind wie Daumen, Zeigefinger, Mittelfinger} \\ \text{der linken Hand.} \end{cases}$$

Allgemeiner gilt für Basen B und C im $\mathbb{R}^n$: $\det B$ und $\det C$ haben genau dann das gleiche Vorzeichen, wenn sich B stetig in C deformieren lässt (durch Drehen, Strecken/Stauchen, Spreizen), sodass während des Deformationsvorgangs jederzeit eine Basis vorliegt. Man sagt in dem Fall „B und C sind gleich orientiert“. Andernfalls kann B nur unter zusätzlicher Verwendung einer Spiegelung (das ist eine sprunghafte Änderung!) in C übergeführt werden und man sagt, dass „B und C verschieden orientiert sind“.

ad (δ): Sei $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ linear und B eine beliebige Basis von $\mathbb{R}^n$ dann gilt:

- Ist P ein Parallelotop und $Q = \varphi(P)$ das Bild von P unter φ, dann ist

$$\mathrm{Vol}(Q) = \left|\det[\varphi]_{B,B}\right| \cdot \mathrm{Vol}(P).$$

- Ist $\det[\varphi]_{B,B} > 0$, dann lässt φ die Orientierung jeder Basis gleich; ist $\det[\varphi]_{B,B} < 0$, dann ist das Bild einer Basis gerade umgekehrt orientiert wie die Ausgangsbasis.

Kursorischer Zwischenbericht über Permutationen

Es gibt verschiedene Zugänge für die systematische Behandlung von Determinanten, im wesentlichen ist eine Wahl zwischen folgenden beiden Möglichkeiten zu treffen:

(a) Zuerst grundlegende Axiome formulieren, dann zunächst daraus die wichtigsten weiteren Eigenschaften und Methoden ableiten und schließlich die Existenz und Eindeutigkeit der Determinantenfunktion mittels der Theorie der Permutationen zeigen.

(b) Gleich konkrete Formeln für die Determinante angeben und damit ihre Eigenschaften nachweisen und die wesentlichen Methoden ableiten, wobei von Anfang an Permutationen gebraucht werden.

Wir haben haben mit 6.1.(iv) also mal brutal den Zugang (b) eröffnet — in der Variante (a) hätten wir die Eigenschaften (D1)-(D3) gemäß unseres Punktes 6.11 als Axiome an die Spitze gestellt. In jedem Fall ist die Theorie der Permutationen eher als vorübergehende Krücke anzusehen, derer wir uns ohnehin entledigen, sobald wir aus den Grundeigenschaften praktischere Berechnungsmethoden für Determinanten entwickelt haben. Wir werden daher auch nicht alle technischen Subtilitäten des Umganges mit Permutationen im Detail durchgehen, sondern uns teilweise mit kursorischen Notizen begnügen.

6.2. Permutationen: Wir hatten für die Determinante einer $n \times n$-Matrix $A = (a_{ij})$ die Formel

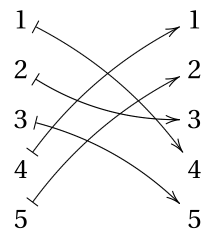
$$\det A = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) \underbrace{a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}}_{(\star)}$$

und $S_n = \{\sigma \colon \{1, \dots, n\} \to \{1, \dots, n\} \mid \sigma \text{ bijektiv}\}$. Daher kommt in $(\star)$ aus jeder Zeile und aus jeder Spalte genau ein Element vor. (Für die Zeilen war dies ohnehin offensichtlich; für die Spalten folgt dies eben aus der Bijektivität von σ.) Mit der Verknüpfung von Abbildungen als Operation bildet S_n eine Gruppe, die sogenannte *symmetrische Gruppe der Ordnung n*, in der die identische Abbildung das neutrale Element ist und die inverse Abbildung σ^{-1} das zu σ inverse Element bildet.

Betrachte zum Beispiel für $n = 5$ die hier gegebene Permutation σ:

wir verwenden hierfür auch die Schreibweise $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 1 & 2 \end{pmatrix}$;

$(\star)$ ergibt für dieses σ: $a_{14} a_{23} a_{35} a_{41} a_{52}$.



Wieviele Elemente hat S_n? Für $\sigma(1)$ gibt es n Möglichkeiten, dann für $\sigma(2)$ noch $n-1$ Möglichkeiten usw. Daher kommen wir auf $n! = n(n-1) \cdots 2 \cdot 1$ Elemente. Daher hat S_5 also 120 Elemente, d.h. für 5×5-Matrizen hat die Determinante 120 Summanden!

Etwas allgemeiner werden übrigens *Permutationen auf einer beliebigen Menge X* als bijektive Abbildungen $X \to X$ definiert. Wir haben oben also stets von Permutationen auf der Menge $\{1, \dots, n\}$ gesprochen und das ist auch der Rahmen, der uns wir für die Determinanten genügt.

6.3. Eine kleine Warnung: Es ist zwar populär, Permutationen auch als „Anordnungen von n Elementen“ oder als „Vertauschungen“ zu bezeichnen, aber bei ersterem ist nicht klar, was Verknüpfungen sein sollen und zweiteres ist ungünstig, weil zur Klarheit immer dazu gesagt werden muss, was eigentlich vertauscht werden soll, nämlich die Elemente auf den Plätzen Nummer so und so oder die Elemente mit den Nummern so und so, egal auf welchem Platz die gerade sind … und dann muss man nochmal extra aufpassen, in welcher Reihenfolge Verknüpfungen auszuwerten sind. In speziellen Anwendungen und Modellen mag das wichtig und sinnvoll sein, wir bleiben aber stur bei der Auffassung von Permutationen als bijektive Abbildungen!

6.4. Das Signum (oder Vorzeichen) einer Permutation: Entsprechend der Erklärung in 6.1.(iv) können wir schreiben

$$\operatorname{sgn} \sigma = (-1)^{F(\sigma)},$$

wobei $F(\sigma)$ die Anzahl der Fehlstände von σ ist. Fehlstand hieß ein Paar $i < j$ mit $\sigma(i) > \sigma(j)$, d.h. $F(\sigma)$ entspricht der Anzahl an Pfeilkreuzungen einer Tafel von Abbildungszuordnungen für σ wie wir sie im Beispiel in 6.2 verwendet hatten. Für das dortige σ zählen wir sieben solcher Pfeilkreuzungen, also ist $F(\sigma) = 7$, somit $\operatorname{sgn}\sigma = -1$.

Für manche Zwecke ist es günstig, eine konkretere Formel für das Signum einer Permutation σ zu haben. Dazu beobachten wir zunächst, dass für $i < j$ der Bruch $\frac{\sigma(j)-\sigma(i)}{j-i}$ negativ ist, falls (i, j) ein Fehlstandspaar ist, andernfalls positiv (und niemals null). Bilden wir das Produkt P über all diese Brüche mit $1 \le i < j \le n$, so besteht dort der Nenner aus allen Differenzen $j - i$ mit $i < j$ als Faktoren und im Zähler kommen wegen der Bijektivität von σ auch genau all diese Paarungen vor, aber für jeden Fehlstand eben in der „falschen" Reihenfolge. Daher können wir P als Produkt von lauter Faktoren $+1$ oder -1 schreiben, wobei genau für jeden Fehlstand ein Faktor -1 zu nehmen ist. Somit haben wir gezeigt:

$$\text{(S)} \qquad \operatorname{sgn}\sigma = \prod_{1\le i<j\le n} \frac{\sigma(j)-\sigma(i)}{j-i}.$$

Auf Dauer sind diese Berechnungen des Signums aber ziemlich kompliziert, daher suchen wir einfachere Methoden, indem wir Permutationen in übersichtliche Grundbausteine zerlegen.

6.5. Transpositionen sind die einfachsten Permutationen (abgesehen von der Identität), nämlich jene von der folgenden Form τ: Es gibt $k, l \in \{1, \dots, n\}$ mit $k \neq l$, sodass gilt

$$\tau(i) = \begin{cases} l & \text{für } i = k, \\ k & \text{für } i = l, \\ i & \text{sonst.} \end{cases}$$

Wir sehen aus dieser Darstellung übrigens sofort, dass jede Transposition selbstinvers ist, d.h. $\tau \circ \tau = \mathrm{id}$.

Im Fall $k < l$ liefern genau die Paare $(k, k+1), \dots, (k, l)$ und $(k+1, l), \dots, (l-1, l)$ alle Fehlstände, das sind $(l - k + 1) + (l - k) = 2(l - k) + 1$ viele, also eine ungerade Zahl. Im Fall $l < k$ sind nur die Rollen von k und l vertauscht, daher die Anzahl der Fehlstände ebenfalls ungerade. Somit gilt

$$\operatorname{sgn}\tau = -1.$$

6.6. Zerlegung einer Permutation in Transpositionen: Ist $\sigma = \mathrm{id}$, so können wir eine beliebige Transposition τ wählen und erhalten $\sigma = \mathrm{id} = \tau \circ \tau$, weil Transpositionen selbstinvers sind.

Ist $\sigma \neq \mathrm{id}$, dann gibt es ein minimales $k_1 \in \{1, \dots, n\}$ mit $\sigma(k_1) \neq k_1$. Wegen der Minimalität von k_1 muss also $l_1 := \sigma(k_1) > k_1$ gelten. Wir definieren die Transposition τ_1 durch

$$\tau_1(i) = \begin{cases} l_1 & \text{für } i = k_1, \\ k_1 & \text{für } i = l_1, \\ i & \text{sonst.} \end{cases}$$

Es gilt dann bereits $\tau_1 \circ \sigma(i) = i$ für $i \le k_1$. Im zweiten Schritt wählen wir das minimale $k_2 \in \{1, \dots, n\}$ mit $\sigma(k_2) \neq k_2$ (es folgt $k_2 > k_1$), also $l_2 := \sigma(k_2) > k_2$ wegen der Minimalität, und definieren τ_2 durch $\tau_2(k_2) = l_2$, $\tau_2(l_2) = k_2$ und $\tau_2(i) = i$ für $i \neq k_2$, $i \neq l_2$. Dann gilt $\tau_2 \circ \sigma(i) = i$ für $i \le k_2$. Wir kommen nach endlich vielen Schritten zur Situation $\tau_r \circ \ldots \circ \tau_1 \circ \sigma = \mathrm{id}$ mit Transpositionen $\tau_1, \ldots, \tau_r$, was nach Verknüpfung von links mit τ_r, dann τ_{r-1} usw. schließlich die Gleichung $\sigma = \tau_1 \circ \ldots \circ \tau_r$ ergibt.

Wir haben also gezeigt: *Jede Permutation kann als Produkt von Transpositionen geschrieben werden.*

Bemerkung: Es gibt allerdings stets mehrere Möglichkeiten der Zerlegung in Transpositionen; z.B. gilt ja die Gleichung $\mathrm{id} = \tau \circ \tau$ für jede Transposition.

6.7. Satz: $\forall \sigma, \rho \in S_n$: $\mathrm{sgn}(\sigma \circ \rho) = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\rho)$.

Mit anderen Worten: „Die Abbildung $\mathrm{sgn}\colon S_n \to \{-1, +1\}$ ist ein Gruppenhomomorphismus, wobei auf $\{-1, +1\}$ die gewöhnliche Multiplikation als Operation betrachtet wird."

Beweis: Wir verwenden die Formel (S) für die Berechnung des Signums von $\sigma \circ \rho$ und erweitern dann jeden Faktor im Zähler und Nenner um den entsprechenden Faktor $\rho(j) - \rho(i)$:

$$\mathrm{sgn}(\sigma \circ \rho) = \prod_{i<j} \frac{\sigma(\rho(j)) - \sigma(\rho(i))}{j-i} = \underbrace{\prod_{i<j} \frac{\sigma(\rho(j)) - \sigma(\rho(i))}{\rho(j) - \rho(i)}}_{(\star)} \cdot \underbrace{\prod_{i<j} \frac{\rho(j) - \rho(i)}{j-i}}_{\mathrm{sgn}(\rho)}.$$

Zu zeigen ist nun noch, dass $(\star) = \mathrm{sgn}(\sigma)$ gilt. Wir lassen in $(\star)$ die Faktoren mit $\rho(i) < \rho(j)$ unverändert und schreiben aber jene Faktoren, wo $\rho(i) > \rho(j)$ gilt, auf den äquivalenten Ausdruck $\frac{\sigma(\rho(i)) - \sigma(\rho(j))}{\rho(i) - \rho(j)}$ um. Wegen der Bijektivität von ρ durchlaufen die Faktoren in dieser Form nun alle Brüche der Form $\frac{\sigma(l) - \sigma(k)}{l-k}$ mit $1 \le k < l \le n$ (setze $k = \rho(i)$, $l = \rho(j)$, falls $\rho(i) < \rho(j)$, und $k = \rho(j)$, $l = \rho(i)$ andernfalls). Also erhalten wir in der Tat $(\star) = \mathrm{sgn}(\sigma)$. □

6.8. Korollar: Ist $\sigma = \tau_1 \circ \ldots \circ \tau_r$ eine Zerlegung der Permutation σ in Transpositionen $\tau_1, \ldots, \tau_r$, so gilt $\mathrm{sgn}\,\sigma = \mathrm{sgn}(\tau_1) \cdots \mathrm{sgn}(\tau_r) = (-1)^r$. [Folgt unmittelbar aus 6.7.]

6.9. Korollar: $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}\,\sigma$.

Beweis: $\sigma \circ \sigma^{-1} = \mathrm{id} \Rightarrow \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}\,\mathrm{id} = 1 \Rightarrow \mathrm{sgn}(\sigma^{-1}) = \frac{1}{\mathrm{sgn}\,\sigma} = \mathrm{sgn}\,\sigma$ (weil $\mathrm{sgn}\,\sigma = \pm 1$). □

Der folgende Satze wird hier zwar höchstens für den Spezialfall $G = S_n$ benötigt, aber es kostet nichts (abgesehen von einer anderen Wahl der Buchstaben) und ist auch interessant, ihn als Aussage für ganz allgemeine Gruppen zu formulieren:

6.10. Proposition: Sei $(G, \circ)$ eine Gruppe und $g \in G$. Dann gilt:

(i) Durchläuft x die Gruppe G, dann auch $x \circ g$ [analog für $g \circ x$]; d.h. die Abbildung $x \mapsto x \circ g$, $G \to G$ ist bijektiv.

(ii) Durchläuft x die Gruppe G, dann auch x^{-1}; d.h. die Abbildung $x \mapsto x^{-1}$ ist bijektiv.

Beweis: (i): Injektivität: $x_1 \circ g = x_2 \circ g \Rightarrow x_1 = x_1 \circ (g \circ g^{-1}) = (x_1 \circ g) \circ g^{-1} = (x_2 \circ g) \circ g^{-1} = x_2 \circ (g \circ g^{-1}) = x_2$.

Surjektivität: Sei $y \in G$ beliebig; setze $x := y \circ g^{-1}$. Dann folgt $x \circ g = y \circ g \circ g^{-1} = y$.

(ii) Injektivität: $x_1^{-1} = x_2^{-1} \Rightarrow x_1 = x_1 \circ (x_2^{-1} \circ x_2) = x_1 \circ (x_1^{-1} \circ x_2) = (x_1 \circ x_1^{-1}) \circ x_2 = x_2$.

Surjektivität: Sei $y \in G$ beliebig; setze $x := y^{-1}$. Dann folgt $x^{-1} = (y^{-1})^{-1} = y$. □

6.11. Eigenschaften der Determinante:

(D1) det ist linear in jeder Zeile der Matrix (bei jeweils fixen restlichen Zeilen).

(D2) Sind zwei Zeilen von A gleich, so ist $\det A = 0$.

(D3) $\det I = 1$.

Die Punkte (D1)-(D3) sind die „Grundeigenschaften" der Determinante (vgl. Satz 6.12 bzw. auch die Axiome in Fischers „Lernbuch", Abschnitt 3.2.1). Weitere Eigenschaften sind:

(D4) Geht die Matrix A' aus der Matrix A durch Platztausch zweier Zeilen hervor, so gilt $\det A' = -\det A$.

(D5) Sei $\lambda \in \mathbb{K}$. Entsteht A' aus A durch Addition des λ-fachen der j-ten Zeile zur i-ten Zeile, dann gilt $\det A' = \det A$.

(D6) $\det A^t = \det A$.

Wegen (D6) gelten also (D1), (D2), (D4) und (D5) auch für Spalten anstelle von Zeilen.

(D7) Ist $A = (a_{ij})$ eine untere oder obere Dreiecksmatrix, d.h.

$$A = \begin{pmatrix} a_{11} & & 0 \\ \vdots & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \ (i < j : a_{ij} = 0) \quad \text{oder} \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{pmatrix} \ (i > j : a_{ij} = 0),$$

so ist $\det A = a_{11} \cdots a_{nn}$.

(D8) Ist A eine „Blockmatrix" mit quadratischen Matrizen $A_1, \dots, A_m$, d.h. von der Gestalt

$$A = \begin{pmatrix} \boxed{A_1} & & 0 \\ & \ddots & \\ \star & & \boxed{A_m} \end{pmatrix} \text{ oder } A = \begin{pmatrix} \boxed{A_1} & & \star \\ & \ddots & \\ 0 & & \boxed{A_m} \end{pmatrix}, \text{ dann gilt } \det A = (\det A_1) \cdots (\det A_m).$$

(D9) $\det A \neq 0 \iff \operatorname{rg} A = n \overset{\text{[nach 4.51]}}{\iff} \exists\, A^{-1}$.

(D10) $\det(A \cdot B) = (\det A) \cdot (\det B)$.

Wir werden einige Teile des Beweis von Satz 6.11 als optionale Zusatzlektüre im Detail ausführen; vor allem sei herausgestrichen, dass (D4) schon allein aus (D1) und (D2) gefolgert werden kann. Außerdem zeigen wir in 6.12 und 6.13 unten, dass (D10) recht elegant auf abstrakte Weise aus der Eindeutigkeit der Determinantenfunktion folgt.

Zu den Beweisen: Wir schreiben $A = \begin{pmatrix} \boxed{a_1} \\ \vdots \\ \boxed{a_n} \end{pmatrix}$ mit Zeilenvektoren $a_k = (a_{k1}, \ldots, a_{kn})$ $(1 \le k \le n)$.

(D1): Ist der i-te Zeilenvektor eine Summe $a_i = a_i' + a_i''$ von Zeilenvektoren, dann folgt

$$\det A = \sum_\sigma a_{1\sigma(1)} \cdots (a'_{i\sigma(i)} + a''_{i\sigma(i)}) \cdots a_{n\sigma(n)} = \sum_\sigma a_{1\sigma(1)} \cdots a'_{i\sigma(i)} \cdots a_{n\sigma(n)} + \sum_\sigma a_{1\sigma(1)} \cdots a''_{i\sigma(i)} \cdots a_{n\sigma(n)}$$
$$= \det \begin{pmatrix} \vdots \\ \boxed{a_i'} \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ \boxed{a_i''} \\ \vdots \end{pmatrix}.$$

Analog erhalten wir

$$\det \begin{pmatrix} \vdots \\ \boxed{\lambda a_i} \\ \vdots \end{pmatrix} = \sum_\sigma a_{1\sigma(1)} \cdots (\lambda a_{i\sigma(i)}) \cdots a_{n\sigma(n)} = \lambda \sum_\sigma a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} = \lambda \det \begin{pmatrix} \vdots \\ \boxed{a_i} \\ \vdots \end{pmatrix}.$$

(D2): Sei für $1 \le i < j \le n$ die i-te Zeile gleich der j-ten Zeile, d.h. $a_i = a_j$.
Sei τ die Transposition mit $\tau(i) = j$, $\tau(j) = i$ und $\tau(l) = l$ für $l \neq i$, $l \neq j$.

Gemäß 6.10.(i) ist die Abbildung $H\colon S_n \to S_n$, $\sigma \mapsto \sigma \circ \tau$ bijektiv. Wir betrachten die Teilmenge $S_n^+ := \{\sigma \in S_n \mid \operatorname{sgn}\sigma = +1\} \subset S_n$ und behaupten, dass

(Z) $$S_n^+ \cap H(S_n^+) = \emptyset \quad \text{und} \quad S_n^+ \cup H(S_n^+) = S_n$$

gilt: Die erste Relation folgt sofort aus $\operatorname{sgn}(\sigma\circ\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) = -\operatorname{sgn}\sigma$; in der zweiten Relation ist die Inklusion „$\subseteq$" klar und es bleibt nur noch $S_n \setminus H(S_n^+) \subseteq S_n^+$ zu zeigen; ist $\sigma \in S_n \setminus H(S_n^+)$, so folgt wegen der Bijektivität von H, dass es ein eindeutiges $\rho \in S_n \setminus S_n^+$ gibt (d.h. $\operatorname{sgn}\rho = -1$), sodass $\sigma = H(\rho)$ gilt; daher ist $\operatorname{sgn}\sigma = \operatorname{sgn}(\rho) \circ \operatorname{sgn}(\tau) = -\operatorname{sgn}\rho = +1$, also $\sigma \in S_n^+$.

Mit Hilfe von (Z) können wir nun die Berechnung der Determinante wie folgt aufspalten:

$$\begin{aligned}
\det A &= \sum_{\sigma\in S_n^+} \underbrace{(\operatorname{sgn}\sigma)}_{+1} a_{1\sigma(1)} \cdots a_{n\sigma(n)} + \sum_{\sigma\in S_n^+} \underbrace{(\operatorname{sgn}\sigma\circ\tau)}_{-1} a_{1\sigma(\tau(1))} \cdots a_{n\sigma(\tau(n))} \\
&= \sum_{\sigma\in S_n^+} a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{j\sigma(j)} \cdots a_{n\sigma(n)} - \sum_{\sigma\in S_n^+} \underbrace{a_{1\sigma(\tau(1))}}_{a_{1\sigma(1)}} \cdots \underbrace{a_{i\sigma(\tau(i))}}_{a_{i\sigma(j)}} \cdots \underbrace{a_{j\sigma(\tau(j))}}_{a_{j\sigma(i)}} \cdots \underbrace{a_{n\sigma(\tau(n))}}_{a_{n\sigma(n)}} \\
&\qquad [\text{beachte: } a_{i\sigma(j)} = a_{j\sigma(j)}, a_{j\sigma(i)} = a_{i\sigma(i)}] \\
&= \sum_{\sigma\in S_n^+} a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{j\sigma(j)} \cdots a_{n\sigma(n)} - \sum_{\sigma\in S_n^+} a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{j\sigma(j)} \cdots a_{n\sigma(n)} = 0.
\end{aligned}$$

Bemerkung: S_n^+ ist eine Untergruppe von S_n (verwende 6.7), sie heißt *alternierende Gruppe.*

(D3): Es ist $\det I_n = \sum_\sigma (\operatorname{sgn}\sigma)\,\delta_{1\sigma(1)} \cdots \delta_{n\sigma(n)}$, wobei nur der Summand mit $\sigma(1) = 1, \ldots, \sigma(n) = n$ nicht verschwindet, d.h. jener für $\sigma = \text{id}$. Somit folgt $\det I_n = (\operatorname{sgn}(\text{id}))\, 1 \cdots 1 = 1$.

