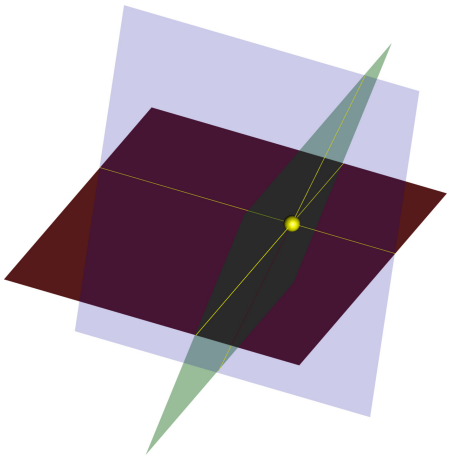
Thomas Krämer

Lineare Algebra

HU Berlin, Winter 24/25



(Version vom 11. Februar 2025)

Inhaltsverzeichnis

Einführung

Eine Quelle der Mathematik ist das Studium von Gleichungen und Funktionen. Die lineare Algebra behandelt den einfachsten Fall: Systeme von

- linearen Gleichungen wie $a_1x_1 + \cdots + a_nx_n = b$,
- linearen Abbildungen wie $(x_1, \ldots, x_n) \mapsto a_1x_1 + \cdots + a_nx_n$.

Diese spielen eine zentrale Rolle in der gesamten Mathematik und darüber hinaus, von Computergraphik, Kryptographie und Datenkompression bis hin zu künstlicher Intelligenz. Lineare Gleichungssysteme können unendlich viele Lösungen haben, aber ihre Lösungsmenge hat eine sehr einfache geometrische Struktur und lässt sich durch endlich viele Daten parametrisieren. Dies wird uns auf das abstrakte Konzept eines Vektorraumes als Ausgangspunkt der linearen Algebra führen. Doch bevor wir damit beginnen, wollen wir zunächst kurz an einige aus der Schule bekannte Grundlagen zu linearen Gleichungssystemen erinnern.

Lineare Gleichungssysteme

Viele Probleme führen in natürlicher Weise auf Systeme von linearen Gleichungen in mehreren Variablen:

Beispiel (ein Mischungsproblem). *Gegeben seien zwei Salzlösungen mit einer Konzentration von* 3 *bzw.* 6 *Gewichtsprozent. Wie müssen wir die beiden Lösungen mischen, um eine Gewichtseinheit einer* 5%*-igen Salzlösung zu erstellen?*

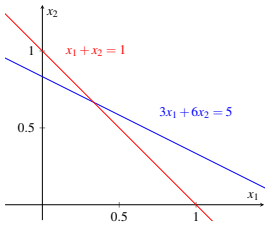
Angenommen, wir mischen x_1 Gewichtseinheiten der ersten Salzlösung mit x_2 Gewichtseinheiten der zweiten Lösung. Da wir insgesamt eine Gewichtseinheit der gewünschten Lösung wollen und diese eine Konzentration von 5 Prozent haben soll, suchen wir eine Lösung des Gleichungssystems

$$\begin{aligned} x_1 + x_2 &= 1, \\ 3x_1 + 6x_2 &= 5. \end{aligned}$$

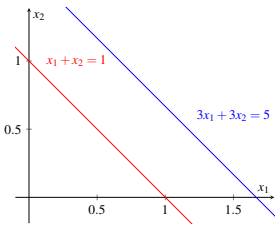
Indem wir von der zweiten Gleichung das dreifache der ersten subtrahieren, erhalten wir das hierzu äquivalente System

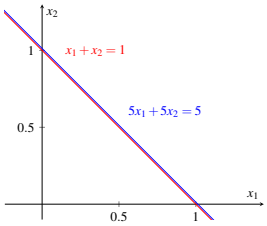
$$\begin{aligned} x_1 + x_2 &= 1, \\ 3x_2 &= 2. \end{aligned}$$