(D4): Sei $1 \leq i < j \leq n$ und A' aus A durch Vertauschen der Zeilen a_i und a_j entstanden. Dann folgt mit (D1) und (D2)

$$\det A + \det A' = \det\begin{pmatrix} \vdots \\ \boxed{a_i} \\ \vdots \\ \boxed{a_j} \\ \vdots \end{pmatrix} + \det\begin{pmatrix} \vdots \\ \boxed{a_j} \\ \vdots \\ \boxed{a_i} \\ \vdots \end{pmatrix} =$$

$$\det\begin{pmatrix} \vdots \\ \boxed{a_i} \\ \vdots \\ \boxed{a_j} \\ \vdots \end{pmatrix} + \underbrace{\det\begin{pmatrix} \vdots \\ \boxed{a_i} \\ \vdots \\ \boxed{a_i} \\ \vdots \end{pmatrix}}_{0} + \det\begin{pmatrix} \vdots \\ \boxed{a_j} \\ \vdots \\ \boxed{a_i} \\ \vdots \end{pmatrix} + \underbrace{\det\begin{pmatrix} \vdots \\ \boxed{a_j} \\ \vdots \\ \boxed{a_j} \\ \vdots \end{pmatrix}}_{0}$$

$$= \det\begin{pmatrix} \vdots \\ \boxed{a_i} \\ \vdots \\ \boxed{a_j + a_i} \\ \vdots \end{pmatrix} + \det\begin{pmatrix} \vdots \\ \boxed{a_j} \\ \vdots \\ \boxed{a_i + a_j} \\ \vdots \end{pmatrix} = \det\begin{pmatrix} \vdots \\ \boxed{a_i + a_j} \\ \vdots \\ \boxed{a_i + a_j} \\ \vdots \end{pmatrix} = 0.$$

(D5)-(D9): Beweise für diese Aussagen, übrigens basierend allein auf den Eigenschaften (D1)-(D4), finden sich zum Beispiel in Fischers „Lernbuch", Abschnitt 3.2.2 (wobei unser (D5) dem dortigen D4 entspricht, (D6) ist D12, (D7) ist D6, (D8) ist D8, (D9) ist D7).

(D10) wird in 6.13 bewiesen. □

Die Determinante ist schon durch die Eigenschaften (D1)-(D3) festgelegt und allein daraus lässt sich auch schon die wichtige Eigenschaft (D10) ableiten, wie wir in den beiden folgenden Aussagen zeigen:

6.12. Satz: Jede Abbildung $F\colon M(n,n;\mathbb{K}) \to \mathbb{K}$, die (D1) und (D2) erfüllt [d.h. jeweils mit „F" statt „det"], ist ein Vielfaches der Determinantenfunktion; genauer gilt $F = \lambda \cdot \det$ mit $\lambda = F(I_n)$. Erfüllt F also zusätzlich auch (D3), dann ist $F = \det$.

Beweis: Ist $A = \begin{pmatrix} \boxed{a_1} \\ \vdots \\ \boxed{a_n} \end{pmatrix}$ mit Zeilenvektoren $a_1, \ldots, a_n$, so erhalten wir

$$F(A) = F\begin{pmatrix} \boxed{\quad a_1 \quad} \\ \vdots \\ \boxed{\quad a_n \quad} \end{pmatrix} = F\begin{pmatrix} \sum_{j_1} a_{1j_1} e^t_{j_1} \\ \vdots \\ \sum_{j_n} a_{1j_n} e^t_{j_n} \end{pmatrix} \overset{\text{(D1)}}{=} \sum_{j_1} a_{1j_1} F \begin{pmatrix} e^t_{j_1} \\ \vdots \\ \sum_{j_n} a_{1j_n} e^t_{j_n} \end{pmatrix} =$$

$$[\text{Linearität in den Zeilen } 2,\dots,n] = \sum_{j_1} \cdots \sum_{j_n} a_{1j_1} \cdots a_{nj_n} \underbrace{F\begin{pmatrix} e^t_{j_1} \\ \vdots \\ e^t_{j_n} \end{pmatrix}}_{(\star)}$$

Nach (D2) ist $(\star) \neq 0$ nur dann, wenn alle j_i verschieden sind, also $j_1 = \sigma(1), \dots, j_n = \sigma(n)$ mit einer Permutation $\sigma \in S_n$ gilt. Nun entsteht $\begin{pmatrix} e^t_{\sigma(1)} \\ \vdots \\ e^t_{\sigma(n)} \end{pmatrix}$ aus I_n durch r Zeilenvertauschungen gemäß einer Zerlegung $\sigma = \tau_1 \circ \cdots \circ \tau_r$ in Transpositionen τ_i. Da (D4) schon direkt aus (D1) und (D2) folgt, dürfen wir auf $(\star) = (-1)^r F(I_n) = (\operatorname{sgn} \sigma) \cdot F(I_n)$ schließen und erhalten insgesamt $F(A) = F(I) \cdot \det A$. □

6.13. Korollar: $\det(A \cdot B) = (\det A) \cdot (\det B)$.

Beweis: Für festes $B \in M(n,n;\mathbb{K})$ erfüllt $F\colon M(n,n;\mathbb{K}) \to \mathbb{K}$, definiert durch $F(A) := \det(A \cdot B)$ die Eigenschaften (D1) und (D2): Sowohl det als auch $A \mapsto A \cdot B$ ist linear in jeder Zeile und zwei gleiche Zeilen in A ergeben zwei gleiche Zeilen in $A \cdot B$. Daher folgt aus 6.12 direkt $\det(A \cdot B) = F(A) = F(I) \cdot \det A = \det(I \cdot B) \cdot (\det A) = \det(B) \cdot (\det A)$. □

Zur *praktischen Berechnung* von $\det A$ ist die allgemeine Formel mit der Summe über alle Permutationen viel zu aufwendig. Besser ist es, A mittels Gauß-Verfahren auf „Zeilenstufenform“ C zu bringen, dann ist $\det A = (-1)^k \det C$, wobei k die Anzahl der dabei ausgeführten Zeilentausche ist (hier wird (D4) und (D1) verwendet). Außerdem ist C eine obere Dreiecksmatrix, weil wir es hier ja nur mit quadratischen Matrizen zu tun haben, sodass $\det C$ einfach nach (D7) berechnet werden kann.

Weiters kann man mit der unten folgenden Formel im Entwicklungssatz von Laplace die Berechnung der Determinante einer $n \times n$-Matrix auf die Berechnung von n Determinanten von $(n-1) \times (n-1)$-Matrizen zurückführen. Wir treffen zunächst die nötigen technischen Vorbereitungen: Sei $A = (a_{ij}) \in M(n,n;K)$ und $1 \le k, l \le n$, dann definieren wir

$A_{[k,l]} \in M(n-1, n-1; \mathbb{K})$, indem wir in A die Zeile k und die Spalte l streichen

und

$A_{\langle k,l\rangle} \in M(n,n;\mathbb{K})$, indem wir in A die Zeile k durch e_l^t und die Spalte l durch e_k ersetzen:

$$A_{\langle k,l\rangle} = \begin{pmatrix} \star & \boxed{0} & \star \\ \boxed{\quad 0 \quad} & 1 & \boxed{\quad 0 \quad} \\ \star & \boxed{0} & \star \end{pmatrix} \qquad (\star \text{ wie in } A).$$

Dann gilt nach mehrfacher Anwendung von (D4) und schließlich noch (D8):

$$\begin{aligned} \det A_{\langle k,l\rangle} &= (-1)^{k-1} \det \begin{pmatrix} \boxed{\quad 0 \quad} & 1 & \boxed{\quad 0 \quad} \\ & \boxed{0} & \end{pmatrix} = (-1)^{k-1}(-1)^{l-1} \det \begin{pmatrix} 1 & \boxed{\quad 0 \quad} \\ \boxed{0} & A_{[k,l]} \end{pmatrix} \\ &= (-1)^{k+l} \det A_{[k,l]}. \end{aligned}$$

6.14. Satz (Laplacescher Entwicklungssatz): Sei $A \in M(n,n;\mathbb{K})$ und i ein Zeilenindex ($1 \le i \le n$). Dann gilt die folgende Formel für die „Entwicklung nach der i-ten Zeile“:

$$\det A = \sum_{j=1}^{n} \underbrace{a_{ij}}_{\substack{\text{Einträge} \\ i\text{-te Zeile}}} \underbrace{(-1)^{i+j}}_{\substack{\text{Schachbrett-} \\ \text{muster}}} \underbrace{\det A_{[i,j]}}_{\substack{\text{streiche Z. u. S.,} \\ \text{wo } a_{ij} \text{ steht}}} . \qquad \text{Schachbrett: } \left((-1)^{i+j}\right)_{i,j} = \begin{array}{|c|c|c|c} + & - & + & \cdots \\ \hline - & + & - & \cdots \\ \hline + & - & + & \cdots \\ \hline \vdots & \vdots & \vdots & \ddots \end{array}$$

Analog gilt für jeden Spaltenindex $1 \le j \le n$:

$$\det A = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} \det A_{[i,j]}.$$

Beweis: Wegen (D6) genügt es, die erste Formel zu beweisen. Es ist mit Zeilenvektoren $a_1 = (a_{11}, \ldots, a_{1n})$ usw.

$$\det A = \det \begin{pmatrix} a_1 \\ \vdots \\ \sum_j a_{ij} e_j^t \\ \vdots \\ a_n \end{pmatrix} \overset{(D1)}{=} \sum_j a_{ij} \det \underbrace{\begin{pmatrix} & \boxed{\Diamond} & \\ \boxed{\quad 0 \quad} & 1 & \boxed{\quad 0 \quad} \\ & \boxed{\Diamond} & \end{pmatrix}}_{\text{Einser in Zeile } i \text{ und Spalte } j},$$

wobei hier in den Determinanten die Einträge in den Bereichen $\boxed{\Diamond}$ nicht vorkommen, weil aus der Zeile i ja der Einser gewählt werden muss, um in der Determinantensummenformel einen Term ungleich null zu bekommen, und somit aber die Spalte j schon vergeben ist. Daher können diese $\boxed{\Diamond}$-Bereiche ohne Schaden jeweils mit 0 gefüllt werden, d.h.

$$\det A = \sum_j a_{ij} \det \underbrace{\begin{pmatrix} & \boxed{0} & \\ \boxed{\;0\;} & 1 & \boxed{\;\;0\;\;} \\ & \boxed{0} & \end{pmatrix}}_{A_{\langle i,j\rangle}} = \sum_j a_{ij} (-1)^{i+j} \det A_{[i,j]}.$$

$\square$

6.15. Korollar: $A^{-1} = \frac{1}{\det A} \left(\det A_{\langle j,i\rangle}\right)_{i,j} = \frac{1}{\det A} \left((-1)^{i+j} \det A_{[j,i]}\right)_{i,j}.$

Beweis: Sei $B := \frac{1}{\det A} \left(\det A_{\langle j,i\rangle}\right)_{i,j}$, dann folgt

$$(A \cdot B)_{ik} = \sum_j a_{ij} \frac{\det A_{\langle k,j\rangle}}{\det A} = \frac{1}{\det A} \sum_j a_{ij} \det A_{\langle k,j\rangle} = \begin{cases} \frac{1}{\det A} \det A = 1 & \text{für } i = k, \\ \frac{1}{\det A} \det \begin{pmatrix} \vdots \\ \text{Zeile } i\text{: } a_i \\ \vdots \\ \text{Zeile } k\text{: } a_i \\ \vdots \end{pmatrix} = 0 & \text{für } i \neq k. \end{cases}$$

$\square$

6.16. Ein Beispiel zu 6.14: Bei der folgenden Determinante bietet sich Entwicklung nach der zweiten Zeile an (d.h. $i = 2$ in 6.14), weil dann schon 2 von 4 Koeffizienten null sind!

$$\det \begin{pmatrix} 7 & 1 & -1 & 5 \\ 2 & 0 & 0 & -1 \\ -3 & -2 & 1 & 4 \\ 2 & 0 & 1 & 0 \end{pmatrix} = -2 \cdot \underbrace{\det \begin{pmatrix} 1 & -1 & 5 \\ -2 & 1 & 4 \\ 0 & 1 & 0 \end{pmatrix}}_{\text{Entw. nach Zeile 3}} + (-1) \cdot \underbrace{\det \begin{pmatrix} 7 & 1 & -1 \\ -3 & -2 & 1 \\ 2 & 0 & 1 \end{pmatrix}}_{\text{Entw. nach Zeile 3}}$$

$$= (-2)(-1 \cdot \det \begin{pmatrix} 1 & 5 \\ -2 & 4 \end{pmatrix}) - (2 \cdot \det \begin{pmatrix} 1 & -1 \\ -2 & 1 \end{pmatrix} + 1 \cdot \det \begin{pmatrix} 7 & 1 \\ -3 & -2 \end{pmatrix})$$

$$= (-2)(-14) - (2(-1) + 1(-11)) = 28 - (-13) = 41.$$

6.17. Proposition: Sei $S \in M(n,n;\mathbb{K})$ invertierbar, dann gilt $\det(S^{-1}) = \dfrac{1}{\det S}$.

Beweis: $S \cdot S^{-1} = I \quad \Rightarrow \quad (\det S)\det(S^{-1}) = \det I = 1$ □

Sei V ein endlichdimensionaler Vektorraum über $\mathbb{K}$. Für Problemstellungen mit linearen Abbildungen $\varphi\colon V \to V$ in Hinblick auf Invertierbarkeit, Volumsänderung oder Orientierungsänderung (siehe 6.1(iii), (α) und (β)) benötigen wir den Begriff der Determinante einer (abstrakten) linearen Abbildung.

Sind B, C Basen von V, dann gilt für die Matrixdarstellungen von φ gemäß 4.39 die Gleichung

$$[\varphi]_{C,C} = S^{-1} \cdot [\varphi]_{B,B} \cdot S,$$

wobei S den Basiswechsel $[B \to C]$ beschreibt. Daraus ergibt sich nun

$$\det[\varphi]_{C,C} = (\det S^{-1})(\det[\varphi]_{B,B})(\det S) \overset{6.17}{=} \det[\varphi]_{B,B},$$

also geht die folgende Definition in Ordnung:

6.18. Definition: Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $\varphi\colon V \to V$ linear. Dann setzen wir

$$\det\varphi := \det[\varphi]_{B,B},$$

wobei B irgendeine Basis von V ist.

6.19. Satz: Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $\varphi\colon V \to V$ linear. Dann gilt:

$$\varphi \text{ ist invertierbar} \quad \Longleftrightarrow \quad \det\varphi \neq 0.$$

Beweis: φ invertierbar $\overset{4.36}{\Leftrightarrow}$ $[\varphi]_{B,B}$ invertierbar $\overset{(D9)}{\Leftrightarrow}$ $\det[\varphi]_{B,B} \neq 0$ $\overset{6.18}{\Leftrightarrow}$ $\det\varphi \neq 0$. □

§7. VEKTORRÄUME MIT SKALARPRODUKT

Das Skalarprodukt im $\mathbb{R}^n$ (vgl. 2.3.3) besitzt die folgenden Eignschaften — in bereits bewährter Manier schreiben wir diese aber gleich etwas abstrakter für Vektorräume über $\mathbb{R}$ oder $\mathbb{C}$ auf:

7.1. Definition: Sei V ein Vektorraum über $\mathbb{R}$ [oder $\mathbb{C}$]. Eine Abbildung $\langle . \mid . \rangle \colon V \times V \to \mathbb{R}$ [$\mathbb{C}$] heißt *Skalarprodukt* auf V, wenn die folgenden Eigenschaften für alle $u, v, w \in V$ und $\lambda \in \mathbb{R}$ [$\mathbb{C}$] gelten:[10]

(S1) $\langle u \mid v + w \rangle = \langle u \mid v \rangle + \langle u \mid w \rangle$ und $\langle v \mid \lambda w \rangle = \lambda \langle v \mid w \rangle$ linear im 2. Argument,

(S2) $\langle u + v \mid w \rangle = \langle u \mid w \rangle + \langle v \mid w \rangle$ und $\langle \lambda v \mid w \rangle = \overline{\lambda} \langle v \mid w \rangle$ [konjugiert] linear im 1. Argument,

(S3) $\langle v \mid w \rangle = \overline{\langle w \mid v \rangle}$ symmetrisch [hermitesch[11]],

(S4) $\langle v \mid v \rangle \geq 0$; $\langle v \mid v \rangle = 0 \Leftrightarrow v = 0$ positiv definit.

Die Eigenschaften (S1) und (S2) bedeuten, dass $\langle . \mid . \rangle$ eine bilineare [sesquilineare] Abbildung ist. Ein Skalarprodukt ist also eine bilineare [sesquilineare], symmetrische [hermitesche], positiv definite Abbildung $V \times V \to \mathbb{R}$ [$\mathbb{C}$]. Ein reeller VR mit Skalarprodukt wird *euklidisch* genannt, ein komplexer VR mit Skalarprodukt heißt *unitär*.

Eigenschaft (S2) folgt eigentlich schon aus (S1) und (S3); wir werden aber später auch bilineare Abbildungen betrachten, die nicht notwendig symmetrisch [hermitesch] sind.

7.2. Sinn und Zweck: Vektorrräume mit Skalarprodukt erlauben es, folgende Begriffe einzuführen und zu studieren: Abstände bzw. Längen durch $\operatorname{dist}(v, w) := \sqrt{\langle v - w \mid v - w \rangle}$ bzw. $\|v\| := \sqrt{\langle v \mid v \rangle}$, Orthogonalität durch die Definition „$v \perp w \Leftrightarrow \langle v \mid w \rangle = 0$“ und Winkel durch $\cos \measuredangle(v, w) := \frac{\langle v \mid w \rangle}{\|v\| \, \|w\|}$ (aus 7.4 und 7.6 sieht man, dass der Betrag dieser Zahl zwischen 0 und 1 liegt).

Lineare Abbildungen, die Längen und Winkel gleich lassen, sind die (Verallgemeinerungen von) Drehungen und Spiegelungen. Symmetrische bilineare Abbildungen $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, in der Gestalt von sogenannten quadratischen Formen, definieren einerseits die bekannten Kegelschnitte und deren höherdimensionale Pendants und spielen andererseits in der Relativitätstheorie eine wesentliche Rolle (Lorentz-Metrik, Minkowski-Raum). Last but not least agieren die sogenannten Observablen in der Quantenmechanik in dem mathematischen „Bühnenbild“ eines komplexen Hilbertraumes, das ist ein unitärer Vektorraum mit einer zusätzlichen Vollständigkeitsbedingung (die im Endlichdimensionalen immer automatisch erfüllt ist).

[10] Der Querstrich für komplex Konjugierte ist im reellen Fall natürlich einfach zu ignorieren.
[11] oft auch hermitisch genannt, geht aber auf Charles Hermite zurück …

7.3. Beispiele: 1. Standardskalarprodukt: $\langle x \mid y\rangle := \sum_{j=1}^{n} x_j\, y_j$ auf $\mathbb{R}^n$, $\langle x \mid y\rangle := \sum_{j=1}^{n} \overline{x_j}\, y_j$ auf $\mathbb{C}^n$.

2. $\mathbb{R}^2$ mit $\langle x \mid y\rangle := x_1 y_1 - 2(x_1 y_2 + x_2 y_1) + 5 x_2 y_2 = (x_1\ \ x_2)\cdot \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

(Beweis von (S1)-(S4) als Übung.)

3. $C[0,1] = \{f\colon\ [0,1] \to \mathbb{C} \mid f \text{ stetig}\}$ mit $\langle f \mid g\rangle := \int_0^1 \overline{f(t)}\, g(t)\, dt$

(S1)-(S3) ist sehr leicht zu zeigen, (S4) ist ein bisschen heikler.

(Nimmt man statt stetiger Funktionen alle Riemann-integrierbaren Funktionen, dann ist (S4) nicht mehr gültig!)

4. $M(n,n;\mathbb{R})$ mit $\langle A \mid B\rangle := \operatorname{trace}(A^t \cdot B)$, wobei $\operatorname{trace} C := \sum_{i=1}^{n} c_{ii}$ die *Spur* von $C = (c_{ij})$ ist;

d.h. $\langle A \mid B\rangle = \sum_{i=1}^{n} (A^t\cdot B)_{ii} = \sum_i \sum_j (A^t)_{ij}\, (B)_{ji} = \sum_{i,j} a_{ji}\, b_{ji}$.

5. $l^2 := \{(x_n)_{n\in\mathbb{N}} \mid x_n \in \mathbb{C},\ \sum_{n=1}^{\infty} |x_n|^2 < \infty\}$ ist ein VR über $\mathbb{C}$ mit Skalarprodukt

$$\langle x \mid y\rangle := \sum_{n=1}^{\infty} \overline{x_n}\, y_n$$

(hier ohne Beweis[12]); dies ist der Prototyp eines sogenannten *Hilbertraumes.*

6. $L^2[0,1] := \{f\colon\ [0,1] \to \mathbb{C} \mid f \text{ Lebesgue-messbar und } \int_0^1 |f|^2 < \infty\}$ mit $\langle f \mid g\rangle := \int_0^1 \overline{f(t)}\, g(t)\, dt$

ist auch ein *Hilbertraum.* (Analog $L^2(\mathbb{R})$, $L^2(\mathbb{R}^3)$ etc; wichtig in der Quantenmechanik!)

7.4. Proposition (Cauchy-Schwarz-Ungleichung): Sei V ein euklidischer oder unitärer VR und $v, w \in V$. Dann gilt

$$|\langle v \mid w\rangle|^2 \le \langle v \mid v\rangle \langle w \mid w\rangle.$$

Außerdem tritt Gleichheit genau dann ein, wenn v, w linear abhängig ist.

Beweis: 1. Fall $w = 0$: dann sind beide Seiten 0 und v, w ist tatsächlich linear abhängig.

2. Fall $w \neq 0$: Für einen beliebigen Skalar λ gilt stets

$$0 \le \langle v - \lambda w \mid v - \lambda w\rangle = \langle v \mid v\rangle - \overline{\lambda}\langle w \mid v\rangle - \lambda \langle v \mid w\rangle + |\lambda|^2 \langle w \mid w\rangle.$$

Nun setzen wir $\lambda := \frac{\langle w|v\rangle}{\langle w|w\rangle}$ (hier verwenden wir $w \neq 0$!) und erhalten aus obiger Ungleichung

$$0 \le \langle v \mid v\rangle - \frac{|\langle v \mid w\rangle|^2}{\langle w \mid w\rangle} - \frac{|\langle v \mid w\rangle|^2}{\langle w \mid w\rangle} + \frac{|\langle v \mid w\rangle|^2}{\langle w \mid w\rangle^2} \langle w \mid w\rangle = \langle v \mid v\rangle - \frac{|\langle v \mid w\rangle|^2}{\langle w \mid w\rangle},$$

[12] Nur soviel: $|x_n y_n| \le \frac{|x_n|^2 + |y_n|^2}{2}$, daher ist die angegebene Summe absolut konvergent.

daher nach Multiplikation mit $\langle w \mid w\rangle$ insgesamt die behauptete Ungleichung.

Gilt $|\langle v \mid w\rangle|^2 = \langle v \mid v\rangle\langle w \mid w\rangle$, dann folgt $\langle v \mid v\rangle - \frac{|\langle v|w\rangle|^2}{\langle w|w\rangle} = 0$ und in obigen Ungleichungsketten sind nur Gleichheitszeichen möglich. Daher folgt $0 = \langle v - \lambda w \mid v - \lambda w\rangle$, somit ist $v - \lambda w = 0$ (wegen (S4)!), d.h. v, w linear abhängig.

Sei v, w linear abhängig: Im Falle $w = \mu v$ ergibt sich $|\langle v \mid w\rangle|^2 = |\mu|^2 \langle v \mid v\rangle^2 = \langle v \mid v\rangle\langle \mu v \mid \mu v\rangle = \langle v \mid v\rangle\langle w \mid w\rangle$; und für den Fall $v = \sigma w$ ist die Rechnung analog. $\square$

Für das Standardskalarprodukt besagt 7.4 übrigens $\left|\sum\limits_{i=1}^{n} \overline{x_i}\, y_i\right| \leq \sqrt{\sum\limits_{i=1}^{n} |x_i|^2} \cdot \sqrt{\sum\limits_{i=1}^{n} |y_i|^2}$.

7.5. Definition: Sei V ein VR über $\mathbb{R}$ oder $\mathbb{C}$. Eine *Norm* auf V ist eine Abbildung $\|.\| : V \to \mathbb{R}$ mit folgenden Eigenschaften:

(N1) $\|v\| \geq 0$; $\|v\| = 0 \iff v = 0$,

(N2) $\|\lambda v\| = |\lambda|\, \|v\|$,

(N3) $\|v + w\| \leq \|v\| + \|w\|$ („Dreiecksungleichung").

Eine Norm dient zur Bestimmung der Länge eines Vektors v (nämlich $\|v\|$) bzw. zur Bestimmung des Abstandes zweier Punkte/Vektoren v und w (nämlich $\|v - w\| = \|w - v\|$). Ein Vektor mit $\|v\| = 1$ heißt *normiert.*

7.6. Proposition: Auf jedem euklidischen/unitären Vektorraum stellt

$$\|v\|_2 := \sqrt{\langle v \mid v\rangle}$$

eine Norm dar. (Die sogenannte euklidische oder 2-Norm.)

Beweis: (N1) folgt direkt aus (S4).

(N2): $\|\lambda v\|_2^2 = \langle \lambda v \mid \lambda v\rangle = \overline{\lambda}\lambda\langle v \mid v\rangle = |\lambda|^2\langle v \mid v\rangle$.

(N3): $\|v + w\|_2^2 = \langle v + w \mid v + w\rangle = \langle v \mid v\rangle + \langle v \mid w\rangle + \underbrace{\langle w \mid v\rangle}_{\overline{\langle v|w\rangle}} + \langle w \mid w\rangle =$

$$\begin{aligned} &= \langle v \mid v\rangle + 2\,\mathrm{Re}\,\langle v \mid w\rangle + \langle w \mid w\rangle \leq \langle v \mid v\rangle + 2\,|\langle v \mid w\rangle| + \langle w \mid w\rangle \\ &\overset{[7.4]}{\leq} \|v\|_2^2 + 2\,\|v\|_2\|w\|_2 + \|w\|_2^2 = (\|v\|_2 + \|w\|_2)^2. \end{aligned}$$

$\square$

Im Spezialfall des Standardskalarprodukts erhalten wir gemäß 7.6 die Norm $\|x\|_2 = \sqrt{\sum\limits_{i=1}^{n} |x_i|^2}$.

Weitere Beispiele für Normen auf $\mathbb{R}^n$ sind $\|x\|_1 := \sum_{i=1}^{n} |x_i|$, deren Abstandsbegriff gemäß $\|x-y\|_1$ übrigens die sogenannte „New Yorker Taximetrik“ ist[13], oder $\|x\|_\infty := \max_{1\le i\le n} |x_i|$.

Bemerkung: Die beiden letztgenannten Normen stammen jedoch nicht von einem Skalarprodukt. Es gilt nämlich: Eine Norm stammt genau dann von einem Skalarprodukt via 7.6 ab, wenn sie die Parallelogrammgleichung

$$\|v+w\|^2 + \|v-w\|^2 = 2\,(\|v\|^2 + \|w\|^2)$$

erfüllt;[14] die Normen $\|.\|_1$ und $\|.\|_\infty$ erfüllen die Parallelogrammgleichung nicht.

7.7. Definition: Sei V ein euklidischer oder unitärer VR und $v, w \in V$. Wir sagen v ist *orthogonal* zu w, in Zeichen $v \perp w$, wenn $\langle v \mid w\rangle = 0$ gilt.

7.8. Definition: Sei V ein euklidischer oder unitärer VR.
Ein System von Vektoren $v_1, \dots, v_k \in V$ heißt

Orthogonalsystem, falls $\langle v_i \mid v_j\rangle = 0$ für $i \neq j$ gilt,

und *Orthonormalsystem* (ONS), falls $\langle v_i \mid v_j\rangle = \delta_{ij}$ für alle i, j gilt (Orthogonalsystem u. $\|v_i\| = 1$).

Ein Orthonormalsystem, das auch eine Basis von V ist, nennen wir *Orthonormalbasis* (ONB).

7.9. Proposition: Jedes Orthonormalsystem ist linear unabhängig.

Beweis: Es sei $0 = \lambda_1 v_1 + \ldots + \lambda_k v_k$. Wir nehmen beide Seiten dieser Gleichung ins Skalarprodukt mit v_i und erhalten

$$0 = \lambda_1\langle v_i \mid v_1\rangle + \ldots + \lambda_k\langle v_i \mid v_k\rangle = \lambda_i \cdot 1.$$

Für $i = 1, \dots, k$ ergibt sich somit $\lambda_1 = \ldots = \lambda_k = 0$. □

7.10. Satz: Aus jedem linear unabhängigen System $v_1, \dots, v_k$ in einem euklidischen/unitären Vektorraum V lässt sich ein Orthonormalsystem $w_1, \dots, w_k$ gewinnen mit der Eigenschaft

$$[v_1, \dots, v_l] = [w_1, \dots, w_l] \quad (1 \le l \le k).$$

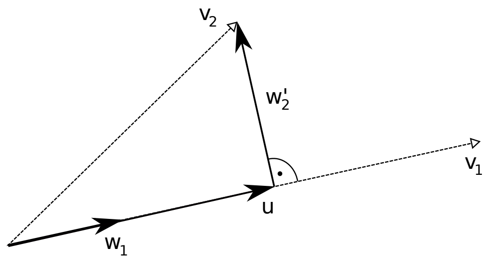
Beweis mittels *Orthogonalisierungsverfahren von Gram-Schmidt*:

- Setze $w_1' := v_1$; es ist $w_1' \neq 0$, da $v_1, \dots, v_k$ linear unabhängig ist; für $w_1 := \frac{1}{\|w_1'\|} w_1'$ gilt auch $\|w_1\| = \frac{1}{\|w_1'\|}\|w_1'\| = 1$ und $[w_1] = [w_1'] = [v_1]$.

[13] Machen Sie eine Skizze im $\mathbb{R}^2$, um diesen Namen zu verstehen!

[14] Einen Beweis dafür finden Sie z.B. am Ende von §1 aus Kapitel 5 in M. Koecher, *Lineare Algebra und analytische Geometrie*, Springer-Verlag, 4. Auflage 1997.

- Die Idee für den nächsten Schritt lässt sich in der „von w_1 und v_2 aufgespannten Ebene“ gut veranschaulichen:



Der Vektor $u := \langle w_1 \mid v_2 \rangle w_1$ entspricht der Projektion von v_2 auf w_1; daher steht $w_2' := v_2 - u$ senkrecht auf v_1.

Setze $w_2' := v_2 - \langle w_1 \mid v_2 \rangle w_1$; es ist $w_2' \neq 0$, weil andernfalls $v_2 = \langle w_1 \mid v_2 \rangle w_1 = \lambda v_1$ im Widerspruch zur linearen Unabhängigkeit von $v_1, \ldots, v_k$ wäre; der Vektor $w_2 := \frac{1}{\| w_2' \|} w_2'$ hat Länge 1 und es gilt

$$\langle w_1 \mid w_2 \rangle = \frac{1}{\| w_2' \|} \langle w_1 \mid v_2 - \langle w_1 \mid v_2 \rangle w_1 \rangle = \frac{1}{\| w_2' \|} \big(\langle w_1 \mid v_2 \rangle - \langle w_1 \mid v_2 \rangle \underbrace{\langle w_1 \mid w_1 \rangle}_{1} \big) = 0,$$

d.h. w_1, w_2 ist ein Orthonormalsystem, also nach 7.9. linear unabhängig; weiters ist nach Konstruktion $[w_1, w_2] = [w_1', w_2'] \subseteq [v_1, v_2, w_1] = [v_1, v_2]$ und wegen Dimensionsgleichheit (nämlich jeweils 2) sind diese linearen Erzeugnisse nach 3.28(ii) sogar gleich.

- Es ist $w_3' := v_3 - \langle w_1 \mid v_3 \rangle w_1 - \langle w_2 \mid v_3 \rangle w_2 \neq 0$, weil sonst v_3 als Linearkombination von v_1 und v_2 darstellbar wäre (im Widerspruch zur linearen Unabhängigkeit von $v_1, \ldots, v_k$); $w_3 := \frac{1}{\| w_3' \|} w_3'$ hat Länge 1 und es gilt für $i = 1$ oder $i = 2$

$$\begin{aligned} \langle w_i \mid w_3 \rangle &= \frac{1}{\| w_3' \|} \langle w_i \mid v_3 - \langle w_1 \mid v_3 \rangle w_1 - \langle w_2 \mid v_3 \rangle w_2 \rangle \\ &= \frac{1}{\| w_3' \|} \big(\langle w_i \mid v_3 \rangle - \langle w_1 \mid v_3 \rangle \underbrace{\langle w_i \mid w_1 \rangle}_{\delta_{i1}} - \langle w_2 \mid v_3 \rangle \underbrace{\langle w_i \mid w_2 \rangle}_{\delta_{i2}} \big) = \frac{\langle w_i \mid v_3 \rangle - \langle w_i \mid v_3 \rangle}{\| w_3' \|} = 0, \end{aligned}$$

daher ist w_1, w_2, w_3 ein Orthonormalsystem und nach 7.9 linear unabhängig; nach Konstruktion ist $[w_1, w_2, w_3] = [w_1', w_2', w_3'] \subseteq [v_1, v_2, v_3, w_1, w_2] = [v_1, v_2, v_3]$, wobei alle linearen Erzeugnisse aber die gleiche Dimension 3 haben, also nach 3.28(ii) gleich sind.

- usw. d.h. eigentlich vollständige Induktion: es ist $w_j' := v_j - \sum_{l=1}^{j-1} \langle w_l \mid v_j \rangle w_l \neq 0$ wegen der linearen Unabhängigkeit von $v_1, \ldots, v_k$; sodann normieren wir $w_j := \frac{1}{\| w_j' \|} w_j'$ und bemerken analog zu oben, dass $[w_1, \ldots, w_j] \subseteq [v_1, \ldots, v_j]$ gilt; schließlich zeigen wir durch explizite Berechnung für $1 \le i < j$, dass $w_1, \ldots, w_j$ ein ONS ist:

$$\begin{aligned}
\langle w_i \mid w_j \rangle &= \frac{1}{\| w_j' \|} \langle w_i \mid v_j - \sum_{l=1}^{j-1} \langle w_l \mid v_j \rangle w_l \rangle \\
&= \frac{1}{\| w_j' \|} \Big(\langle w_i \mid v_j \rangle - \underbrace{\sum_{l=1}^{j-1} \langle w_l \mid v_j \rangle \underbrace{\langle w_i \mid w_l \rangle}_{\delta_{il}}}_{\langle w_i | v_j \rangle} \Big) = \frac{1}{\| w_j' \|} \big(\langle w_i \mid v_j \rangle - \langle w_i \mid v_j \rangle \big) = 0.
\end{aligned}$$

Daher ist $w_1, \ldots, w_j$ nach 7.9 linear unabhängig ist und $[w_1, \ldots, w_j] = [v_1, \ldots, v_j]$ folgt wiederum aus 3.28(ii). □

7.11. Korollar: Jeder endlichdimensionale euklidische oder unitäre Vektorraum besitzt Orthonormalbasen.

Beweis: Wende das Gram-Schmidt-Verfahren auf eine beliebige Basis an. □

7.12. Beispiele: 1. Die Standardbasis $e_1, \ldots, e_n$ ist eine ONB in $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ bzgl. des Standardskalarprodukts.

2. In $V = \mathbb{R}^2$ mit dem Standardskalarprodukt wenden wir das Gram-Schmidt-Verfahren auf die Basis $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ an: $w_1' = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; $w_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und

$$w_2' = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \langle \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \tfrac{1}{2} \cdot 1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ +\frac{1}{2} \end{pmatrix} = \tfrac{1}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix}; \quad w_2 = \tfrac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Orthonormalbasen haben viele Vorteile. Zwei davon zeigt die folgende Proposition: Bei der Darstellung eines Vektors bzgl. einer ONB kann man sich die Berechnung mittels Gauß-Verfahren ersparen, denn es gibt eine explizite Formel für die Koeffizienten; und bzgl. einer ONB sieht das betreffende (abstrakte) Skalarprodukt aus wie das Standardskalarprodukt.

7.13. Proposition: Sei $b_1, \ldots, b_n$ eine ONB in dem euklidischen oder unitären VR V:

(i) Ist $v = \sum_{i=1}^{n} \lambda_i b_i$, so gilt $\lambda_k = \langle b_k \mid v \rangle$, d.h. $\boxed{v = \sum_{i=1}^{n} \langle b_i \mid v \rangle b_i}$.

(ii) Ist $v = \sum_{i=1}^{n} \lambda_i b_i$ und $w = \sum_{j=1}^{n} \mu_j b_j$, so gilt $\boxed{\langle v \mid w \rangle = \sum_{i=1}^{n} \overline{\lambda_i} \mu_i}$ $(= \sum \overline{\langle b_i \mid v \rangle} \langle b_i \mid w \rangle$ nach (i)).

Beweis: (i) $\langle b_k \mid v \rangle = \langle b_k \mid \sum \lambda_i b_i \rangle = \sum \lambda_i \underbrace{\langle b_k \mid b_i \rangle}_{\delta_{ki}} = \lambda_k$.

(ii) $\langle v \mid w \rangle = \langle \sum_i \lambda_i b_i \mid \sum_j \mu_j b_j \rangle = \sum_{i,j} \overline{\lambda_i} \mu_j \underbrace{\langle b_i \mid b_j \rangle}_{\delta_{ij}} = \sum_i \overline{\lambda_i} \mu_i$. □

Bemerkung: Das Ergebnis $\lambda_k = \langle b_k \mid v\rangle$ sollte nicht allzu sehr verwundern, denn die linearen Funktionale $v \mapsto \langle b_k \mid v\rangle$ wirken auf den b_i ja gerade wie die Elemente f_k der Dualbasis (nämlich mit Resultat δ_{ik}), somit auf jedem Vektor $v \in V$ wie f_k, also $\langle b_k \mid v\rangle = f_k(v) = \lambda_k$.

7.14. Definition: Sei V ein euklidischer oder unitärer VR und $S \subseteq V$. Die Menge

$$S^\perp := \{v \in V \mid \langle v \mid w\rangle = 0\ \forall w \in S\}$$

heißt *orthogonales Komplement* von S.

7.15. Proposition: (i) $S^\perp$ ist immer ein Teilraum von V.

(ii) $S \cap S^\perp = \{0\}$.

Beweis: (i) $\langle v_1 \mid w\rangle = 0$ und $\langle v_2 \mid w\rangle = 0$ impliziert $\langle v_1 + v_2 \mid w\rangle = 0$ sowie $\langle \lambda v_1 \mid w\rangle = 0$.

(ii) $v \in S \cap S^\perp \ \Rightarrow\ \langle v \mid v\rangle = 0 \ \Rightarrow\ v = 0$. □

7.16. Beispiele: 1. Im $\mathbb{R}^2$ mit Standardskalarprodukt: Für einen 1-dim. TR W, d.h. eine Gerade durch 0, ist $W^\perp$ die auf W senkrecht stehende Gerade durch 0.

2. Im $\mathbb{R}^3$ mit Standardskalarprodukt: Für einen 2-dim. TR W, d.h. eine Ebene durch 0, ist $W^\perp$ die auf W senkrecht stehende Gerade durch 0; für einen 1-dim. TR W, d.h. eine Gerade durch 0, ist $W^\perp$ die auf W senkrecht stehende Ebene durch 0.

Wir versammeln einige Eigenschaften orthogonaler Komplemente in den folgenden beiden Propositionen, deren Beweise jeweils durch kurze Überlegungen und Rechnungen zu jedem Punkt als freiwillige Zusatzübung erbracht werden können.

7.17. Proposition: Sei V ein euklidischer oder unitärer VR und S, S_1, S_2 Teilmengen von V. Es gilt:

(i) $\{0\}^\perp = V$ und $V^\perp = \{0\}$

(ii) $S_1 \subseteq S_2 \implies S_1^\perp \supset S_2^\perp$

(iii) $S \subseteq (S^\perp)^\perp =: S^{\perp\perp}$

(iv) $(S^{\perp\perp})^\perp = S^\perp$

(v) $(S_1 \cup S_2)^\perp = S_1^\perp \cap S_2^\perp$

(vi) $S^\perp = [S]^\perp$

7.18. Proposition: Sei V ein endlichdimensionaler euklidischer oder unitärer VR und $S \subseteq V$. Dann gilt $S^{\perp\perp} = [S]$; insbesondere ist $W^{\perp\perp} = W$, falls W ein Teilraum ist.

7.19. Satz: Sei V ein endlichdimensionaler euklidischer oder unitärer VR und W ein TR von V. Dann gilt

$$V = W \oplus W^\perp$$

und damit insbesondere

$$\dim V = \dim W + \dim W^\perp.$$

Beweis: Wähle eine ONB $b_1,\dots,b_k$ von W und ergänze diese zu einer Basis von V. Mittels Gram-Schmidt-Verfahren wird daraus eine ONB $b_1,\dots,b_n$ von V.

Behauptung: $b_{k+1},\dots,b_n$ ist eine ONB von $W^\perp$.

Die lineare Unabhängigkeit ist klar, weil die genannten Vektoren Teil eines Orthonormalsystems sind. Wir zeigen, dass sie auch ein Erzeugendensystem von $W^\perp$ bilden: Ein beliebiger Vektor $v \in W^\perp \subseteq V$ lässt sich jedenfalls bzgl. der ONB von V darstellen durch $v = \sum_{i=1}^{n} \langle b_i \mid v\rangle\, b_i$.

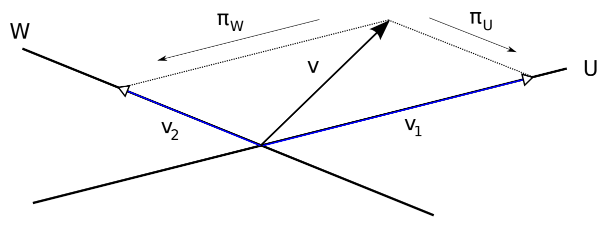
Wegen $v \in W^\perp$ gilt aber $v \perp b_i$ für $1 \le i \le k$, also folgt $v = \sum_{i=k+1}^{n} \langle b_i \mid v\rangle\, b_i \in [b_{k+1},\dots,b_n]$ und die Behauptung ist bewiesen.

Nun haben wir bereits die Dimensionsgleichung gezeigt, weil aus der Behauptung $\dim V = n$, $\dim W = k$ und $\dim W^\perp = n-k$ folgt.

Wegen $W \cap W^\perp = \{0\}$ (vgl. 7.15(ii)) müssen wir nur noch $V = W + W^\perp$ zeigen. Dies folgt aber aus der Beobachtung $v = \sum_{i=1}^{n} \langle b_i \mid v\rangle\, b_i = \underbrace{\sum_{i=1}^{k} \langle b_i \mid v\rangle\, b_i}_{\in W} + \underbrace{\sum_{i=k+1}^{n} \langle b_i \mid v\rangle\, b_i}_{\in W^\perp}$ für jedes $v \in V$. □

7.20. Wir besprechen in einem kurzen thematischen Einschub den Begriff der Projektion in allgemeinen Vektorräumen (also ohne euklidische/unitäre Struktur), bevor wir dann in 7.24 speziell zum Begriff der Orthogonalprojektion kommen. Die Punkte 7.21-7.23 können aber auch übersprungen werden, wenn die in 7.24 am Schluss angegebene Formel als Definition gelesen wird. Andernfalls folgt diese Formel dort eben aus dem Halbsatz vor „ , d.h. …“.

7.21. Definition: Sei V ein beliebiger $\mathbb{K}$-VR und U, W TR von V, sodass $V = U \oplus W$. Dann definieren wir die *Projektion von V auf U parallel zu W* durch $\pi_U \colon V \to V$, $\pi_U(v) = v_1$, wobei $v = v_1 + v_2$ die eindeutige Zerlegung von v mit $v_1 \in U$ und $v_2 \in W$ ist. (Analog definieren wir die *Projektion von V auf W parallel zu U* durch $\pi_W \colon V \to V$, $\pi_W(v) = v_2$.)



7.22. Proposition: Die Projektion π_U (und ebenso π_W) ist linear.

Beweis: Seien $v, z \in V$ und $v = v_1 + v_2$, $z = z_1 + z_2$ die Zerlegungen bzgl. $U \oplus W$. Dann gilt

$$\pi_U(v+z) = \pi_U((v_1+v_2)+(z_1+z_2)) = \pi_U(\underbrace{(v_1+z_1)}_{\in U}+\underbrace{(v_2+z_2)}_{\in W}) = v_1+z_1 = \pi_U(v)+\pi_U(z)$$

und $\pi_U(\lambda v) = \pi_U(\lambda(v_1+v_2)) = \pi_U(\underbrace{\lambda v_1}_{\in U} + \underbrace{\lambda v_2}_{\in W}) = \lambda v_1 = \lambda \pi_U(v)$. □

Offenbar gilt $\operatorname{im}\pi_U = \{v_1 \mid v_1 \in U\} = U$ und $\ker\pi_U = \{v = v_1 + v_2 \mid v_1 = 0\} = \{v_2 \mid v_2 \in W\} = W$; analog ist $\operatorname{im}\pi_W = W$, $\ker\pi_W = U$; daher gilt $\pi_U + \pi_W = \mathrm{id}_V$.

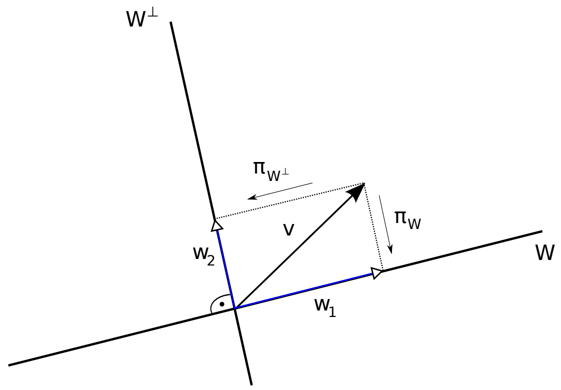
7.23. Proposition: Eine lineare Abbildung $\varphi\colon V \to V$ ist genau dann die Projektion auf einen Teilraum U von V, wenn $\varphi \circ \varphi = \varphi$ gilt (bzw. $P^2 = P$ für die Matrix von φ bzgl. einer beliebigen Basis von V).