Dieses System hat obere Dreiecksform und besitzt daher genau eine Lösung, die wir ablesen als $(x_1, x_2) = (1/3, 2/3)$. Wir können die Lösung auch wie in der folgenden Skizze visualisieren, wobei die rote Gerade die gewünschte Gesamtmenge und die blaue Gerade die gewünschte Konzentration beschreibt. Die gesuchte Lösung ist somit geometrisch eindeutig bestimmt als Schnittpunkt der beiden Geraden:

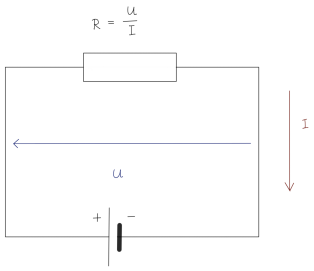


Mischen von Salzlösungen der gleichen Konzentration ändert die Konzentration nicht. Unser lineares Gleichungssystem besitzt dann entweder gar keine Lösung oder unendlich viele:

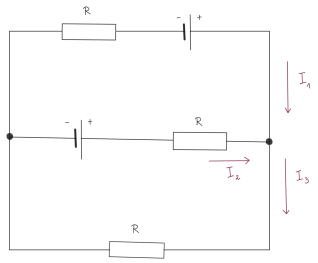




Lineare Gleichungssysteme treten auch bei der Untersuchung von elektrischen Netzwerken auf. Dazu erinnern wir uns zunächst an das *Ohmsche Gesetz*: Die durch einen elektrischen Leiter fließende Stromstärke I ist proportional zur anliegenden Spannung U. Die Konstante $R = U/I$ heißt der *Widerstand* des Leiters.



Beispiel (Ein elektrisches Netzwerk). *Zwei gleiche Batterien mit der Spannung U und drei gleiche Widerstände mit dem Widerstandswert R seien wie folgt zu einer Schaltung verlötet:*



Wie groß sind die in der Schaltung fließenden Stromstärken I_1, I_2, I_3?

Wir benötigen dazu etwas mehr Physik: Die *Knotenpunktregel* besagt, dass die Summe der in jeden Knoten einfließenden Ströme übereinstimmt mit der Summe der ausfließenden Ströme, im obigen Beispiel also $I_1 + I_2 = I_3$. Die *Maschenregel* besagt, dass entlang jeder Masche die Summe der abfallenden Spannungen Null ist, im obigen Beispiel also $RI_1 - RI_2 = RI_2 + RI_3 - U = 0$. Wir erhalten das lineare Gleichungssystem:

$$\begin{aligned} I_1 + I_2 - I_3 &= 0, \\ I_1 - I_2 \quad &= 0, \\ I_2 + I_3 &= U/R. \end{aligned}$$

Ziehen wir von der zweiten Gleichung die erste ab, so erhalten wir das äquivalente System

$$\begin{aligned} I_1 + I_2 - I_3 &= 0, \\ -2I_2 + I_3 &= 0, \\ I_2 + I_3 &= U/R. \end{aligned}$$

Addieren wir nun zur letzten Gleichung die Hälfte der zweiten, so erhalten wir das äquivalente System

$$\begin{aligned} I_1 + I_2 - I_3 &= 0, \\ -2I_2 + I_3 &= 0, \\ \tfrac{3}{2} I_3 &= U/R. \end{aligned}$$

Dies hat obere Dreiecksform und hat somit eine eindeutige Lösung, die wir hier ablesen als

$$(I_1, I_2, I_3) = (I, I, 2I) \quad \text{mit} \quad I = U/3R.$$

Der Gauß-Algorithmus

Das obige Verfahren der sukzessiven Umformung auf Dreiecksform kennen viele aus der Schule. Wir wollen es hier allgemein formulieren:

Definition. Ein *lineares Gleichungssystem* ist ein Gleichungssystem der Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_1 \end{aligned} \qquad (*)$$

mit Koeffizienten $a_{ij}, b_i \in \mathbb{R}$ und Variablen $x_1, \ldots, x_n$. Das System heißt

- *homogen*, falls $b_i = 0$ für alle i ist,
- *inhomogen*, falls ein i existiert mit $b_i \neq 0$.

Im homogenen Fall ist alle Information über das lineare Gleichungssystem gegeben durch die *Koeffizientenmatrix*, die alle Koeffizienten anordnet in einer in Klammen stehenden Tabelle

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

Im inhomogenen Fall nutzt man die *erweiterte Koeffizientenmatrix*

$$(A|b) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right) \quad \text{mit} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Die *Lösungsmenge* des linearen Gleichungssystems bezeichnen wir im Folgenden kurz mit

$$\mathscr{L}(A \mid b) = \{(c_1, \ldots, c_n) \in \mathbb{R}^n \mid (*) \text{ gilt, wenn man } x_i = c_i \text{ für alle } i \text{ setzt}\}.$$

Die Idee des Gauß-Algorithmus ist es, die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems zu vereinfachen, ohne dabei seine Lösungsmenge zu verändern. Dazu benutzen wir folgende Umformungen:

Definition. Eine *elementaren Zeilenumformung* einer Matrix ist eine der folgenden Operationen:

a) Vertauschung von zwei Zeilen.

b) Multiplikation jedes Eintrags einer gegebenen Zeile mit einer festen Zahl $\lambda \neq 0$.

c) Addition des λ-fachen einer Zeile zu einer echt anderen Zeile. Dabei sind das Vielfache und die Addition von Zeilen Eintrag für Eintrag zu verstehen.

Wir geben solche Zeilenumformungen wie im folgenden Beispiel durch Pfeile am rechten Rand der Matrix an:

$$\begin{pmatrix} 2 & 4 & 5 \\ 1 & 2 & 2 \\ 3 & 6 & 12 \end{pmatrix} \begin{matrix} \leftarrow \\ \leftarrow \\ \mid :3 \end{matrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 5 \\ 1 & 2 & 4 \end{pmatrix} \begin{matrix} \cdot(-2) \\ + \\ \end{matrix} \begin{matrix} \cdot(-1) \\ \\ + \end{matrix} \rightsquigarrow \begin{pmatrix} 1 & 3 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{matrix} + \\ \cdot(-2) \\ \\ \end{matrix} \begin{matrix} \\ \\ \cdot(-2) \\ + \end{matrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Hier haben wir zunächst die ersten beiden Zeilen vertauscht und die dritte mit $1/3$ multipliziert. Dann haben wir von der zweiten Zeile das doppelte der ersten Zeile abgezogen, usw. Die Lösungsmenge von linearen Gleichungssystemen ändert sich bei solchen Umformungen nicht:

Lemma. *Gegeben sei ein lineares Gleichungssystem wie in $(*)$. Seine erweiterte Koeffizientenmatrix $(A \mid b)$ werde durch endlich viele elementare Zeilenoperationen umgeformt zu einer neuen Matrix $(A' \mid b')$. Dann gilt $\mathscr{L}(A,b) = \mathscr{L}(A',b')$.*

Beweis. Das Vertauschen zweier Zeilen vertauscht lediglich Gleichungen in $(*)$, was die Lösungsmenge offensichtlich nicht ändert. Ebenso ändert das Multiplizieren der i-ten Zeile mit einer Zahl $\lambda \neq 0$ die Lösungsmenge nicht: Wenn $(c_1,\ldots,c_n) \in \mathbb{R}^n$ eine Lösung von $(*)$ ist, dann gilt $a_{i1}c_1 + \cdots + a_{in}c_n = b$. Durch Multiplikation mit λ folgt

$$\lambda a_{i1}c_1 + \cdots + \lambda a_{in}c_n \;=\; \lambda(a_{i1}c_1 + \cdots + a_{in}c_n) \;=\; \lambda b,$$

sodass $(c_1,\ldots,c_n)$ auch eine Lösung des umgeformten linearen Gleichungssystems ist. Umgekehrt ist jede Lösung des umgeformten Systems auch eine von $(*)$, wie man durch Multiplikation der i-ten Gleichung mit λ^{-1} sieht.