Beweis: Wenn $\varphi = \pi_U$ gilt, dann folgt für jedes $v = v_1 + v_2$ mit $v_1 \in U$, $v_2 \in W$, stets $\varphi\circ\varphi(v) = \pi_U(\pi_U((v)) = \pi_U(v_1) = \pi_U(v_1 + 0) = v_1 = \pi_U(v) = \varphi(v)$, d.h. $\varphi \circ \varphi = \varphi$.

Ist andererseits $\varphi \circ \varphi = \varphi$, dann setzen wir $U := \operatorname{im}\varphi$ und $W := \ker\varphi$. Wir bemerken, dass $v - \varphi(v) \in \ker\varphi = W$ ist, weil $\varphi(v - \varphi(v)) = \varphi(v) - \varphi(\varphi(v)) = \varphi(v) - \varphi(v) = 0$ gilt. Weil wir stets $v = \varphi(v) + (v - \varphi(v))$ schreiben können, ist somit offensichtlich $V = U + W$. Es bleibt nur noch zu zeigen, dass dies eine direkte Summe ist, denn dann gilt $\pi_U(v) = \varphi(v)$ nach Konstruktion von U. Sei $v \in U \cap W = \operatorname{im}\varphi \cap \ker\varphi$. Weil $v \in \operatorname{im}\varphi$ ist, gibt es ein $z \in V$ mit $v = \varphi(z)$, und, weil $v \in \ker\varphi$ ist, gilt $\varphi(v) = 0$. Daher folgt $v = \varphi(z) = \varphi(\varphi(z)) = \varphi(v) = 0$, d.h. $U \cap W = \{0\}$. □

Nach 7.19 gilt $V = W \oplus W^\perp$ für jeden TR W eines endlichdimensionalen euklidischen oder unitären VR und wir haben den folgenden Projektionsbegriff.

7.24. Definition: Sei V ein endlichdimensionaler euklidischer oder unitärer VR und W ein TR von V. Die *Orthogonalprojektion* π_W von V auf W ist die Projektion von V auf W parallel zu $W^\perp$, d.h. für $v = w_1 + w_2$ mit eindeutigen $w_1 \in W$ und $w_2 \in W^\perp$ ist $\pi_W(v) = w_1$.



7.25. Bemerkung: Der Beweis von 7.19 zeigt, dass $\pi_W(v)$ durch folgende Formel gegeben ist:

$$\pi_W(v) = \sum_{i=1}^{k} \langle b_i \mid v\rangle\, b_i,$$

wobei $b_1, \dots, b_k$ eine (beliebige) ONB von W ist (vgl. 7.13(i)).

7.26. Beispiele: 1. In $V = \mathbb{R}^2$ oder $\mathbb{R}^3$ mit dem Standardskalarprodukt betrachten wir die orthogonale Projektion auf eine Gerade W durch 0: Sei $W = \{\lambda a \mid \lambda \in \mathbb{R}\}$ mit $a \in V \setminus \{0\}$. Es ist dann $\frac{1}{\|a\|}\, a$ eine ONB von W und die orthogonale Projektion eines Vektors $v \in V$ auf W ist gegeben durch $\pi_W(v) = \langle \frac{1}{\|a\|}\, a \mid v\rangle \frac{1}{\|a\|}\, a = \frac{\langle a \mid v\rangle}{\|a\|^2}\, a = \frac{\langle a \mid v\rangle}{\langle a \mid a\rangle}\, a.$

2. In $V = \mathbb{R}^3$ mit Standardskalarprodukt betrachten wir die Orthogonalprojektion auf eine Ebene W durch 0: Wir haben also $\dim W = 2$ und $W^\perp$ ist eine Gerade durch 0; für $v = v_1 + v_2$ mit eindeutigen Vektoren $v_1 \in W$ und $v_2 \in W^\perp$ ist $\pi_W(v) = v_1$ und wir haben zwei Möglichkeiten, um explizite Formeln zu erhalten:

(i) Bestimme eine ONB b_1, b_2 von W (z.B. mittels Gram-Schmidt-Verfahren), dann ist

$$\pi_W(v) = \langle b_1 \mid v\rangle\, b_1 + \langle b_2 \mid v\rangle\, b_2.$$

(ii) Beachte $\pi_W(v) = v_1 = v - v_2 = v - \pi_{W^\perp}(v)$. Ist also n ein Normalvektor von W, dann ist $\frac{1}{\|n\|}\, n$ eine ONB von $W^\perp$ und somit

$$\pi_W(v) = v - \langle \frac{1}{\|n\|}\, n \mid v\rangle\, \frac{1}{\|n\|}\, n = v - \frac{\langle n \mid v\rangle}{\langle n \mid n\rangle}\, n.$$

(Bemerkung: b_1, b_2, n ist eine ONB von $\mathbb{R}^3$.)

Einige wichtige Klassen von Matrizen

7.27. Definition: Für $A \in M(m,n;\mathbb{C})$ definieren wir die *zu A adjungierte Matrix* $A^+ = (b_{ij}) \in M(n,m;\mathbb{C})$ durch $b_{ij} := \overline{a_{ji}}$ $(i = 1,\dots,n,\ j = 1,\dots,m)$.

Das heißt $A^+ = \overline{A^t} = \overline{A}^t$ ist die komponentenweise konjugiert komplexe transponierte Matrix oder auch die transponierte der komponentenweise konjugiert komplexen Matrix; für relle Matrizen ist $A^+ = A^t$. (Es ist zwar weithin die Notation A^* für die Adjungierte im Sinne von 7.27 gebräuchlicher, aber wir verwenden im Rahmen dieser VO die $*$-Notation lieber nur für die dualen Abbildungen und Räume.)

7.28. Proposition: (i) $(A^+)^+ = A$.

(ii) $(A+B)^+ = A^+ + B^+$ und $(\lambda A)^+ = \overline{\lambda}\, A^+$.

(iii) $(A \cdot B)^+ = B^+ \cdot A^+$.

(iv) Ist A invertierbar, so auch A^+, und es gilt $(A^+)^{-1} = (A^{-1})^+$.

Beweis: Wie jener für 4.46 mit zusätzlicher komplexer Konjugation. □

7.29. Bemerkung: Das Standardskalarprodukt auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ lässt sich somit schreiben als

$$\langle x \mid y\rangle = \sum_{i=1}^n \overline{x_i}\, y_i = \begin{pmatrix} \overline{x_1} & \cdots & \overline{x_n} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x^+ \cdot y \quad (= x^t \cdot y \text{ für } \mathbb{R}^n).$$

7.30. Definition: Eine Matrix $A \in M(n,n;\mathbb{C})$ bzw. $A \in M(n,n;\mathbb{R})$ heißt

1. *unitär*, falls $A^+ \cdot A = A \cdot A^+ = I$ gilt (d.h. $A^+ = A^{-1}$)
   orthogonal, wenn A reell ist und $A^t \cdot A = A \cdot A^t = I$ gilt (d.h. $A^t = A^{-1}$)
2. *selbstadjungiert (hermitesch)*, falls $A^+ = A$ gilt
   selbstadjungiert (symmetrisch), wenn A reell ist und $A^t = A$ gilt
3. *antiselbstadjungiert (schiefhermitesch)*, falls $A^+ = -A$ gilt
   antiselbstadjungiert (schiefsymmetrisch), wenn A reell ist und $A^t = -A$ gilt
4. *normal*, falls $A^+ \cdot A = A \cdot A^+$ gilt
5. *nichtnegativ*, falls $A^+ = A$ und $\langle x \mid Ax\rangle \geq 0 \ \forall x \in \mathbb{C}^n$ gilt ($\langle . \mid . \rangle$ Standardskalarprodukt)
6. *positiv (definit)*, falls $A^+ = A$ und $\langle x \mid Ax\rangle > 0 \ \forall x \in \mathbb{C}^n \setminus \{0\}$ gilt ($\langle . \mid . \rangle$ Standardskalarprodukt)
7. *Orthogonalprojektion*[15], falls $A \cdot A = A$ und $A^+ = A$ gilt

7.31. Beispiele:

ad 1. Wir bestimmen alle (reellen) orthogonalen 2×2-Matrizen: Für $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ gilt

$$A^t \cdot A = I \iff \begin{pmatrix} \boxed{a \ \ b} \\ \boxed{c \ \ d} \end{pmatrix} \cdot \begin{pmatrix} \boxed{\begin{matrix} a \\ b \end{matrix}} & \boxed{\begin{matrix} c \\ d \end{matrix}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \iff \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \text{ ist ein ONS, also eine ONB!}$$

Daher können wir für passendes $\alpha \in [0, 2\pi[$ schreiben: $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$ und $\begin{pmatrix} c \\ d \end{pmatrix} = \pm \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$; also erhalten wir

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \quad \det A = 1,$$

was die schon aus Kapitel 4 bekannte Drehung um den Winkel α beschreibt, oder

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}, \quad \det A = -1,$$

was sich durch eine Skizze als Spiegelung an der Geraden mit Winkel $\frac{\alpha}{2}$ entlarven lässt.

Beispiele für unitäre 3×3-Matrizen sind $\begin{pmatrix} 0 & i & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -i \end{pmatrix}$ und $\frac{1}{4}\begin{pmatrix} \sqrt{3}+2 & \sqrt{3}-2 & -i\sqrt{2} \\ \sqrt{3}-2 & \sqrt{3}+2 & -i\sqrt{2} \\ \sqrt{2} & \sqrt{2} & 2i\sqrt{3} \end{pmatrix}$.

[15] Diese Bezeichnung ist eigentlich durch 7.24 schon vergeben, passt aber wegen 7.44 hier auch.

Die 3×3-Matrix $\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$ ist orthogonal und beschreibt eine Drehung um die z-Achse im $\mathbb{R}^3$.

ad 2. Die Matrix $\begin{pmatrix} 1 & 2+i \\ 2-i & -3 \end{pmatrix}$ ist selbstadjungiert. Beachte: Für die Diagonalelemente einer selbstadjungierten Matrix gilt $\overline{a_{ii}} = a_{ii}$, daher müssen also die a_{ii} reell sein.
Die Matrix $\begin{pmatrix} 2 & 7 \\ 7 & 5 \end{pmatrix}$ ist reell und selbstadjungiert.

ad 3. Die Matrix $\begin{pmatrix} 0 & 2+i \\ -2+i & 3i \end{pmatrix}$ ist antiselbstadjungiert. Beachte: Hier müssen die Diagonalelemente wegen $\overline{a_{ii}} = -a_{ii}$ rein imaginär sein.
Die Matrix $\begin{pmatrix} 0 & 4 \\ -4 & 0 \end{pmatrix}$ ist reell und antiselbstadjungiert.

ad 4. Alle Matrizen der Klassen 1.-3. und 5.-6. sind normal. Ein Beispiel für eine Matrix, die nicht normal ist, ist $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$: $A^t \cdot A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \neq \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = A \cdot A^t$.

ad 5. $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ ist nichtnegativ, weil selbstadjungiert und $\langle x \mid Ax \rangle = \overline{x_1}\, x_1 = |x_1|^2 \geq 0$.

ad 6. Die Matrix $\begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}$ ist positiv, weil die Formel in 7.3.2 ein Skalarprodukt definiert, das ja insbesondere positiv definit ist.

ad 7. Betrachte die Matrix A der Projektion auf die 1. Mediane im $\mathbb{R}^2$, d.h. $A \cdot e_1 = A \cdot e_2 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$, was die Matrix $A = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ ergibt.

Welche Bedeutung haben die obigen Arten von Matrizen jeweils bzw. wo treten sie beispielsweise auf?

ad 1. $x \mapsto A \cdot x$ lässt das Standardskalarprodukt und somit Längen und Winkel gleich — das sind also Drehungen und Drehspiegelungen im Fall von $\mathbb{R}^3$ bzw. gewisse unitäre 2×2-Matrizen für den Spin in der Quantenmechanik.

ad 2. Selbstadjungierte Matrizen (bzw. lineare Abbildungen oder Operatoren) spielen eine wesentliche Rolle bei der Behandlung von Kegelschnitten, als Trägheitstensor eines starren Körpers bei der Rotation und als Observable in der Quantenmechanik.

ad 3. Die elektromagnetische Feldstärke in der speziellen Relativitätstheorie lässt sich als schiefsymmetrische 4×4-Matrix auffassen (eigentlich ein schiefsymmetrischer Tensor zweiter Stufe).

ad 4. Die normalen Matrizen sind gerade die allgemeinsten „unitär diagonalisierbaren" Matrizen (siehe Kapitel 8).

ad 5. Sei $p\colon \mathbb{R}^2 \to \mathbb{R}$ eine rein quadratische Polynomfunktion (in zwei reellen Variablen), d.h. $p(x,y) = ax^2 + 2bxy + cx^2 = \langle \begin{pmatrix} x \\ y \end{pmatrix} \mid \begin{pmatrix} a & b \\ b & c \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \rangle$. Es gilt $p(x,y) \geq 0$ für alle (x,y) genau dann, wenn die Matrix $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ nichtnegativ ist.

ad 6. Die positiven Matrizen beschreiben die allgemeinste Form des Skalarprodukts auf $\mathbb{R}^n$ oder $\mathbb{C}^n$ (siehe 7.51).

ad 7. Die Bezeichnung „Orthogonalprojektion im $\mathbb{R}^n$ bzw. $\mathbb{C}^n$" wird durch 7.44 gerechtfertigt. Zum Teil ist dies schon klar, weil eine Projektion φ gemäß 7.23 die Eigenschaft $\varphi \circ \varphi = \varphi$ hat, daher für deren Matrix A also $A \cdot A = A$ gilt.

Die ganze Liste 7.30 gibt es auch im „abstrakten Fall" einer linearen Abbildung φ zwischen euklidischen bzw. unitären[16] Räumen.

7.32. Satz: Für jede unitäre Matrix A gilt $|\det A| = 1$; und $\det A = \pm 1$ für orthogonales A.

Beweis: $A^+ \cdot A = I \;\Rightarrow\; 1 = \det I = \det(A^+ \cdot A) = \det(\overline{A}^t \cdot A) = (\det \overline{A})(\det A) = (\overline{\det A})(\det A) = |\det A|^2$. Wenn A reell ist, dann auch $\det A$ und somit $\det A = \pm 1$. □

7.33. Satz: Für $A \in M(n,n;\mathbb{C})$ [bzw. $M(n,n;\mathbb{R})$] sind folgende Bedingungen äquivalent:

(i) A ist unitär [bzw. orthogonal].

(ii) Die Spalten von A bilden ein ONS bzgl. des Standardskalarprodukts (daher nach 7.9 also sogar eine ONB).

(iii) Die Zeilen von A bilden ein ONS bzgl. des Standardskalarprodukts (daher nach 7.9 also sogar eine ONB).

(iv) $\langle Ax \mid Ay \rangle = \langle x \mid y \rangle$ für alle Vektoren x, y. (Standardskalarprodukt)

(v) $\|Ax\| = \|x\|$ für alle Vektoren x. ($\|x\| = \sqrt{\langle x \mid x \rangle}$)

(vi) $\varphi_A\colon x \mapsto A \cdot x$ bildet jede ONB auf eine ONB ab.

(vii) $\varphi_A\colon x \mapsto A \cdot x$ bildet eine ONB auf eine ONB ab.

[16] Dafür muss eben noch φ^+ definiert werden und deshalb kam die Liste 7.30 auch erst in diesem Kapitel vor und nicht schon in Kapitel 4.

Beweis: Bevor wir loslegen, veranschaulichen wir zunächst die gesamte Implikationsstruktur dieses Äquivalenzbeweises im folgenden Diagramm

$$\begin{tikzcd}
(iii) \arrow[r, Leftrightarrow] & (i) \arrow[rr, Rightarrow] \arrow[d, Leftrightarrow] & & (iv) \arrow[r, Leftrightarrow] \arrow[d, Rightarrow] & (v) \\
 & & (vii) \arrow[ul, Rightarrow] & & \\
 & (ii) & & (vi) \arrow[ll, Rightarrow] \arrow[ul, Rightarrow] &
\end{tikzcd}$$

Nun zu den einzelnen Verbindungsstücken im obigen Diagramm:

(i) $\Rightarrow$ (iv): $\langle Ax \mid Ay\rangle = (Ax)^+ Ay = x^+ A^+ Ay = x^+ I y = x^+ y = \langle x \mid y\rangle$.

(iv) $\Leftrightarrow$ (v): ($\Rightarrow$) $\|Ax\|^2 = \langle Ax \mid Ax\rangle = \langle x \mid x\rangle = \|x\|^2$.

($\Leftarrow$) in $\mathbb{C}^n$ gilt: $\langle x \mid y\rangle = \frac{1}{4}\left(\|x+y\|^2 - i\|x+iy\|^2 - \|x-y\|^2 + i\|x-iy\|^2\right)$,

in $\mathbb{R}^n$ gilt: $\langle x \mid y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$ *Polarisierungsidentitäten* (ohne Bew.).

Laut (v) ändert A nichts an $\|.\|$, daher auch an $\langle . \mid . \rangle$ nichts.

(iv) $\Rightarrow$ (vi): $\langle b_i \mid b_j\rangle = \delta_{ij} \Rightarrow \langle Ab_i \mid Ab_j\rangle \overset{\text{(iv)}}{=} \delta_{ij}$.

(vi) $\Rightarrow$ (ii): $e_1, \ldots, e_n$ ist ONB, daher auch $Ae_1, \ldots, Ae_n$; letzteres sind aber die Spalten von A.

(i) $\Leftrightarrow$ (ii): $A^+ \cdot A = I \iff \begin{pmatrix} \vdots \\ \boxed{\quad s_i^+ \quad} \\ \vdots \end{pmatrix} \cdot \begin{pmatrix} \cdots & \boxed{s_j} & \cdots \end{pmatrix} = (\delta_{ij})$;

wegen 4.32 folgt aus $A^+ \cdot A = I$ auch schon $A \cdot A^+ = I$ automatisch.

(i) $\Leftrightarrow$ (iii): $A \cdot A^+ = I \iff \begin{pmatrix} \vdots \\ \boxed{\quad z_i \quad} \\ \vdots \end{pmatrix} \cdot \begin{pmatrix} \cdots & \boxed{z_j^+} & \cdots \end{pmatrix} = (\delta_{ij})$;

wegen 4.32 folgt aus $A \cdot A^+ = I$ auch schon $A^+ \cdot A = I$ automatisch.

(vi) $\Rightarrow$ (vii): ist klar.

(vii) $\Rightarrow$ (i): Sei $b_1, \ldots, b_n$ eine ONB, sodass auch $c_1 = A \cdot b_1, \ldots, c_n = A \cdot b_n$ eine ONB ist; setze

$$B := \begin{pmatrix} \cdots & \boxed{b_i} & \cdots \end{pmatrix} \quad \text{und} \quad C := \begin{pmatrix} \cdots & \boxed{c_i} & \cdots \end{pmatrix};$$

beide Matrizen sind wegen der bereits bewiesenen Äquivalenz (i) $\Leftrightarrow$ (ii) unitär bzw. orthogonal. Nach Konstruktion gilt $A \cdot B = C$, daher $A = C \cdot B^{-1} = C \cdot B^+$ und weiter $A^+ A = (CB^+)^+ CB^+ = BC^+ CB^+ = BIB^+ = I$ und wegen 4.32 gilt somit auch $AA^+ = I$, also ist A unitär. $\square$

Bemerkung: 7.33(v) besagt, dass die Abbildung $x \mapsto A \cdot x$ längentreu ist und 7.33(iv) sogar, dass sie längen- und winkeltreu ist, denn es gilt ja

$$\|x\| = \sqrt{\langle x \mid x \rangle} \quad \text{und} \quad \cos\measuredangle(x, y) = \frac{\langle x \mid y \rangle}{\sqrt{\langle x \mid x \rangle \langle y \mid y \rangle}}.$$

7.34. Einige Matrizengruppen: Bezüglich der Matrizenmultiplikation sind folgende Teilmengen der quadratischen Matrizen jeweils Gruppen:[17]

$GL(n,\mathbb{K}) = \{A \in M(n,n;\mathbb{K}) \mid \det A \neq 0\}$ .... allgemeine lineare Gruppe der Ordnung n

$SL(n,\mathbb{K}) = \{A \in M(n,n;\mathbb{K}) \mid \det A = 1\}$ .... spezielle lineare Gruppe (der Ordnung n)

$U(n) = \{U \in M(n,n;\mathbb{C}) \mid U \text{ unitär}\}$ .... unitäre Gruppe der Ordnung n

$O(n) = \{A \in M(n,n;\mathbb{R}) \mid A \text{ orthogonal}\}$ .... orthogonale Gruppe (der Ordnung n)

$SO(n) = \{A \in M(n,n;\mathbb{R}) \mid A \text{ orthogonal und } \det A = 1\}$ .... spezielle orthogonale Gruppe

Ein weiteres Beispiel, nämlich die Lorentzgruppe $O(3,1)$ findet sich in 7.54.

Nach 7.31.ad 1. ist $O(2) = \{\begin{pmatrix} \cos\alpha & \mp\sin\alpha \\ \sin\alpha & \pm\cos\alpha \end{pmatrix}\}$, besteht also aus den Drehungen und Spiegelungen, und $SO(2) = \{\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\}$, besteht also aus den Drehungen.

Vor allem in der theoretischen Physik spielen die Linearformen $f \in V^* = L(V,\mathbb{K})$ auf Vektorräumen V eine wichtige Rolle. Ist V ein endlichdimensionaler euklidischer oder unitärer VR [oder allgemeiner ein Hilbertraum], dann können wir in gewisser Hinsicht die Linearformen (und damit den Dualraum) aus der Theorie „rausschmeißen"; die Vektoren selbst zusammen mit dem Skalarprodukt übernehmen nämlich die Rolle der Linearformen, und zwar so:

7.35. Sei V ein endlichdimensionaler euklidischer oder unitärer VR und $f \in V^*$. Wir wählen eine ONB $b_1, \ldots, b_n$ in V und erhalten gemäß 5.10 die dazu duale Basis $f_1, \ldots, f_n$ in V^*.
Sei $f = \sum\limits_{j=1}^{n} \mu_j f_j$ die entsprechende Basisdarstellung von f und setze $v_f := \sum\limits_{j=1}^{n} \overline{\mu_j}\, b_j \in V$.

Für $v = \sum \lambda_i b_i \in V$ gilt dann $f(v) = (\sum \mu_j f_j)(\sum \lambda_i b_i) = \sum \mu_j \lambda_i f_j(b_i) = \sum \mu_i \lambda_i$
und $\langle v_f \mid v \rangle = \langle \sum \overline{\mu_j}\, b_j \mid \sum \lambda_i b_i \rangle = \sum \overline{\overline{\mu_j}}\, \lambda_i \langle b_j \mid b_i \rangle = \sum \mu_i \lambda_i$, also

$(\star)$
$$\boxed{f(v) = \langle v_f \mid v \rangle}$$

[17] Hier ohne Beweise, die aber sehr leicht zu erbringen sind.

Die folgende Proposition zeigt, dass v_f durch $(\star)$ eindeutig bestimmt ist:

7.36. Proposition: Sei V ein euklidischer oder unitärer VR und $w_1, w_2 \in V$, dann gilt:

$$\langle w_1 \mid v\rangle = \langle w_2 \mid v\rangle \;\; \forall v \in V \quad \Longrightarrow \quad w_1 = w_2.$$

Beweis: Für beliebiges $v \in V$ gilt $\langle w_1 - w_2 \mid v\rangle = 0$; setzen wir speziell $v = w_1 - w_2$ ein, dann folgt wegen Eigenschaft (S4) des Skalarprodukts sofort $w_1 - w_2 = 0$. □

7.37. Bemerkung: (i) Da v_f in $(\star)$ also gemäß 7.36 eindeutig bestimmt ist, hätten wir auch bei Verwendung einer anderen ONB dasselbe v_f erhalten.

(ii) Die Abbildung $\gamma\colon f \mapsto v_f$, $\gamma(\sum \mu_j f_j) = \sum \overline{\mu_j}\, b_j$, ist offenbar bijektiv $V^* \to V$ (und wegen (i) auch unabhängig von der Wahl der ONB $B = (b_1, \ldots, b_n)$ von V; die Basis $f_1, \ldots, f_n$ ist ja durch B als deren duale Basis bestimmt). Die Gleichung $(\star)$ schreibt sich mit Hilfe der Abbildung γ nun so:

$$\boxed{f(v) = \langle \gamma(f) \mid v\rangle}$$

Wir haben also $f \in V^*$ durch $\gamma(f) \in V$ und $\langle . \mid . \rangle$ ersetzt. Die konkrete Gestalt von γ wollen wir in einigen Beispielen anschauen:

1. Im $\mathbb{R}^n$ mit dem Standardskalarprodukt ist $\gamma\colon (\mathbb{R}^n)^* \to \mathbb{R}^n$ gegeben durch

$$\gamma((a_1, \ldots, a_n)) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix},$$

denn $f \in (\mathbb{R}^n)^*$ ist durch einen Zeilenvektor $(a_1, \ldots, a_n)$ gegeben und

$$f(x) = (a_1, \ldots, a_n) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + \ldots a_n x_n = \langle \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mid \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \rangle = \langle \gamma(f) \mid x\rangle.$$

2. Ähnlich ist für $\mathbb{C}^n$ mit dem Standardskalarprodukt $\gamma\colon (\mathbb{C}^n)^* \to \mathbb{C}^n$ gegeben durch

$$\gamma((a_1, \ldots, a_n)) = \begin{pmatrix} \overline{a_1} \\ \vdots \\ \overline{a_n} \end{pmatrix}.$$

3. In $V = \mathbb{R}^2$ mit dem Skalarprodukt $\langle x \mid y\rangle = x^t \cdot A \cdot y$, wobei $A = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}$ ist (vgl. 7.3.2), erhalten wir

$$\gamma((a_1, a_2)) = \big((a_1 \;\; a_2) \cdot A^{-1}\big)^t = (A^{-1})^t \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 5a_1 + 2a_2 \\ 2a_1 + a_2 \end{pmatrix},$$

denn $f(x) = (a_1 \;\; a_2) \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (a_1 \;\; a_2) \cdot (A^{-1} \cdot A) \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \big((a_1 \;\; a_2) \cdot A^{-1}\big) \cdot A \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$$= \langle \big((a_1 \;\; a_2) \cdot A^{-1}\big)^t \mid \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \rangle = \langle \gamma(f) \mid x\rangle.$$

Jetzt schmeißen wir auch noch die dualen Abbildungen $\varphi^*\colon W^* \to V^*$ aus der Theorie raus und ersetzen sie durch passende Abbildungen $\varphi^+\colon W \to V$: Seien V und W euklidische oder unitäre VR und $\varphi\colon V \to W$ linear. Gemäß 5.11 ist die duale Abbildung gegeben durch $\varphi^*(g) = g \circ \varphi$ für jedes $g \in W^*$, d.h. $\varphi^*(g)(v) = g(\varphi(v))$ für jedes $v \in V$; als Diagramm

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[dr, "g\circ\varphi"'] & W \arrow[d, "g"] \\ & \mathbb{K} \end{tikzcd}$$

Nun definieren wir $\boxed{\varphi^+ := \gamma \circ \varphi^* \circ \gamma^{-1}}$ bzw. in Diagrammform

$$\begin{tikzcd} V^* \arrow[d, "\gamma"'] & W^* \arrow[l, "\varphi^*"'] \arrow[d, "\gamma"] \\ V & W \arrow[l, "\varphi^+"] \end{tikzcd}$$

Eine explizite Formel für φ^+ ohne γ erhalten wir so:

$$\begin{aligned} \langle \varphi^+(x) \mid v\rangle = \langle \gamma(\varphi^*(\gamma^{-1}(w))) \mid v\rangle &\overset{[7.37]}{=} \varphi^*(\gamma^{-1}(w))(v) \overset{[5.11]}{=} \gamma^{-1}(w)(\varphi(v)) \\ &\overset{[7.37]}{=} \langle \gamma(\gamma^{-1}(w)) \mid \varphi(v)\rangle = \langle w \mid \varphi(v)\rangle, \end{aligned}$$

d.h. zusammengefasst $\boxed{\langle \varphi^+(w) \mid v\rangle = \langle w \mid \varphi(v)\rangle}$. Daher können wir jetzt φ^+ sogar ohne Zuhilfenahme von φ^* definieren.[18]

7.38. Definition: Seien V und W euklidische oder unitäre VR und $\varphi\colon V \to W$ linear. Sei $w \in W$. Die Abbildung $f\colon v \mapsto \varphi(v) \mapsto \langle w \mid \varphi(v)\rangle$ ist linear, also eine Linearform auf V. Nach 7.35 und 7.36 gibt es ein eindeutig bestimmtes $v_0 \in V$ mit $f(v) = \langle v_0 \mid v\rangle$ für alle $v \in V$, d.h. $\langle w \mid \varphi(v)\rangle = \langle v_0 \mid v\rangle$; wir definieren $\varphi^+(w) := v_0$ und erhalten dadurch eine Abbildung $\varphi^+\colon W \to V$, die durch auch die Gleichung

$$\boxed{\langle w \mid \varphi(v)\rangle = \langle \varphi^+(w) \mid v\rangle}$$

eindeutig bestimmt ist und *adjungierte Abbildung* zu φ heißt.

7.39. Proposition: φ^+ ist linear

Beweis: Seien $w_1, w_2 \in W$. Für jedes $v \in V$ gilt $\langle \varphi^+(w_1 + w_2) \mid v\rangle = \langle w_1 + w_2 \mid \varphi(v)\rangle = \langle w_1 \mid \varphi(v)\rangle + \langle w_2 \mid \varphi(v)\rangle = \langle \varphi^+(w_1) \mid v\rangle + \langle \varphi^+(w_2) \mid v\rangle = \langle \varphi^+(w_1) + \varphi^+(w_2) \mid v\rangle$ und daher wegen 7.36 also $\varphi^+(w_1 + w_2) = \varphi^+(w_1) + \varphi^+(w_2)$.

Ist $w \in W$ und $\lambda \in \mathbb{R}$ (bzw. $\mathbb{C}$), dann gilt $\langle \varphi^+(\lambda w) \mid v\rangle = \langle \lambda w \mid \varphi(v)\rangle = \overline{\lambda}\langle w \mid \varphi(v)\rangle = \overline{\lambda}\langle \varphi^+(w) \mid v\rangle = \langle \lambda\varphi^+(w) \mid v\rangle$ für jedes $v \in V$, daher folgt $\varphi^+(\lambda w) = \lambda\varphi^+(w)$ aus 7.36. □

Wir geben in den folgenden Punkten ohne Beweise einige Eigenschaften der adjungierten Abbildung an, geben das Verhältnis zu den entsprechenden Matrizen an und liefern auch die abstrakten Begriffe passend zu jenen in 7.30 nach. (Die ausgelassenen Beweise sind relativ einfach zu führen und finden sich in den meisten Lehrbüchern.)

[18] Was zumindest jenen zugute kommt, die „sich mit dem Begriff duale Abbildung auf Kriegsfuß befinden".

7.40. Proposition: Seien V, W, Z endlichdimensionale euklidische oder unitäre VR und die Abbildungen $\varphi, \psi\colon V \to W$ sowie $\sigma\colon W \to Z$ linear. Dann gilt:

(i) $(\varphi^+)^+ = \varphi$

(ii) $(\varphi + \psi)^+ = \varphi^+ + \psi^+$, $(\lambda\varphi)^+ = \overline{\lambda}\varphi^+$

(iii) $(\sigma \circ \varphi)^+ = \varphi^+ \circ \sigma^+$, ${\mathrm{id}_V}^+ = \mathrm{id}_V$

(iv) Ist φ bijektiv, dann auch φ^+ und es gilt $(\varphi^+)^{-1} = (\varphi^{-1})^+$.

7.41. Spezialfall: Für $V = \mathbb{K}^n$ und $W = \mathbb{K}^m$ mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und mit dem Standardskalarprodukt betrachten wir eine durch eine Matrix $A \in M(m, n; \mathbb{K})$ gegebene lineare Abbildung $\varphi = \varphi_A\colon x \mapsto A \cdot x$. Wie sieht die Matrix von φ_A^+ aus?

Für jedes y gilt $\langle \varphi_A^+(x) \mid y\rangle = \langle x \mid \varphi_A(y)\rangle = x^+ \cdot (A \cdot y) = (x^+ \cdot A) \cdot y = (A^+ \cdot x)^+ \cdot y = \langle A^+ \cdot x \mid y\rangle$, daher erhalten wir $\boxed{\varphi_A^+(x) = A^+ \cdot x}$ für $\mathbb{K} = \mathbb{C}$ bzw. $\boxed{\varphi_A^+(x) = A^t \cdot x}$ für $\mathbb{K} = \mathbb{R}$. Die adjungierte Abbildung ist hier also durch die adjungierte Matrix gegeben. Das gilt bezügich ONB (!) auch im „abstrakten" Fall:

7.42. Satz: Seien V und W euklidische oder unitäre VR und $\varphi\colon V \to W$ linear. Sei B eine ONB in V, C eine ONB in W. Dann gilt:

$$[\varphi^+]_{B,C} = {[\varphi]_{B,C}}^+.$$

Das abstrakte Gegenstück zu 7.30 ist

7.43. Definition: Seien V und W euklidische oder unitäre VR.

1. Eine lineare Abbildung $\varphi\colon V \to W$ heißt *unitär* bzw. *orthogonal*, falls $\varphi^+ \circ \varphi = \mathrm{id}_V$, $\varphi \circ \varphi^+ = \mathrm{id}_W$ (d.h. $\varphi^+ = \varphi^{-1}$).

Eine lineare Abbildung $\varphi\colon V \to V$ heißt

2. *selbstadjungiert (hermitesch)*, falls $\varphi^+ = \varphi$,
3. *antiselbstadjungiert (schiefhermitesch)*, falls $\varphi^+ = -\varphi$,
4. *normal*, falls $\varphi^+ \circ \varphi = \varphi \circ \varphi^+$,
5. *nichtnegativ*, falls $\varphi^+ = \varphi$ und $\langle v \mid \varphi(v)\rangle \geq 0 \ \forall\, v \in V$,
6. *positiv*, falls $\varphi^+ = \varphi$ und $\langle v \mid \varphi(v)\rangle > 0 \ \forall\, v \in V \setminus \{0\}$.

7.44. Proposition: φ Orthogonalprojektion $\iff$ $\varphi \circ \varphi = \varphi$ und $\varphi^+ = \varphi$.

Wir fügen dies quasi als 7. Eigenschaft zur Liste in 7.43 hinzu.

7.45. Aufgrund von 7.41 gilt für $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ mit dem Standardskalarprodukt und $j = 1, \ldots, 7$:

$\boxed{\varphi_A\colon x \mapsto A \cdot x \text{ hat die Eigenschaft Nr. } j \text{ aus 7.43} \iff A \text{ hat die Eigenschaft Nr. } j \text{ aus 7.30}}$

Aufgrund von 7.42 gilt für euklidische/unitäre VR V, W mit ONB B, C:

$\boxed{\varphi\colon V \to W \text{ hat die Eigenschaft Nr. 1 aus 7.43} \iff [\varphi]_{B,C} \text{ hat die Eigenschaft Nr. 1 aus 7.30}}$

und für $j = 2, \ldots, 7$

$\boxed{\varphi\colon V \to V \text{ hat die Eigenschaft Nr. } j \text{ aus 7.43} \iff [\varphi]_{B,B} \text{ hat die Eigenschaft Nr. } j \text{ aus 7.30}}$

7.46. Bemerkung: Für Eigenschaft 5 und 6 in 7.43 benötigt man auch 7.13(ii), nämlich gleich am Beginn der folgenden Rechnung:

$\langle v \mid \varphi(v)\rangle = \sum \overline{\langle b_i \mid v\rangle}\,\langle b_i \mid \varphi(v)\rangle = \langle [v]_B \mid [\varphi(v)]_B\rangle \overset{[4.12]}{=} \langle [v]_B \mid [\varphi]_{B,B}\cdot [v]_B\rangle \ldots$ entspricht $\langle x \mid Ax\rangle$ in 7.30.

Das Pendant zu 7.33 lautet nun so:

7.47. Satz: Seien V und W euklidische oder unitäre VR und $\varphi\colon V \to W$ linear. Dann sind folgende Eigenschaften äquivalent:

(i) φ ist unitär [bzw. orthogonal].

(ii) $\langle \varphi(v) \mid \varphi(w)\rangle = \langle v \mid w\rangle$ für alle $v \in V$ und $w \in W$ und $\dim V = \dim W$.

(iii) $\|\varphi(v)\| = \|v\|$ für alle alle $v \in V$ und $\dim V = \dim W$. $(\|.\| = \sqrt{\langle . \mid . \rangle})$

(iv) φ bildet jede ONB auf eine ONB ab.

(v) φ bildet eine ONB auf eine ONB ab.

7.48. In der Physik wird oft folgende Terminologie verwendet:

Vektor u	$\lvert u\rangle$	ket-Vektor
Linearform $v \mapsto \langle w \mid v\rangle$	$\langle w\rvert$	bra- Vektor
lineare Abbildung		(linearer) Operator
lineare Abbildung $v \mapsto \langle w \mid v\rangle\, u$		$\lvert u\rangle\langle w\rvert$

Ist $b_1,\ldots,b_n$ eine ONB, dann bedeutet die Gleichung

$$I = \sum_{i=1}^{n} |b_i\rangle\langle b_i|, \quad \text{d.h.} \quad |v\rangle = I|v\rangle = \sum_{i=1}^{n} |b_i\rangle\langle b_i|v\rangle \quad \forall v,$$

(unter Beachtung von 7.13(i)) nichts anderes als

$$v = Iv = \sum_{i=1}^{n} \langle b_i \mid v\rangle\, b_i \quad \forall v.$$

Analog können wir die Orthogonalprojektion P auf $W := [b_1,\ldots,b_k]$ mit 7.25 schreiben als

$$P = \sum_{i=1}^{k} |b_i\rangle\langle b_i|.$$

Falls für den Operator A gilt, dass mit passenden Skalaren λ_i die Gleichungen $Ab_i = \lambda_i b_i$ ($i = 1,\ldots,n$) gelten (d.h. b_i ist Eigenvektor zum Eigenwert λ_i; siehe Kapitel 8), so folgt $Av = A(\sum\langle b_i \mid v\rangle b_i) = \sum\langle b_i \mid v\rangle Ab_i = \sum\langle b_i \mid v\rangle\lambda_i b_i$, also

$$A = \sum_{i=1}^{n} \lambda_i\, |b_i\rangle\langle b_i|.$$

Für eine kurze Behandlung der Raum-Zeit-Metrik, wie sie in der speziellen Relativitätstheorie verwendet wird, müssen wir allgemeinere bilineare Abbildungen in den Grundkörper betrachten als die Skalarprodukte aus 7.1. Es sei also $\mathbb{K}$ vorübergehend wieder ein beliebiger Körper.

7.49. Definition: Es seien V, W, Z VR über $\mathbb{K}$. Eine Abbildung $\beta\colon V \times W \to Z$ heißt *bilinear*, wenn gilt:

β ist linear im 1. Argument, d.h. $\beta(v_1 + v_2, w) = \beta(v_1, w) + \beta(v_2, w)$ und $\beta(\lambda v, w) = \lambda\beta(v, w)$,

und

β ist linear im 2. Argument, d.h. $\beta(v, w_1 + w_2) = \beta(v, w_1) + \beta(v, w_2)$ und $\beta(v, \lambda w) = \lambda\beta(v, w)$.

Eine *Bilinearform* auf $V \times W$ ist eine bilineare Abbildung $V \times W \to \mathbb{K}$.

Analog werden *multilineare (n-lineare)* Abbildungen $V_1 \times \ldots \times V_n \to Z$ und *Multilinearformen* $V_1 \times \ldots \times V_n \to \mathbb{K}$ definiert.

Ein Skalarprodukt auf einem reellen VR V ist demnach eine positiv definite symmetrische Bilinearform auf $V \times V$. Die Determinante ist eine „alternierende“ [d.h. Platztausch kostet ein Minus] Multilinearform auf $\underbrace{\mathbb{K}^n \times \ldots \times \mathbb{K}^n}_{n \text{ mal}}$ (n Spaltenvektoren einer $n \times n$-Matrix).

7.50. (i) Bilinearformen auf $\mathbb{K}^n \times \mathbb{K}^n$: Sei $\beta\colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ bilinear, dann gilt:

$$\beta(x, y) = \beta(\sum_i x_i e_i, \sum_j y_j e_j) = \sum_i \sum_j x_i y_j \underbrace{\beta(e_i, e_j)}_{b_{ij}} = x^t \cdot B \cdot y.$$

Jede Bilinearform sieht also so aus, und umgekehrt stellt jede derartige Formel eine Bilinearform dar (vgl. 4.3). [Ein analoger Beweis zeigt übrigens , dass die Sesquilinearformen auf $\mathbb{C}^n \times \mathbb{C}^n$ immer von der Form $\gamma(x, y) = x^+ \cdot C \cdot y$ sind.]

(ii) Matrix einer Bilinearform: Seien V und W VR über $\mathbb{K}$, $\beta\colon V \times W \to \mathbb{K}$ bilinear und $B = (b_1, \ldots, b_m)$ bzw. $C = (c_1, \ldots, c_n)$ eine Basis von V bzw. W. Mittels Basisdarstellungen von $v \in V$ und $w \in W$ erhalten wir

$$\beta(v, w) = \beta(\sum_i \lambda_i b_i, \sum_j \mu_j c_j) = \sum_i \sum_j \lambda_i \mu_j \underbrace{\beta(b_i, c_j)}_{b_{ij}} = [v]_B^t \cdot B \cdot [w]_C,$$

wobei wir $[\beta]_{B,C} := B$ die Matrix von β bzgl. B, C nennen.

(iii) Verhalten bei Basiswechsel: Betrachte die Situation wie in (ii), wobei zusätzlich weitere Basen $B' = (b'_1, \ldots, b'_m)$, $C' = (c'_1, \ldots, c'_n)$ von V, W gegeben seien. Es sei weiters $S = [B \to B']$ und $T = [C \to C']$, d.h. $[v]_B = S \cdot [v]_{B'}$ und $[w]_C = T \cdot [w]_{C'}$. Dann folgt

$$\beta(v, w) = [v]_B^t \cdot [\beta]_{B,C} \cdot [w]_C = [v]_{B'}^t \cdot S^t \cdot [\beta]_{B,C} \cdot T \cdot [w]_{C'}$$

und der Vergleich mit $\beta(v, w) = [v]_{B'}^t \cdot [\beta]_{B',C'} \cdot [w]_{C'}$ liefert $\boxed{[\beta]_{B',C'} = S^t \cdot [\beta]_{B,C} \cdot T}$ (vgl. 4.39).

7.51. Bemerkung: Aus 7.50(i) ist ersichtlich, wie alle Skalarprodukte auf $\mathbb{R}^n$ [$\mathbb{C}^n$] aussehen: (S1)-(S2), also die Bi[Sesqui]linearität, ist zur Gestalt $x^t \cdot B \cdot y$ [$x^+ \cdot B \cdot y$] äquivalent; (S4) ist dazu äquivalent, dass B positiv ist; und (S3) dazu, dass B symmetrisch [selbstadjungiert] ist: (S3) impliziert $b_{ji} = \langle e_j \mid e_i \rangle = \overline{\langle e_i \mid e_j \rangle} = \overline{b_{ij}}$; ist andererseits B symmetrisch [selbstadjungiert], d.h. $B^+ = B$, so folgt $\overline{\langle y \mid x \rangle} = \underbrace{\overline{y^+ Bx} = \overline{(y^+ Bx)^t}}_{y^+Bx=(y^+Bx)^t \in \mathbb{K}!} = (y^+ Bx)^+ = x^+ By = \langle x \mid y \rangle$.

7.52. Definition: Der Minkowski-Raum (oder die 4-dimensionale Raum-Zeit) ist der Vektorraum $\mathbb{R}^4 = \{\begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{pmatrix} \mid u_i \in \mathbb{R}\ (i = 0,1,2,3)\} = \{\begin{pmatrix} t \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid t \in \mathbb{R}, x_1, x_2, x_3 \in \mathbb{R}\}$ zusammen mit der Bilinearform $\mu\colon \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ als *Raum-Zeit-Metrik*, gegeben durch[19]

$$\mu(u,w) := -u_0 w_0 + u_1 w_1 + u_2 w_2 + u_3 w_3 = u^t \cdot \underbrace{\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{E=(\eta_{ij})} \cdot w.$$

Die Raum-Zeit-Metrik ist zwar symmetrisch und bilinear, aber nicht positiv definit (z.B. ist $\mu(e_0, e_0) = -1$), also kein Skalarprodukt im Sinne von 7.1 (bloß seine Einschränkung auf den Teilraum mit $t = 0$ erfüllt (S4)).

Auf dem Teilraum der Vektoren $\begin{pmatrix} 0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$ kann μ wie üblich zur Bestimmung von Längen und Winkeln verwendet werden, insgesamt auf $\mathbb{R}^4$ hat es jedoch eine andere Funktion, z.B. dient es zur Berechnung der Eigenzeit von bewegten Beobachtern.

Obzwar (S4) nicht erfüllt ist, gilt für μ eine wichtige (etwas schwächere) Eigenschaft:

7.53. Definition: Eine Bilinearform $\beta\colon V \times W \to \mathbb{K}$ heißt *nicht ausgeartet*, wenn gilt:

$$\text{(NA)} \quad \begin{array}{lll} \forall u \neq 0 & \exists w \neq 0: & \beta(u,w) \neq 0 \\ \forall w \neq 0 & \exists u \neq 0: & \beta(u,w) \neq 0 \end{array}$$

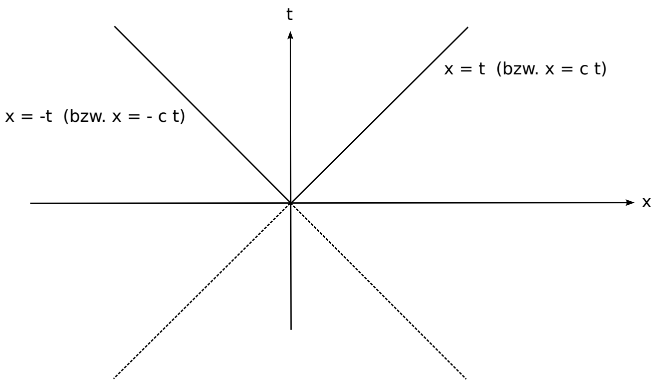
Im endlichdimensionalen Fall ist das äquivalent zur Invertierbarkeit von $[\beta]_{B,C}$ (daher folgt dann insbesondere $\dim V = \dim W$).

Aus (S4) folgt stets (NA) (vgl. 7.36): zu $u \neq 0$ wähle $w := u$, dann ist $\beta(u,u) > 0$; zu gegebenem $w \neq 0$ wähle analog $u := w$.
Auch mit (NA) anstatt (S4) kann man „Dualräume aus der Theorie rausschmeißen“!

---

[19] Hier bequemer Weise mit Lichtgeschwindigkeit $c = 1$, sonst wäre $\mu(u,w) := -c^2 u_0 w_0 + u_1 w_1 + u_2 w_2 + u_3 w_3$.

Die Raum-Zeit-Metrik μ erfüllt (NA): zu gegebenem $u = \begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{pmatrix} \neq 0$ wähle $w := \begin{pmatrix} -u_0 \\ u_1 \\ u_2 \\ u_3 \end{pmatrix}$, dann ist $\mu(u, w) = \sum u_i^2 \neq 0$.



Zweidimensionale Skizze für Ausbreitung eines Lichtblitzes zu $t = 0$ im Ursprung.

Für Vektoren u aus der Kegelfläche mit der Gleichung $\|x\|^2 = t^2$ im $\mathbb{R}^4$ gilt $\mu(u, u) = 0$; diese u heißen *lichtartig* und definieren den sogenannten *Lichtkegel*.

Für Vektoren u im Inneren des Doppelkegels gilt $\mu(u, u) < 0$; diese u heißen *zeitartig*. Es gibt dann ein Intertialsystem, in dem 0 und u am gleichen Ort (zu verschiedenen Zeiten) sind.

Für Vektoren u außerhalb des Doppelkegels gilt $\mu(u, u) > 0$; diese u heißen *raumartig*. Es gibt dann ein Inertialsystem, in dem 0 und u gleichzeitig (an verschiedenen Orten) sind.

7.54. Definition: $O(3,1) := \{L \in M(4,4;\mathbb{R}) \mid \mu(Lu, Lv) = \mu(u, v)\ \forall u, v \in \mathbb{R}^4\}$ ist die sogenannte *Lorentz-Gruppe*.

Die Bedingung $\mu(Lu, Lv) = \mu(u, v)$ für alle $u, v \in \mathbb{R}^4$ ist äquivalent zu $(L \cdot u)^t \cdot E \cdot (L \cdot v) = u^t \cdot E \cdot v$ für alle $u, v \in \mathbb{R}^4$; und letztere ist wiederum gleichbedeutend mit

$$(\star) \qquad L^t \cdot E \cdot L = E,$$

was sich gut mit der Bedingung $O^t \cdot I \cdot O = O^t \cdot O = I$ für orthogonale Matrizen vergleichen lässt.

Die Elemente der Lorentz-Gruppe heißen *Lorentztransformationen*. Sie beschreiben (als Matrizen von Basiswechsel im $\mathbb{R}^4$) den Übergang von einem Koordinatensystem zu einem zweiten, das sich relativ zum ersten geradlinig gleichförmig bewegt. Die Bedingung $(\star)$ ergibt, dass E (und somit μ) dabei ungeändert bleibt (denn an und für sich geht ja E nach 7.50(iii) in $L^t EL$

über, wenn ein Basiswechsel mit der Matrix L durchgeführt wird.) Die Invarianz von E bedeutet die Invarianz der Lichtgeschwindigkeit (siehe unten).

Ein typisches Element der Lorentz-Gruppe (Bewegung in x-Richtung mit Geschwindigkeit v) sieht so aus:

$$L = \begin{pmatrix} \gamma & \gamma v & 0 & 0 \\ \gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{mit } \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \overset{[c=1]}{=} \frac{1}{\sqrt{1-v^2}};$$

es ist $L \cdot [u]_{\text{bewegt}} = [u]_{\text{ruhend}} = u$.

Wir rechnen nach, dass L zu $O(3,1)$ gehört: Damit μ invariant bleibt, muss $L^t E L = E$ gelten; wir müssen das nur in den linken oberen 2×2-Kästchen nachprüfen und berechen

$$\begin{pmatrix} \gamma & \gamma v \\ \gamma v & \gamma \end{pmatrix}^t \cdot \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \gamma & \gamma v \\ \gamma v & \gamma \end{pmatrix} = \begin{pmatrix} \gamma & \gamma v \\ \gamma v & \gamma \end{pmatrix} \cdot \begin{pmatrix} -\gamma & -\gamma v \\ \gamma v & \gamma \end{pmatrix} = \begin{pmatrix} -\gamma^2 + \gamma^2 v^2 & -\gamma^2 v + \gamma^2 v \\ -\gamma^2 v + \gamma^2 v & -\gamma^2 v^2 + \gamma^2 \end{pmatrix}$$
$$= \begin{pmatrix} -\gamma^2 (1-v^2) & 0 \\ 0 & \gamma^2(1-v^2) \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Insbesondere gilt also: Aus $\mu([u]_{\text{ruhend}}, [u]_{\text{ruhend}}) = 0$ folgt auch $\mu([u]_{\text{bewegt}}, [u]_{\text{bewegt}}) = 0$. Somit ist auch für den bewegten Beobachter die Lichtgeschwindigkeit 1 (bzw. c).

§8. EIGENWERTE

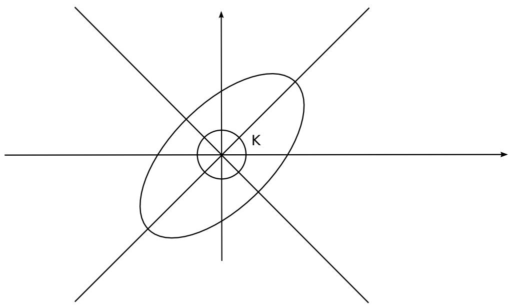
8.1. Motivation: Wenn wir für eine gegebene quadratische Matrix $A \in M(n,n;\mathbb{K})$ eine Basis $B = (b_1,\dots,b_n)$ von $\mathbb{K}^n$ finden können, sodass jedes b_i auf ein Vielfaches seiner selbst abgebildet wird (d.h. $Ab_i = \lambda_i b_i$), dann hat das eine Reihe von Vorteilen:

(i) Die Matrix von $\varphi_A\colon x \mapsto A\cdot x$ bzgl. B, gegeben gemäß 4.39 durch $S^{-1}\cdot A\cdot S$ mit Basiswechselabbildung $S = [E \to B]$ zwischen der Standardbasis E und der Basis B, hat eine sehr einfache, nämlich *Diagonalgestalt*:

$$[\varphi_A]_{B,B} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} =: \operatorname{diag}(\lambda_1,\dots,\lambda_n)$$

(ii) Die *geometrische Wirkung* von φ_A ist leicht zu durchschauen: Was macht z.B. $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ mit den Punkten des $\mathbb{R}^2$ geometrisch?

In diesem Fall können wir die ONB $b_1 := \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $b_2 := \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ von $\mathbb{R}^2$ angeben, für die $A\cdot b_1 = 4b_1$ und $A\cdot b_2 = 2b_2$ gilt. Daher streckt A einfach in Richtung der 1. Mediane um den Faktor 4 und in Richtung der 2. Mediane um den Faktor 2, d.h. aus dem Einheitskreis K wird eine Ellipse.



(iii) *Potenzen von A* sind leicht zu berechnen: Ist $D := \operatorname{diag}(\lambda_1,\dots,\lambda_n) = [\varphi_A]_{B,B} = S^{-1}\cdot A\cdot S$ mit $S = [E\to B] = (b_1 \cdots b_n)$, so folgt $A = S\cdot D\cdot S^{-1}$ und weiter

$$A^k = \underbrace{(S\cdot D\cdot S^{-1})\cdots(S\cdot D\cdot S^{-1})}_{k \text{ Klammern}} = S\cdot D^k\cdot S^{-1} = S\cdot\begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix}\cdot S^{-1} = S\cdot\operatorname{diag}(\lambda_1^k,\dots,\lambda_n^k)\cdot S^{-1}.$$

Zum Beispiel $\begin{pmatrix}3&1\\1&3\end{pmatrix}^5 = \frac{1}{\sqrt{2}}\begin{pmatrix}1&-1\\1&1\end{pmatrix}\cdot\begin{pmatrix}4^5&0\\0&2^5\end{pmatrix}\cdot\underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\-1&1\end{pmatrix}}_{S^{-1}=S^t} = \frac{1}{2}\begin{pmatrix}4^5&-2^5\\4^5&2^5\end{pmatrix}\cdot\begin{pmatrix}1&-1\\1&1\end{pmatrix} = \begin{pmatrix}528&496\\496&528\end{pmatrix}$.

(iv) Sogar *wildere Funktionen von A* als Potenzen und Polynome können berechnet werden, wie zum Beispiel

$$e^A := \sum_{k=0}^{\infty}\frac{1}{k!}A^k = \sum_{k=0}^{\infty}\frac{1}{k!}S\cdot D^k\cdot S^{-1} = S\cdot\Big(\sum_{k=0}^{\infty}\frac{1}{k!}\begin{pmatrix}\lambda_1^k&0&\cdots&0\\0&\lambda_2^k&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&\lambda_n^k\end{pmatrix}\Big)\cdot S^{-1}$$
$$= S\cdot\begin{pmatrix}e^{\lambda_1}&0&\cdots&0\\0&e^{\lambda_2}&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&e^{\lambda_n}\end{pmatrix}\cdot S^{-1} = S\cdot\operatorname{diag}(e^{\lambda_1},\dots,e^{\lambda_n})\cdot S^{-1};$$

oder für die Potenzreihe der Sinus-Funktion ergibt sich $\sin A = S\cdot\operatorname{diag}(\sin\lambda_1,\dots,\sin\lambda_n)\cdot S^{-1}$; diese Techniken gehören zum sogenannten *Funktionalkalkül* und sind auch analog auf dem abstrakten Level für lineare Abbildungen $\varphi\colon V\to V$ verfügbar, sobald eine Basis $B=(b_1,\dots,b_n)$ von V mit den Eigenschaften $\varphi(b_i)=\lambda_i b_i$ $(i=1,\dots,n)$ existiert. (Es ergibt ja dann wieder $[\varphi]_{B,B}=\operatorname{diag}(\lambda_1,\dots,\lambda_n)$ usw.)

8.2. Definition: Sei $A\in M(n,n;\mathbb{K})$, $\lambda\in\mathbb{K}$ und $x\in\mathbb{K}^n$. Dann heißt x *Eigenvektor von A zum Eigenwert* λ, falls $x\neq 0$ und $A\cdot x=\lambda x$ gilt.

8.3. Definition: Sei V ein $\mathbb{K}$-VR, $\varphi\colon V\to V$ linear, $\lambda\in\mathbb{K}$ und $v\in V$. Dann heißt v *Eigenvektor von* φ *zum Eigenwert* λ, falls $v\neq 0$ und $\varphi(v)=\lambda v$ gilt.

8.4. Definition: (i) Für $A\in M(n,n;\mathbb{K})$ und $\lambda\in\mathbb{K}$ heißt

$$E(A,\lambda) := \{\text{ Eigenvektoren von } A \text{ zum Eigenwert } \lambda\ \}\cup\{0\}\subseteq\mathbb{K}^n$$

der *Eigenraum von A zum Eigenwert* λ.

(ii) Für V ein $\mathbb{K}$-VR, $\varphi\colon V\to V$ linear und $\lambda\in\mathbb{K}$ heißt

$$E(\varphi,\lambda) := \{\text{ Eigenvektoren von } \varphi \text{ zum Eigenwert } \lambda\ \}\cup\{0\}\subseteq V$$

der *Eigenraum von* φ *zum Eigenwert* λ.

Wegen $\big(\varphi(v)=\lambda v \Leftrightarrow (\varphi-\lambda\,\mathrm{id})(v)=0\big)$ ist $\boxed{E(\varphi,\lambda)=\ker(\varphi-\lambda\,\mathrm{id})}$.

Die Eigenräume sind Teilräume von $\mathbb{K}^n$ bzw. V.

Weiters gilt offensichtlich:

x ist Eigenvektor (EV) von A zum Eigenwert (EW) $\lambda \quad\Longleftrightarrow\quad x$ ist EV von φ_A zum EW λ,

v ist EV von φ zum EW $\lambda \quad\Longleftrightarrow\quad [v]_B$ ist EV von $[\varphi]_{B,B}$ zum EW λ für jede/eine Basis B.

In diesem Sinne sind also 8.2 und 8.3 gegeneinander austauschbar.

8.5. Bestimmung von Eigenwerten (in endlichdimensionalen VR):

$$
\begin{aligned}
A \text{ hat einen EV } x \neq 0 \text{ zum EW } \lambda \quad &\Longleftrightarrow \quad \exists x \neq 0\colon\ A \cdot x = \lambda x \\
&\Longleftrightarrow \quad \exists x \neq 0\colon\ A \cdot x - \lambda x = 0 \\
&\Longleftrightarrow \quad \exists x \neq 0\colon\ (A - \lambda I) \cdot x = 0 \\
&\Longleftrightarrow \quad (A - \lambda I) \cdot x = 0 \text{ hat nichttriviale Lösungen} \\
&\overset{5.6}{\Longleftrightarrow} \quad A - \lambda I \text{ ist nicht invertierbar} \\
&\overset{6.11\ (D9)}{\Longleftrightarrow} \quad \det(A - \lambda I) = 0,
\end{aligned}
$$

$$
\begin{aligned}
\varphi \text{ hat einen EV } v \neq 0 \text{ zum EW } \lambda \quad &\Longleftrightarrow \quad \exists v \neq 0\colon\ \varphi(v) = \lambda v \\
&\Longleftrightarrow \quad \exists v \neq 0\colon\ \varphi(v) - \lambda v = 0 \\
&\Longleftrightarrow \quad \exists v \neq 0\colon\ (\varphi - \lambda\,\mathrm{id})(v) = 0 \\
&\Longleftrightarrow \quad \varphi - \lambda\,\mathrm{id} \text{ ist nicht injektiv} \\
&\overset{4.25}{\Longleftrightarrow} \quad \varphi - \lambda\,\mathrm{id} \text{ ist nicht bijektiv} \\
&\overset{6.19}{\Longleftrightarrow} \quad \det(\varphi - \lambda\,\mathrm{id}) = 0 \\
&\Longleftrightarrow \quad \det([\varphi]_{B,B} - \lambda I) = 0 \text{ für jede Basis} B.
\end{aligned}
$$

Die Eigenwerte λ sind also Lösungen der Gleichung

$$0 = \det(\underbrace{A - tI}_{B}) = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) b_{1\sigma(1)} \cdots b_{n\sigma(n)} \quad \text{mit } \begin{cases} b_{ij} = a_{ij} & (i \neq j) \text{ Typ 1,} \\ b_{ii} = a_{ii} - t & \text{Typ 2,} \end{cases}$$

daher ist $\det(A - tI)$ ein Polynom $p(t)$ mit Koeffizienten aus $\mathbb{K}$. Die höchste Potenz von t erhalten wir aus dem Summanden, wo alle $b_{i\sigma(i)}$ vom Typ 2 sind, was nur für $\sigma = \mathrm{id}$ möglich ist. Der entsprechende Term lautet

$$1 \cdot (a_{11} - t) \cdots (a_{nn} - t) = (-t)^n + (-t)^{n-1} (\underbrace{a_{11} + \ldots + a_{nn}}_{\operatorname{trace} A}) + q(t).$$

Hier ist $q(t)$ ein Polynom vom Grad $n-2$. Die Potenz t^{n-1} können wir nur aus Summanden gewinnen, wo alle $b_{i\sigma(i)}$ bis auf höchstens eines vom Typ 2 sind, d.h. es gilt $\sigma(i) = i$ mindestens $(n-1)$ mal; damit muss σ aber schon die Identität sein und wir sind also wieder beim obigen Term gelandet. Daher wissen wir jetzt schon, dass

$$p(t) = (-t)^n + (-t)^{n-1} (a_{11} + \ldots + a_{nn}) + Q(t)$$

gilt, wobei $Q(t)$ ein Polynom vom Grad höchstens $n-2$ ist. Das konstante Glied des Polynoms ist $p(0) = \det(A - 0I) = \det A$. Somit haben wir insgesamt gezeigt, dass

$$\det(A - tI) = p(t) = (-1)^n t^n + (-1)^{n-1}(\operatorname{trace} A)\, t^{n-1} + \ldots + \det A$$

gilt. Die Eigenwerte von A sind genau die Nullstellen von $p(t)$ in $\mathbb{K}$.

8.6. Definition: (i) Für $A \in M(n,n;\mathbb{K})$ heißt $p_A(t) := \det(A - tI)$ das *charakteristische Polynom von A*. Das *Spektrum von A* ist die Menge aller Eigenwerte von A (gleichbedeutend: Menge aller Nullstellen von $p_A(t)$).

(ii) Für eine lineare Abbildung $\varphi\colon V \to V$ des endlichdimensionalen Vektorraumes V heißt $p_\varphi(t) := \det(\varphi - tI)$ das *charakteristische Polynom von φ*. Das *Spektrum von φ* ist die Menge aller Eigenwerte von φ (gleichbedeutend: Menge aller Nullstellen von $p_\varphi(t)$).

8.7. Satz: Die Eigenwerte von φ bzw. A sind die Nullstellen des charakteristischen Polynoms von φ bzw. A.

Beweis: Haben wir in 8.5 gesehen $\square$

8.8. Beispiel: Für $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ ist $A - \lambda I = \begin{pmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{pmatrix}$, daher

$$0 = \det(A - \lambda I) = (3-\lambda)^2 - 1 \quad\Longleftrightarrow\quad 3 - \lambda = \pm 1,\ \text{d.h. } \lambda = 3 \mp 1,$$

daher ist das Spektrum von A gleich $\{4, 2\}$.

Ein Eigenvektor $v_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \in \mathbb{R}^2$ zum Eigenwert $\lambda_1 = 4$ ist eine nichttriviale Lösung des Gleichungssystems $(A - 4I)v_1 = 0$, d.h.

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \text{ daher } v_1 = \begin{pmatrix} t \\ t \end{pmatrix} \quad (t \in \mathbb{R}, t \neq 0).$$

Einen Eigenvektor $v_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ zum Eigenwert $\lambda_2 = 2$ erhalten wir als nichttriviale Lösung von $(A - 2I)v_2 = 0$, d.h.

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \text{ daher } v_2 = \begin{pmatrix} -s \\ s \end{pmatrix} \quad (s \in \mathbb{R}, s \neq 0).$$

Wir bemerken, dass $v_1 \perp v_2$ gilt — das ist kein Zufall, siehe später; wir normieren die Vektoren damit wir eine ONB $B = (b_1, b_2)$ von $\mathbb{R}^2$ erhalten:

$b_1 := \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ist EV zum EW $\lambda_1 = 4$ und $b_2 := \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ist EV zum EW $\lambda_2 = 2$.

Mit der Basis B gilt nun $[\varphi_A]_{B,B} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$, was wir zur Probe auch mit Satz 4.39 und der orthogonalen Basiswechselmatrix $S = [E \to B] = (b_1\ b_2)$ nachrechnen können:

$$S^{-1} \cdot A \cdot S = S^t \cdot A \cdot S = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 4 & 4 \\ -2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 8 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}.$$

Kurzer Einschub über Nullstellen von Polynomen[20]

8.9. Proposition: Hat das Polynom $p(t)$ mit Koeffizienten aus $\mathbb{K}$ die Nullstelle $\lambda \in \mathbb{K}$, so ist $p(t)$ ohne Rest durch $(t - \lambda)$ teilbar.

8.10. Fundamentalsatz der Algebra: Jedes Polynom $p(t) = a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0$ mit Koeffizienten aus $\mathbb{C}$, $n \geq 1$ und $a_n \neq 0$ hat n komplexe Nullstellen, wenn man die Vielfachheiten mitzählt, d.h.: Es gibt ein $k \in \mathbb{N}$ mit $k \leq n$ und k paarweise verschiedene komplexe Zahlen $\mu_1, \ldots, \mu_k$ mit Vielfachheiten $1 \leq n_1, \ldots, n_k \leq n$, die $n_1 + \ldots + n_k = n$ erfüllen, sodass

$$p(t) = a_n (t - \mu_1)^{n_1} \cdots (t - \mu_k)^{n_k}$$

gilt. Es ist also μ_j jeweils eine n_j-fache Nullstelle von p.

8.11. Korollar: $a_0 = a_n (-\mu_1)^{n_1} \cdots (-\mu_k)^{n_k} = a_n \cdot (-1)^n \cdot$ Produkt aller Nullstellen.

Wegen $\mathbb{R} \subseteq \mathbb{C}$ gilt der Satz auch für Polynome mit REELLEN Koeffizienten, ABER niemand garantiert uns, dass die Nullstellen dann ebenfalls reell sind: z.B. hat das reelle Polynom $p(t) = t^2 + 1$ die komplexen Nullstellen $\pm i$.

8.12. Einige Fakten: (i) Bei reellen Polynomen treten nicht-reelle Nullstellen immer als Paare konjugiert komplexer Zahlen auf; daher haben wir für ein Polynom von ungeradem Grad immer mindestens eine reelle Nullstelle.

(ii) Ein Polynom der Form $p(t) = t^n + a_{n-1} t^{n-1} + \ldots + a_0$ mit Koeffizienten $a_j \in \mathbb{Z}$ hat in $\mathbb{R}$ nur ganzzahlige oder irrationale Nullstellen (also keine echt rationalen). Wegen 8.11 ist jede ganzzahlige Nullstelle ein Teiler von a_0 — was die Suche auf endlich viele Möglichkeiten beschränkt.

(iii) Ein Polynom $p(t) = a_n t^n + a_{n-1} t^{n-1} + \ldots + a_0$ mit $a_j \in \mathbb{Z}$, $a_n \neq 0$, geht durch Substitution $t = \frac{s}{a_n}$ und Multiplikation mit a_n^{n-1} über in $q(s) = s^n + a_{n-1} s^{n-1} + a_{n-2} a_n s^{n-2} \ldots + a_0 a_n^{n-1}$. Nach (ii) hat $q(s)$ in $\mathbb{R}$ nur ganzzahlige oder irrationale Nullstellen, daher hat $p(t)$ in $\mathbb{R}$ nur rationale Nullstellen der Gestalt $\frac{s}{a_n}$ mit $s \in \mathbb{Z}$, s teilt $a_0 a_n^{n-1}$, oder irrationale Nullstellen.

8.13. Definition:

Es sei V ein endlichdimensionaler VR über $\mathbb{K}$. Eine lineare Abbildung $\varphi\colon V \to V$ heißt *diagonalisierbar*, wenn es eine Basis $B = (b_1, \ldots, b_n)$ von V aus Eigenvektoren von φ gibt.

[20] Mehr Information und Beweise oder Hinweise zu Beweisen finden sich in den bekannten Lehrbüchern; vgl. z.B. in Fischers *Lineare Algebra* das Kapitel über Grundbegriffe.

In diesem Falle ist $[\varphi]_B = \operatorname{diag}(\lambda_1,\dots,\lambda_n) = \begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}$ (wir schreiben ab jetzt $[\varphi]_B$ statt $[\varphi]_{B,B}$), wobei λ_j EW zu b_j ist, und $p_\varphi(t) = \det([\varphi]_B - tI) = (\lambda_1 - t)\cdots(\lambda_n - t)$.

8.14. Definition: Eine Matrix $A \in M(n,n;\mathbb{K})$ heißt *diagonalisierbar über* $\mathbb{K}$, wenn es eine Basis $B = (b_1,\dots,b_n)$ von $\mathbb{K}^n$ aus Eigenvektoren von A gibt.

In diesem Fall gilt mit $S = [E \to B] = (b_1 \cdots b_n)$ die Gleichung $S^{-1}\cdot A\cdot S = [\varphi_A]_B = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$ und $p_A(t) = \det(A - tI) = \det([\varphi_A]_E - tI) \overset{[6.18]}{=} \det([\varphi_A]_B - tI) = (\lambda_1 - t)\cdots(\lambda_n - t)$.

8.15. Beispiel: Die Matrix der Drehung um den Winkel $\alpha \in [0,2\pi[$ im $\mathbb{R}^2$ ist gegeben durch $A = \begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix} \in M(2,2;\mathbb{R}) \subseteq M(2,2;\mathbb{C})$. Das charakteristische Polynom

$$p_A(\lambda) = \det(A - \lambda I) = \det\begin{pmatrix}\cos\alpha-\lambda & -\sin\alpha\\ \sin\alpha & \cos\alpha-\lambda\end{pmatrix} = (\cos\alpha-\lambda)^2 + \sin^2\alpha = \lambda^2 - 2\lambda\cos\alpha + 1$$

hat die Nullstellen $\lambda_{1,2} = \cos\alpha \pm \sqrt{\cos^2\alpha - 1} = \cos\alpha \pm i\sqrt{1-\cos^2\alpha} = \cos\alpha \pm i\sin\alpha$.

1. Fall $\alpha = 0$: Es ist $A = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ bereits diagonal und $\lambda_1 = \lambda_2 = 1$.

2. Fall $\alpha = \pi$: Es ist $A = \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$ bereits diagonal und $\lambda_1 = \lambda_2 = -1$.

3. Fall $0 < \alpha < 2\pi$ und $\alpha \neq \pi$: Wegen $\sin\alpha \neq 0$ sind λ_1 und λ_2 nicht reell und wir stellen fest:

- A ist NICHT *diagonalisierbar über* $\mathbb{R}$ (was eigentlich klar sein sollte, weil eine Drehung um α keinen Vektor $\neq 0$ auf ein Vielfaches von sich selbst abbildet!),
- A ist *diagonalisierbar über* $\mathbb{C}$ zu $\begin{pmatrix}\cos\alpha + i\sin\alpha & 0\\ 0 & \cos\alpha - i\sin\alpha\end{pmatrix} = \begin{pmatrix}e^{i\alpha} & 0\\ 0 & e^{-i\alpha}\end{pmatrix}$.

8.16. Untersuchung der Diagonalisierbarkeit mit Hilfe des charakteristischen Polynoms

Alle folgenden Überlegungen gelten gleichermaßen für lineare Abbildungen $\varphi\colon V \to V$ auf einem endlichdimensionalen Vektorraum V und für quadratische Matrizen A. Wir formulieren jeweils nur für den Fall von φ, die Matrix-Versionen der Sätze ergeben sich sofort, wenn man $\varphi = \varphi_A\colon x \mapsto A\cdot x$ setzt.

<u>1. Fall:</u> $p_\varphi(t)$ lässt sich über $\mathbb{K}$ NICHT in ein Produkt von Linearfaktoren $(\lambda_j - t)$ zerlegen.

Dann sind die Diagonalisierungsbemühungen aussichtslos: Hätten wir nämlich für eine Basis B die Gleichung $[\varphi]_B = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$, so wäre $p_\varphi(t) = (\lambda_1 - t)\cdots(\lambda_n - t)$ — ein Widerspruch!

Zur Illustration sei an den Fall $0 < \alpha < 2\pi$, $\alpha \neq \pi$ aus 8.15 erinnert, wo wir gesehen haben, dass die Drehmatrix für den Winkel α über $\mathbb{R}$ nicht diagonalisierbar ist.

<u>2. Fall:</u> $p_\varphi(t)$ zerfällt in $(\lambda_1 - t)\cdots(\lambda_n - t) = (\mu_1 - t)^{n_1}\cdots(\mu_k - t)^{n_k}$.
(Über $\mathbb{C}$ ist das nach dem Fundamentalsatz der Algebra immer der Fall.)

Dann gilt:

Sind alle λ_j verschieden (d.h. alle $n_j = 1$), so ist φ diagonalisierbar (siehe 8.18).

Sind die λ_j nicht notwendig verschieden, dann müssen die Eigenräume größtmöglich sein, damit φ diagonalisierbar ist (siehe 8.20).

8.17. Lemma: Sei $\varphi\colon V \to V$ linear. Weiters seien $\lambda_1, \ldots, \lambda_k$ paarweise verschiedene Eigenwerte von φ.

(i) Seien $v_j \in E(\varphi, \lambda_j)$ $(j = 1, \ldots, k)$, dann gilt:

$$v_1 + \ldots + v_k = 0 \quad \Longrightarrow \quad v_j = 0 \ \ (j = 1, \ldots, k)$$

(ii) Eigenvektoren zu paarweise verschiedenen Eigenwerten sind linear unabhängig.

Beweis: (i) durch Induktion nach k: Der Fall $k = 1$ ist klar. (Beachte: Der Nullvektor ist nie Eigenvektor, gehört aber natürlich zu jedem Eigenraum.)

Induktionsschluss $k-1 \mapsto k$: Aus $\sum\limits_{j=1}^{k} v_j = 0$ folgt einerseits durch Anwendung von φ auf beide Seiten

$$(\star) \qquad 0 = \varphi(0) = \varphi(\sum v_j) = \sum \varphi(v_j) = \sum \lambda_j v_j$$

und andererseits nach Multiplikation mit λ_k auch

$$(\circ) \qquad 0 = \lambda_k \sum v_j = \sum \lambda_k v_j.$$

Wir bilden die Differenz auf beiden Seiten von $(\star)$ und $(\circ)$ und erhalten

$$0 = \underbrace{(\lambda_1 - \lambda_k) v_1}_{\in E(\varphi,\lambda_1)} + \ldots + \underbrace{(\lambda_{k-1} - \lambda_k) v_{k-1}}_{\in E(\varphi,\lambda_{k-1})} + \underbrace{0 v_k}_{0}.$$

Nach Induktionsvoraussetzung gilt $(\lambda_j - \lambda_k) v_j = 0$ $(j = 1, \ldots, k-1)$ und somit wegen $\lambda_j - \lambda_k \neq 0$ auch $v_j = 0$ für $j = 1, \ldots, k-1$. Daher impliziert die oben angenommene Gleichung $v_1 + \ldots + v_k = 0$ nun auch $v_k = 0$. Somit ist die Aussage bewiesen.

(ii) Ist w_j EV zum EW λ_j und $\sum\limits_{j=1}^{k} \beta_j w_j = 0$ für Skalare β_j, dann gilt ja $\beta_j w_j \in E(\varphi, \lambda_j)$ und somit nach (i) jeweils $\beta_j w_j = 0$, was aber wegen $w_j \neq 0$ sofort $\beta_j = 0$ impliziert. $\square$

8.18. Satz: Sei V ein endlichdimensionaler VR mit $\dim V = n$ und $\varphi\colon V \to V$ linear. Das charakteristische Polynom von φ habe n paarweise verschiedene Nullstellen, d.h. es ist $p_\varphi(t) = (\lambda_1 - t)\cdots(\lambda_n - t)$ mit $\lambda_l \neq \lambda_j$ für $l \neq j$, dann ist φ diagonalisierbar.

Beweis: Sei b_j EV zu λ_j $(j = 1,\dots,n)$, dann ist $b_1,\dots,b_n$ gemäß 8.17(ii) linear unabhängig und somit wegen $\dim V = n$ eine Basis von V. □

8.19. Lemma: Sei V ein endlichdimensionaler VR, $\varphi\colon V \to V$ linear, $k \geq 1$ und $\lambda \in \mathbb{K}$ eine k-fache Nullstelle des charakteristischen Polynoms $p_\varphi(t)$, d.h. $p_\varphi(t) = (\lambda - t)^k q(t)$ für ein Polynom q mit $q(\lambda) \neq 0$. Dann gilt: $1 \leq \dim E(\varphi,\lambda) \leq k$.

Wir nennen $\dim E(\varphi,\lambda)$ die *geometrische Vielfachheit* und k die *algebraische Vielfachheit* des Eigenwertes λ.

Beweis: Sei $b_1,\dots,b_l$ eine Basis von $E(\varphi,\lambda)$; wir ergänzen diese zu einer Basis B von V und erhalten

$$[\varphi]_B = \begin{pmatrix} \begin{matrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{matrix} & \boxed{*} \\ \boxed{0} & \boxed{C} \end{pmatrix},$$

wobei λ hier l mal auftritt. Daher berechnen wir mittels Eigenschaft 6.11 (D8) für Determinanten von Blockmatrizen

$$p_\varphi(t) = \det([\varphi]_B - tI) = (\lambda - t)^l \det(C - tI)$$

und somit muss $l \leq k$ gelten. □

8.20. Satz: Sei V ein endlichdimensionaler VR mit $\dim V = n$ und $\varphi\colon V \to V$ linear mit charakteristischem Polynom

$$p_\varphi(t) = (\mu_1 - t)^{n_1}\cdots(\mu_k - t)^{n_k},$$

wobei $\mu_1,\dots,\mu_k \in \mathbb{K}$ paarweise verschiedenen sind und $n_1 + \ldots + n_k = n$ gilt. Dann sind die folgenden beiden Aussagen äquivalent:

(i) φ ist diagonalisierbar

(ii) $\dim E(\varphi,\mu_j) = n_j$ für alle $j = 1,\dots,k$ (d.h. alle geometr. und algebr. Vielfachh. stimmen überein).

Beweis: (i) $\Rightarrow$ (ii): Sei B eine Basis von V bestehend aus EV, von denen für $j = 1,\dots,k$ jeweils m_j viele EV im Eigenraum $E(\varphi,\mu_j)$ liegen. Dann gilt nach 8.19 immer $m_j \leq n_j$ und weiter

$$n = m_1 + \ldots + m_k \leq n_1 + \ldots + n_k = n,$$

was nur für $m_j = n_j$ möglich ist, weil wir hier durchwegs natürliche Zahlen vorliegen haben. Nach 8.17.(ii) ist $E(\varphi,\mu_r) \cap E(\varphi,\mu_s) = \{0\}$ für $r \neq s$ (kein EV kann in zwei Eigenräumen zu verschiedenen EW liegen), daher folgt $\dim E(\varphi,\mu_j) = m_j = n_j$ für $j = 1,\dots,k$.

(ii) $\Rightarrow$ (i): Für $j = 1, \ldots, k$ wählen wir jeweils eine Basis $B_j = (b_1^{(j)}, \ldots, b_{n_j}^{(j)})$ von $E(\varphi, \mu_j)$. Das System $B = (B_1, \ldots, B_k)$ besteht aus $n_1 + \ldots + n_k = n$ Vektoren. Außerdem ist B linear unabhängig: Ist nämlich

$$\underbrace{\sum_{l=1}^{n_1} \lambda_l^{(1)} b_l^{(1)}}_{\in E(\varphi,\mu_1)} + \ldots + \underbrace{\sum_{l=1}^{n_k} \lambda_l^{(k)} b_l^{(k)}}_{\in E(\varphi,\mu_k)} = 0,$$

dann sind nach 8.17(i) alle Einzelsummen 0 und somit wegen der linearen Unabhängigkeit jedes Teilsystems B_j schließlich alle $\lambda_l^{(j)} = 0$.

Daher ist B eine Basis von V, die aus EV besteht, also ist φ diagonalisierbar. $\square$

8.21. Proposition: Ist $\varphi\colon V \to V$ eine Projektion, d.h. $\varphi \circ \varphi = \varphi$, dann kommen nur 0 und 1 als Eigenwerte von φ in Frage.

Beweis: Für $v \neq 0$ mit $\varphi(v) = \lambda v$ gilt $\lambda v = \varphi(v) = \varphi \circ \varphi(v) = \varphi(\varphi(v)) = \varphi(\lambda v) = \lambda\lambda v = \lambda^2 v$, daher $\lambda^2 v = \lambda v$ und weiter $\lambda(\lambda - 1)v = 0$. Wegen $v \neq 0$ ist somit $\lambda(\lambda - 1) = 0$. $\square$

Aus 8.20 ist ersichtlich: Gibt es zu einem EW μ_j weniger als n_j linear unabhängige EV, dann ist φ nicht diagonalisierbar. Das folgende Beispiel zeigt den Extremfall $1 = \dim E(\varphi, \mu_j) < n_j = n$.

8.22. Beispiel: Wir betrachten die obere Dreiecksmatrix der Form

$$A = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & \lambda \end{pmatrix} \in M(n, n; \mathbb{K}).$$

Das charakteristische Polynom berechnen wir mittels 6.11 (D7) einfach zu

$$p_A(t) = \det \begin{pmatrix} \lambda - t & 1 & 0 & \ldots & 0 \\ 0 & \lambda - t & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & \lambda - t \end{pmatrix} = (\lambda - t)^n,$$

daher hat der EW λ die algebraische Vielfachheit n. Wir bestimmen den Eigenraum zu λ aus den Lösungen der Gleichung $(A - \lambda I)x = 0$:

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & 0 \end{pmatrix} \cdot x = 0 \quad \Longrightarrow \quad x_n = \ldots = x_2 = 0, x_1 = t \quad \Longrightarrow \quad E(A, \lambda) = \{\begin{pmatrix} t \\ 0 \\ \vdots \\ 0 \end{pmatrix} \mid t \in \mathbb{R}\}.$$

Somit ist die geometrische Vielfachheit von λ gleich $\dim E(A,\lambda) = 1$ und A ist also hochgradig nicht-diagonalisierbar!

Ein näherer Blick auf die Matrix $N := A - \lambda I$ in diesem Beispiel zeigt, dass N die Standardbasisvektoren wie folgt abbildet: $e_n \mapsto e_{n-1}$, $e_{n-1} \mapsto e_{n-2}, \dots, e_2 \mapsto e_1$, $e_1 \mapsto 0$. Daher bildet die n-malige Anwendung von N alle e_j auf 0 ab, d.h. $N^n = 0$ (die Einserkette in der Matrix wandert beim Potenzieren nach rechts oben und schrumpft).

8.23. Definition: Eine lineare Abbildung $\varphi\colon V \to V$ heißt *nilpotent*, wenn es ein $k \in \mathbb{N}$ gibt mit $\varphi^k := \underbrace{\varphi \circ \cdots \circ \varphi}_{k \text{ mal}} = 0$. (Analog für eine Matrix A mit $A^k = 0$.)

Für nicht-diagonalisierbare lineare Abbildungen kann man durch Wahl geeigneter Basen immerhin die unten beschriebene Jordansche Normalform erreichen, d.h. schlimmer als in Beispiel 8.22 kann es nicht werden.

8.24. Satz über die Jordansche Normalform: Sei V ein endlichdimensionaler VR, $\varphi\colon V \to V$ linear mit charakteristischem Polynom $p_\varphi(t)$, das vollständig in Linearfaktoren zerfällt (wie z.B. stets im Fall $\mathbb{K} = \mathbb{C}$):

$$p_\varphi(t) = (\mu_1 - t)^{n_1} \cdots (\mu_k - t)^{n_k}, \quad \text{alle } \mu_j \text{ paarweise verschieden}, n_1 + \ldots + n_k = n.$$

Dann existiert eine Basis B von V, sodass $[\varphi]_B$ in *Jordanscher Normalform* gegeben ist, d.h.

$$[\varphi]_B = \begin{pmatrix} \boxed{J_1} & & \\ & \ddots & \\ & & \boxed{J_k} \end{pmatrix},$$

wobei jedes „Jordan-Kästchen“ $J_l \in M(n_l, n_l; \mathbb{K})$ die folgende Gestalt hat:

$$J_l = \begin{pmatrix} \boxed{J_1^{(l)}} & & \\ & \ddots & \\ & & \boxed{J_{p_l}^{(l)}} \end{pmatrix}$$

und jedes $J_r^{(l)}$ ist entweder gleich μ_l oder hat die Form $\begin{pmatrix} \mu_l & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \dots & \dots & \mu_l \end{pmatrix}$.

Für eine quadratische Matrix A sagt der Satz, dass eine Basis B existiert, sodass mit dem Basiswechsel $S = [E \to B]$ die Matrix $S^{-1} \cdot A \cdot S = [\varphi_A]_B$ Jordansche Normalform wie oben annimmt.

Beweise dieses Satzes sind langwierig und finden sich z.B. in Fischers Büchern.

8.25. Beispiel: Die Matrix $J = \begin{pmatrix} \boxed{\begin{matrix}\boxed{\lambda_1} & \\ & \boxed{\begin{matrix}\lambda_1 & 1\\ & \lambda_1\end{matrix}}\end{matrix}} & & \\ & \boxed{\begin{matrix}\boxed{\begin{matrix}\lambda_2 & 1 & \\ & \lambda_2 & 1\\ & & \lambda_2\end{matrix}} & \\ & \boxed{\begin{matrix}\lambda_2 & 1\\ & \lambda_2\end{matrix}}\end{matrix}} & \\ & & \boxed{\lambda_3}\end{pmatrix}$

entspricht der Jordanschen Normalform einer Matrix mit charakteristischem Polynom gleich jenem von J, nämlich $p_J(t) = (\lambda_1 - t)^3(\lambda_2 - t)^5(\lambda_3 - t)$, insbesondere ist $n_1 = 3, n_2 = 5, n_3 = 1$; die EV finden sich genau in den Spalten ohne Einser, also ist $\dim E(J,\lambda_1) = 2 < n_1$, $\dim E(J,\lambda_2) = 2 < n_2$ und $\dim E(J,\lambda_3) = 1 = n_3$.

Eine Matrix ist offensichtlich genau dann diagonalisierbar, wenn ihre Jordansche Normalform nur 1×1-Kästchen, also keine Einserketten, enthält.

8.26. Definition: Zwei quadratische Matrizen $A, B \in M(n,n;\mathbb{K})$ heißen *ähnlich* zueinander, in Zeichen $A \sim B$, wenn es eine invertierbare Matrix $S \in GL(n,\mathbb{K})$ gibt, sodass $B = S^{-1} \cdot A \cdot S$.

Die Matrizen A und B sind also ähnlich, wenn sie dieselbe lineare Abbildung bzgl. verschiedener Basen beschreiben. Die Ähnlichkeit ist eine Äquivalenzrelation und erzeugt daher eine Klasseneinteilung in der Menge aller $n \times n$-Matrizen. In jeder dieser Klassen ist die Jordansche Normalform eine Vertreterin möglichst einfacher Gestalt. Ist A diagonalisierbar, dann ist die Jordansche Normalform sogar eine Diagonalmatrix.

Eigenwertprobleme in euklidischen oder unitären Vektorräumen

8.27. Satz:

$$\text{Die Eigenwerte von} \left\{\begin{matrix}\text{orthogonalen/unitären}\\ \text{selbstadjungierten}\\ \text{antiselbstadjungierten}\\ \text{nichtnegativen}\\ \text{positiven}\end{matrix}\right\} \left\{\begin{matrix}\text{Matrizen oder}\\ \text{linearen Abbildungen}\end{matrix}\right\} \text{ sind } \left\{\begin{matrix}\text{vom Betrag 1}\\ \text{reell}\\ \text{rein imaginär}\\ \text{nichtnegativ}\\ \text{positiv}\end{matrix}\right\}.$$

Beweis: Sei jeweils $\varphi(v) = \lambda v$ und $v \neq 0$.

φ unitär: $\langle v \mid v\rangle = \langle \varphi^+(\varphi(v)) \mid v\rangle = \langle \varphi(v) \mid \varphi(v)\rangle = \langle \lambda v \mid \lambda v\rangle = \overline{\lambda}\lambda\langle v \mid v\rangle = |\lambda|^2\langle v \mid v\rangle$, daher $|\lambda|^2 = 1$, weil $\langle v \mid v\rangle \neq 0$.

φ selbstadiungiert: $\lambda\langle v \mid v\rangle = \langle v \mid \lambda v\rangle = \langle v \mid \varphi(v)\rangle = \langle \varphi(v) \mid v\rangle = \langle \lambda v \mid v\rangle = \overline{\lambda}\langle v \mid v\rangle$, daher $\overline{\lambda} = \lambda$.

φ antiselbstadiungiert: $\lambda\langle v \mid v\rangle = \langle v \mid \lambda v\rangle = \langle v \mid \varphi(v)\rangle = -\langle \varphi(v) \mid v\rangle = -\langle \lambda v \mid v\rangle = -\overline{\lambda}\langle v \mid v\rangle$, daher $\overline{\lambda} = -\lambda$.

φ nichtnegativ: $\lambda\langle v \mid v\rangle = \langle v \mid \lambda v\rangle = \langle v \mid \varphi(v)\rangle \geq 0$, daher $\lambda \geq 0$. Ist φ sogar positiv, dann erhalten wir jeweils $>$ statt $\geq$. □

8.28. Bemerkung: (i) Warum kommen in 8.27 die „normalen“ Matrizen nicht vor? Weil in dem Fall jedes $\lambda \in \mathbb{C}$ als EW auftreten kann: z.B. ist $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ für alle $\lambda_1, \lambda_2 \in \mathbb{C}$ normal bzgl. des Standardskalarproduktes und hat die EW λ_1, λ_2.

(ii) Warum kommen in 8.27 keine „Orthogonalprojektionen“ vor? Weil diese schon durch 8.21 erledigt sind, wonach für jede Projektion nur die EW 0 und 1 in Frage kommen.

Orthogonale bzw. unitäre Diagonalisierbarkeit

8.29. Besonders nett ist es, wenn in einem euklidischen/unitären VR V die zum Diagonalisieren benützte Basis $B = (b_1, \ldots, b_n)$ als ONB gewählt werden kann — eine Matrix A bzw. lineare Abbildung φ heißt in diesem Fall *orthogonal/unitär diagonalisierbar*:

(i) $Ab_j = \lambda b_j$ beschreibt dann eine Stauchung/Streckung in Richtung von orthogonalen Achsen; vgl. Beispiel 8.8;

(ii) für die Basiswechselmatrix $S = [E \to B]$ ist $S^{-1} = S^t$ (oder $= S^+$) dann besonders leicht zu berechnen;

(iii) wegen $S^{-1}AS = S^tAS$ haben wir dann mit einem Schlag nicht nur lineare Abbildungen, sondern auch Bilinearformen durch Basiswechsel auf Diagonalform gebracht; das kann in der sogenannten „Hauptachsentransformation“ zum Entlarven von Kegelschnitten benützt werden, wie wir im Folgenden an einem einfachen Beispiel zeigen.

8.30. Beispiel: Welchen Kegelschnitt beschreibt die folgende Gleichung?

$$(\star) \qquad 3x_1^2 + 2x_1x_2 + 3x_2^2 = 8$$

Diese Gleichung können wir auch so schreiben:

$$(x_1\ x_2) \cdot \underbrace{\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}}_{A} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 8.$$

Die linke Seite ist also eine Bilinearform $\beta(x, y) = x^t \cdot A \cdot y$ mit symmetrischer Matrix A, in der $y = x$ gesetzt wurde. So etwas nennt man *quadratische Form.* [Allgemein lassen sich „verallgemeinerte Kegelschnitte“, sogenannte *Quadriken* in $\mathbb{R}^n$ mit Hilfe quadratischer Formen definieren.]

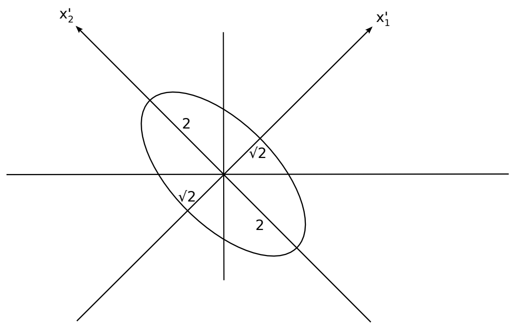
Wir übernehmen aus Beispiel 8.8: Ist B die ONB bestehend aus $b_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$, $b_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}$, daher mit Basiswechsel $S = (b_1\ b_2) = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}$, dann ist $S^{-1}AS = \begin{pmatrix}4 & 0\\0 & 2\end{pmatrix}$; da S orthogonal ist, aber auch $S^tAS = \begin{pmatrix}4 & 0\\0 & 2\end{pmatrix}$. Die Bilinearform β hat daher bzgl. B gemäß 7.50(iii) die Gestalt

$$\beta(x,y) = [x]_B^t S^t AS[y]_B = [x]_B^t \begin{pmatrix}4 & 0\\0 & 2\end{pmatrix}[y]_b.$$

Daher lautet die gegebene Gleichung ($\star$) in Koordinaten $[x]_B = \begin{pmatrix}x_1'\\x_2'\end{pmatrix}$ bzgl. B nun einfach

$$8 = \beta(x,x) = (x_1'\ x_2')\cdot\begin{pmatrix}4 & 0\\0 & 2\end{pmatrix}\cdot\begin{pmatrix}x_1'\\x_2'\end{pmatrix} = 4{x_1'}^2 + 2{x_2'}^2.$$

Es handelt sich also um eine Ellipse mit großer Halbachse 2 auf der 2. Mediane und kleiner Halbachse $\sqrt{2}$ auf der 1. Mediane.



Kurz zusammengefasst: Substituiere in der Gleichung $x^tAx = c$ einfach x durch $S[x]_B$, wobei B die Matrix A orthogonal diagonalisiert und $S = [E \to B]$ ist.

Im Fall von $\mathbb{K} = \mathbb{C}$ ist die Lage ausgezeichnet: Jede normale lineare Abbildung und jede normale Matrix lässt sich unitär diagonalisieren — und damit alle linearen Abbildungen und Matrizen aus den Listen 7.30 und 7.43!

8.31. Satz:[21] Sei V ein endlichdimensionaler unitärer VR und $\varphi\colon V \to V$ linear. Es gilt: φ ist genau dann normal, wenn φ unitär diagonalisierbar ist, d.h. wenn es eine ONB B von V gibt, sodass $[\varphi]_B = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

8.32. Korollar: Jede reelle symmetrische $n \times n$-Matrix lässt sich orthogonal diagonalisieren und ihre Eigenwerte sind reell. Darüberhinaus sind EV zu verschiedenen EW orthogonal.

[21] Für einen Beweis siehe z.B. Kapitel V, §4 in Bröckers Buch.

Beweis: Sei $A \in M(n,n;\mathbb{R})$ symmetrisch, dann ist $\varphi_A\colon x \mapsto Ax$ selbstadjungiert bzgl. des Standardskalarproduktes, also normal. Daher gibt es ein $T \in U(n)$, sodass $T^+AT = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$ gilt. Nach 8.27 sind die EW λ_j von A reell, also sind die homogenen linearen Gleichungssysteme $(A-\lambda_j I)x = 0$, mit denen die EV bestimmt werden, ebenfalls reell. Das Gauß-Verfahren liefert daher reelle Eigenvektoren $b_1',\dots,b_n' \in \mathbb{R}^n \subseteq \mathbb{C}^n$ zu den EW $\lambda_1,\dots,\lambda_n \in \mathbb{R}$.

In jedem der Eigenräume $E(A,\lambda_j)$ können wir das Orthogonalisierungsverfahren von Gram-Schmidt anwenden und erhalten eine neue Basis $b_1,\dots,b_n$ von V. Ist nun $b_l \in E(A,\lambda_l)$ und $b_j \in E(A,\lambda_j)$ mit $\lambda_l \neq \lambda_j$, so folgt

$$\lambda_j\langle b_l \mid b_j\rangle = \langle b_l \mid \lambda_j b_j\rangle = \langle b_l \mid Ab_j\rangle = \langle Ab_l \mid b_j\rangle = \langle \lambda_l b_l \mid b_j\rangle = \lambda_l\langle b_l \mid b_j\rangle,$$

daher muss $\langle b_l \mid b_j\rangle = 0$ gelten, d.h. $b_l \perp b_j$.

Somit bilden die $b_1,\dots,b_n$ auch insgesamt ein ONS (also nicht nur in jedem Eigenraum separat) und ergeben somit eine ONB aus EV. □

8.33. Beispiel: Wie sehen die Matrizen aus $SO(3)$ bzgl. ihrer geometrischen Wirkung aus? (Erinnere: $SO(2)$ besteht aus den Drehungen in der Ebene.)

Sei $A \in SO(3)$, d.h. A ist eine 3×3-Matrix mit $A^tA = I$ und $\det A = 1$.

Das charakteristsiche Polynom von A ist vom Grad 3 und zerfällt über $\mathbb{C}$:

$$p_A(t) = a_0 + a_1 t + a_2 t^2 - t^3 = (\lambda_1 - t)(\lambda_2 - t)(\lambda_3 - t).$$

Nach 8.5 ist $a_0 = \det A = 1$, daher gilt $\lambda_1\lambda_2\lambda_3 = 1$, und gemäß 8.27 ist $|\lambda_1| = |\lambda_2| = |\lambda_3| = 1$.

Nach 8.12(i) gibt es zwei Möglichkeiten:

1. Fall: Alle drei $\lambda_1,\lambda_2,\lambda_3$ sind reell, also gleich ± 1.

Dann kommen für die EW nur die Werte $1,1,1$ oder $1,-1,-1$ in Frage.

2. Fall: Zwei der EW sind nicht-reell und konjugiert komplex zueinander.

OBdA sind λ_2 und λ_3 diese; dann folgt $\lambda_2\lambda_3 = \lambda_2\overline{\lambda_2} = |\lambda_2|^2 = 1$ und somit muss $\lambda_1 = 1$ gelten.

In jedem Fall haben wir einen EW 1 und wir können eine ONB $B = (b_1,b_2,b_3)$ von $\mathbb{R}^3$ wählen mit $Ab_3 = b_3$. Dann gilt

$$[\varphi_A]_B = \begin{pmatrix} \boxed{\begin{matrix} * & * \\ * & * \end{matrix}} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} c_{31} & c_{32} \end{matrix} & 1 \end{pmatrix}$$

Nach 7.45 ist $[\varphi_A]_B$ wiederum orthogonal, also bilden die Zeilen eine ONB. Daher muss $c_{31} = c_{32} = 0$ sein. Die linke obere 2×2-Matrix ist daher orthogonal und hat Determinante 1, ist also eine ebene Drehung um einen Winkel α.

Somit erhalten wir insgesamt: A bewirkt eine Drehung um die Achse b_3 und sieht nach geeignetem Basiswechsel so aus

$$[\varphi_A]_B = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

8.34. Quadriken im $\mathbb{R}^n$: Diese sind definiert als Lösungsmenge einer Gleichung für $x \in \mathbb{R}^n$ der Gestalt

$$x^t Ax + 2a^t x + \alpha = 0,$$

wobei $A \in M(n,n;\mathbb{R})$ symmetrisch ist, $a \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$. Wir haben also neben Termen zweiter Ordnung (das sind rein quadratische und gemischte) im Allgemeinen auch noch lineare und konstante Glieder. Diese Gleichung kann mit einer $(n+1) \times (n+1)$-Matrix umgeschrieben werden in die Form

$$(1\ x^t) \cdot \underbrace{\begin{pmatrix} \alpha & a^t \\ a & A \end{pmatrix}}_{\text{symmetrisch}} \cdot \begin{pmatrix} 1 \\ x \end{pmatrix} = 0.$$

Zunächst beseitigt man mittels orthogonaler Transformation die „gemischten“ Glieder (durch Diagonalisierung der symmetrischen Matrix) und bringt dann in weiteren Schritten, die auch Translationen beinhalten, die linearen Glieder bis auf höchstens eines zum Verschwinden. Man erhält dann eine der in diversen Listen aufgeführten „Normalformen“, der man direkt die Art der Quadrik ansieht. Ihre ursprüngliche Lage erkennt man durch Zurückverfolgen der ausgeführten Transformationen. Beschreibungen und Illustrationen der typischen Quadriken im $\mathbb{R}^3$ finden sich zum Beispiel in Fischers Büchern.

8.35. Spektralsätze: Sätze, die Aussagen über die Diagonalisierbarkeit linearer Abbildungen $\varphi\colon V \to V$ machen, werden oft *Spektralsätze* genannt. Sie werden meist so formuliert, dass V, id_V sowie φ in einzelne „Stücke“ zerlegt werden, die den verschiedenen EW μ_j, also dem *„Spektrum“*, entsprechen:

$$\begin{aligned} V &= \bigoplus_{\mu \in \text{Spektrum}(\varphi)} E(\varphi,\mu), \\ \mathrm{id}_V &= \sum_\mu \pi_\mu \qquad (\pi_\mu \dots \text{ Projektion auf } E(\varphi,\mu)), \\ \varphi &= \sum_\mu \mu\pi_\mu. \end{aligned}$$

Sei speziell V endlichdimensional, φ diagonalisierbar und $\mu_1,\dots,\mu_k$ die paarweise verschiedenen EW mit Vielfachheiten $n_1,\dots,n_k$, $n_1+\dots+n_k = n = \dim V$. Die ausführlichen Bezeichnungen für eine Basis $B = (b_1,\dots,b_n)$ aus EV seien wie folgt: $b_l^{(1)} := b_l$ für $l = 1,\dots,n_1$ seien die EV zum EW μ_1, $b_l^{(2)} := b_{l+n_1}$ für $l = 1,\dots,n_2$ seien die EV zum EW μ_2 usw. …

Dann gilt $E(\varphi,\mu_j) = [b_1^{(j)},\ldots,b_{n_j}^{(j)}]$ und jedes $v \in V$ lässt sich schreiben in der Form

$$v = \sum_j \sum_l \alpha_l^{(j)} b_l^{(j)} = \underbrace{\sum_{l=1}^{n_1} \alpha_l^{(1)} b_l^{(1)}}_{\pi_{\mu_1}(v)\in E(\varphi,\mu_1)} + \ldots + \underbrace{\sum_{l=1}^{n_k} \alpha_l^{(k)} b_l^{(k)}}_{\pi_{\mu_k}(v)\in E(\varphi,\mu_k)}$$

sowie $\varphi(v) = \varphi(\sum_j \pi_{\mu_j}(v)) = \sum_j \varphi(\pi_{\mu_j}(v)) = \sum_j \mu_j \pi_{\mu_j}(v)$. Somit erhalten wir:

$$\begin{aligned} V &= E(\varphi,\mu_1) \oplus \ldots \oplus E(\varphi,\mu_k), \\ \mathrm{id}_V &= \pi_{\mu_1} + \ldots + \pi_{\mu_k}, \\ \varphi &= \mu_1 \pi_{\mu_1} + \ldots + \mu_k \pi_{\mu_k}. \end{aligned}$$

Falls V ein euklidischer/unitärer VR ist und B eine ONB aus EV, dann gilt $v = \sum_l \langle b_l \mid v\rangle b_l$, $\pi_{\mu_j}(v) = \sum_{l:\ \lambda_l=\mu_j} \langle b_l \mid v\rangle b_l$ und $\varphi(v) = \sum_l \lambda_l \langle b_l \mid v\rangle b_l$. Somit in bra-ket-Notation:

$$\begin{aligned} \mathrm{id}_V &= \sum_l \mid b_l\rangle\langle b_l \mid, \\ \pi_{\mu_j} &= \sum_{l:\ \lambda_l=\mu_j} \mid b_l\rangle\langle b_l \mid, \\ \varphi &= \sum_l \lambda_l \mid b_l\rangle\langle b_l \mid. \end{aligned}$$

ANHANG

A0. Körper

Die unten angeführten Rechenregeln (K1)-(K10) gelten in den aus der Schule bekannten Zahlenmengen $\mathbb{R}$ (reelle Zahlen), $\mathbb{Q}$ (rationale Zahlen, Bruchzahlen) und $\mathbb{C}$ (komplexe Zahlen). Für den algebraischen Begriff eines (kommutativen) Körpers stellt man diese Regeln als Axiome voran.

Definition: Eine Menge $\mathbb{K}$ zusammen mit zwei Operationen

$$+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}, \qquad (\lambda, \mu) \mapsto \lambda + \mu,$$
$$\cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}, \qquad (\lambda, \mu) \mapsto \lambda \cdot \mu,$$

heißt *Körper*, wenn die folgenden Rechenregeln für alle Elemente $\lambda, \mu, \nu \in \mathbb{K}$ gelten:

(K1) $(\lambda + \mu) + \nu = \lambda + (\mu + \nu)$ (Assoziativität der Addition).

(K2) $\lambda + \mu = \mu + \lambda$ (Kommutativität der Addition).

(K3) Es gibt ein Element $0 \in \mathbb{K}$ mit $0 + \lambda = \lambda + 0$.

(K4) Zu jedem λ existiert ein Element $-\lambda$ in $\mathbb{K}$ mit $\lambda + (-\lambda) = 0$.

(K5) $(\lambda \cdot \mu) \cdot \nu = \lambda \cdot (\mu \cdot \nu)$ (Assoziativität der Multiplikation).

(K6) Es gibt ein Element $1 \in \mathbb{K}$ mit $1 \cdot \lambda = \lambda$.

(K7) Zu jedem λ mit $\lambda \neq 0$ existiert ein Element λ^{-1} mit $\lambda \cdot \lambda^{-1} = 1$.

(K8) $\lambda \cdot (\mu + \nu) = \lambda \cdot \mu + \lambda \cdot \nu$ (Distributivität).

(K9) $1 \neq 0$.

(K10) $\lambda \cdot \mu = \mu \cdot \lambda$ (Kommutativität der Multiplikation).

Bemerkung: (K1)-(K4) besagen gerade, dass $\mathbb{K}$ bzgl. der Addition eine kommutative Gruppe ist; (K5)-(K7) und (K9) besagen gerade, dass die Menge $\mathbb{K}$ ohne das Element 0 bzgl. der Multiplikation eine kommutative Gruppe bildet.

A1. Das Vektorprodukt im $\mathbb{R}^3$

Definition: $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} := \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$

Geometrisch: Sind $a, b \in \mathbb{R}^3$ linear unabhängig (andernfalls ist $a \times b = 0$, vgl. (iv)), dann steht $a \times b$ normal sowohl auf a als auch auf b (vgl. (i)), die Länge von $a \times b$ ist gleich dem Flächeninhalt des von a und b aufgespannten Parallelogramms (vgl. (vii)) und $a, b, a \times b$ bildet ein Rechtssystem (vgl. (viii)).

Einige Eigenschaften:

(i) $\langle a \times b \mid a \rangle = \langle a \times b \mid b \rangle = 0$

(ii) $a \times b = -b \times a$

(iii) $a \times b$ ist linear in a und in b, d.h.

$$(a + c) \times b = a \times b + c \times b \qquad a \times (b + c) = a \times b + a \times c$$
$$(\lambda a) \times b = \lambda (a \times b) \qquad a \times (\lambda b) = \lambda (a \times b)$$

(iv) $a \times b = 0 \quad \Longleftrightarrow \quad a, b,$ linear abhängig

(v) $a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = 0$ (Jacobi-Identität)

(vi) $a \times (b \times c) = \langle a \mid c \rangle b - \langle a \mid b \rangle c$ und $(a \times b) \times c = \langle a \mid c \rangle b - \langle b \mid c \rangle a$

(vii) $\| a \times b \| = \|a\| \, \|b\| \sin \measuredangle(a, b)$ und $\langle a \mid b \rangle = \|a\| \, \|b\| \cos \measuredangle(a, b)$,
woraus nach Quadrieren und Addieren folgt: $\| a \times b \|^2 + |\langle a \mid b \rangle|^2 = \|a\|^2 \|b\|^2$

(viii) $\langle a \times b \mid c \rangle = \det(a, b, c)$ *Spatprodukt* von a, b, c (Spat = Parallelepiped im $\mathbb{R}^3$);
Absolutbetrag davon = Volumen des von a, b, c aufgespannten Spates.
Aus den Rechenregeln für die Determinante (Spaltenvertauschung!) folgt:

$$\langle a \times b \mid c \rangle = \langle b \times c \mid a \rangle = \langle c \times a \mid b \rangle.$$

[Beweise durch Einsetzen der jeweiligen Definitionen und Nachrechnen.]

A2. Gleichungen von Geraden und Ebenen

	Vektorformen (p Punkt; v, w Richtungsvektoren; n Normalvektor)		Koordinatenform
Gerade im $\mathbb{R}^2$	$x = p + \lambda v$	$\langle x - p \mid n\rangle = 0$	$ax + by = e$; $y = kx + d$
Gerade im $\mathbb{R}^3$	$x = p + \lambda v$	$(x - p) \times v = 0$	
Ebene im $\mathbb{R}^3$	$x = p + \lambda v + \mu w$	$\langle x - p \mid n\rangle = 0$	$ax + by + cz = e$

Normalabstand des Punktes q von der Geraden $\langle x - p \mid n\rangle = 0$ im $\mathbb{R}^2$: $\quad |\langle q - p \mid \frac{n}{\|n\|}\rangle|$

des Punktes q von der Ebene $\langle x - p \mid n\rangle = 0$ im $\mathbb{R}^3$: $\quad |\langle q - p \mid \frac{n}{\|n\|}\rangle|$

des Punktes q von der Geraden $x = p + \lambda v$ im $\mathbb{R}^3$: $\quad \|(q - p) \times \frac{v}{\|v\|}\|$

der Geraden $y = q + \mu v$ von

der parallelen Geraden $x = p + \lambda v$ im $\mathbb{R}^3$: $\quad \|(q - p) \times \frac{v}{\|v\|}\|$

der Geraden $y = q + \mu w$ von

der dazu windschiefen Geraden $x = p + \lambda v$ im $\mathbb{R}^3$: $\quad |\langle q - p \mid \frac{v \times w}{\|v \times w\|}\rangle|$.

Definition: Eine *Hyperebene H* in einem $\mathbb{K}$-Vektorraum V ist ein affiner Teilraum „mit einer uns 1 kleineren Dimension", d.h. ein affiner Teilraum $H = p + L$ (L Teilraum), sodass für $x \in V \setminus L$ stets $[L, x] = V$ gilt („eine Dimension dazu ergibt schon ganz V").

(Es gilt übrigens: $\forall x \in V \setminus L\colon [L, x] = V \iff \exists x \in V \setminus L\colon [L, x] = V$.)

Es ist in diesem Fall immer möglich, eine Linearform $f\colon V \to \mathbb{K}$ zu konstruieren mit $L = \ker f$, d.h. $L = \{x \in V \mid f(x) = 0\}$ [vgl. „Normalvektor"!]]. Setzt man $e := f(p)$, dann ist $H = \{x \in V \mid f(x) = e\}$, d.h. *$H$ wird durch die Gleichung $f(x) = e$ beschrieben* ($f \in V^*$, $e \in \mathbb{K}$).

Daraus folgt: Eine Hyperebene in einem $\mathbb{K}$-Vektorraum der endlichen Dimension n ist ein $(n-1)$-dimensionaler affiner Teilraum; dieser lässt sich durch eine Gleichung $f(x) = e$ ($f \in V^*$, $e \in \mathbb{K}$) beschreiben. Spezialfall $V = \mathbb{K}^n$: Gleichung der Hyperebene: $a_1 x_1 + \ldots + a_n x_n = e$, wobei $(a_1, \ldots, a_n) \in (\mathbb{K}^n)^*$ und $e \in \mathbb{K}$.

Die Hyperebenen im $\mathbb{R}^2$ sind die Geraden, im $\mathbb{R}^3$ die Ebenen.

A3. Algebren

Auf der Menge $M(n,n;\mathbb{K})$ der $n \times n$-Matrizen hatten wir die Operationen der Addition, der Multiplikation mit einem Skalar und der Matrizenmultiplikation. Die ersten beiden Operationen verleihen $M(n,n;\mathbb{K})$ eine Vektorraumstruktur, während die erste und dritte Operation dieselbe Menge auch zum Ring machen. Eine abstrakte Version so einer „Verschmelzung von Vektorraum- und Ringstruktur" finden wir im folgenden Begriff.

Definition: Eine *Algebra* über einem Körper $\mathbb{K}$ ist eine Menge A $(\neq \emptyset)$ mit drei Verknüpfungen $+: A\times A \to A$, $\cdot: \mathbb{K}\times A \to A$ und $\bullet: A\times A \to A$, sodass folgende Eigenschaften gelten: $(A,+,\cdot)$ ist ein $\mathbb{K}$-Vektorraum, d.h. + und $\cdot$ erfüllen die Axiome $(V1)-(V8)$, und zusätzlich haben wir für alle $a,b,c \in A$ und $\lambda \in \mathbb{K}$

(A9) $a \bullet (b \bullet c) = (a \bullet b) \bullet c$ (d.h. $\bullet$ ist assoziativ)

(A10) $a \bullet (b + c) = a \bullet b + a \bullet c$ und $(a + b) \bullet c = a \bullet c + b \bullet c$

(A11) $a \bullet (\lambda b) = \lambda(a \bullet b)$ und $(\lambda a) \bullet b = \lambda(a \bullet b)$.

Bemerkung: Die Axiome (A10)-(A11) besagen gerade, dass $\bullet$ eine $\mathbb{K}$-bilineare Abbildung ist. Die Axiome (V1)-(V4) zusammen mit (A9)-(A10) ergeben, dass $(A,+,\bullet)$ ein Ring ist.

Eine Algebra heißt *kommutativ*, wenn für alle $a,b \in A$ die Relation $a \bullet b = b \bullet a$ gilt; sie heißt *Algebra mit Eins(element)*, wenn ein $e \in A$ existiert mit $e \bullet a = a \bullet e = a$ für alle $a \in A$.

Beispiele:

1. $C[a,b]$ mit den üblichen Rechenoperationen für Funktionen ist eine kommutative Algebra mit Einselement.
2. $C_0(\mathbb{R}) = \{f: \mathbb{R} \to \mathbb{R} \mid f \text{ stetig}, \lim\limits_{x\to\pm\infty} f(x) = 0\}$ ist eine kommutative Algebra *ohne* Einselement, denn dies könnte nur die konstante Funktion mit Wert 1 sein, und die gehört aber nicht zu $C_0(\mathbb{R})$.
3. $M(n,n;\mathbb{K})$ $(n \geq 2)$ ist eine nichtkommutative Algebra mit Eins (siehe Kaiptel 2).