Es bleibt die elementare Umformung zu diskutieren, bei der die Matrix $(A' \mid b')$ entsteht, indem zur j-ten Zeile der Matrix $(A \mid b)$ das λ-fache der i-ten Zeile addiert wird. Wenn $(c_1,\ldots,c_n) \in \mathbb{R}^n$ eine Lösung des gegebenen Gleichungssystems $(*)$ ist, dann gilt insbesondere

$$\begin{aligned} a_{i1}c_1 + \cdots + a_{in}c_n &= b_i \\ a_{j1}c_1 + \cdots + a_{jn}c_n &= b_j. \end{aligned}$$

Hieraus folgt

$$\begin{aligned} \lambda b_i + b_j &= \lambda(a_{i1}c_1 + \cdots + a_{in}c_n) + (a_{j1}c_1 + \cdots + a_{jn}c_n) \\ &= (\lambda a_{i1} + a_{j1})c_1 + \cdots + (\lambda a_{in} + a_{jn})c_n, \end{aligned}$$

sodass $(c_1,\ldots,c_n)$ auch eine Lösung des umgeformten Gleichungssystem ist. Auch hier gilt die Umkehrung, denn wenn man im neuen Gleichungssystem von der j-ten Zeile das λ-fache der i-ten Zeile subtrahiert, erhält man das alte zurück. □

In dem Zahlenbeispiel, das wir vor dem obigen Lemma betrachtet haben, stand am Ende eine Matrix von besonders einfacher Form:

Definition. Eine Matrix A hat *reduzierte Zeilenstufenform*, wenn sie von folgender Form ist:

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

Dabei steht $*$ für beliebige Einträge, und weißen Stellen bedeuten Nullen. Genauer habe die Matrix m Zeilen, und es bezeichne a_{ij} den Eintrag der i-ten Zeile und j-ten Spalte der Matrix. Die Matrix hat reduzierte Zeilenstufenform, wenn gilt:

a) Es gibt ein $r \geq 0$, sodass die letzten $m-r$ Zeilen von A nur Nullen enthalten und jede vorige Zeilen mindestens einen von Null verschiedenen Eintrag besitzt.

b) Für $i = 1, 2, \ldots, r$ sei $j_i = \min\{j \mid a_{ij} \neq 0\}$, dann gilt:

- Zeilenstufenform: Es ist $j_1 < j_2 < \cdots < j_r$.
- Reduziertheit: Es ist $a_{ij_i} = 1$ und $a_{kj_i} = 0$ für alle $k < i$.

Die Spaltenindices $j_1, \ldots, j_r$ der 'Stufen' in der Matrix nennt man *Pivotindices*. Es gilt:

Satz (Gauß-Algorithmus). *Jede Matrix A lässt sich durch eine endliche Abfolge elementarer Zeilenumformungen auf reduzierte Zeilenstufenform bringen.*

Beweis. Wie zuvor sei a_{ij} der Eintrag der i-ten Zeile und j-ten Spalte von A. Wir dürfen annehmen, dass mindestens einer dieser Einträge von Null verschieden ist, denn sonst ist nichts zu zeigen. Wir gehen nun wie folgt vor:

- Wähle j minimal, sodass ein i existiert mit $a_{ij} \neq 0$. Nach Vertauschen der ersten und i-ten Zeile ist $a_{1j} \neq 0$. Multiplikation der ersten Zeile mit $1/a_{1j}$ liefert dann die Matrix:
$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & a_{2j} & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{mj} & * & \cdots & * \end{pmatrix}$$
- Wir subtrahieren nun sukzessive für $i = 2, \ldots, m$ von der i-ten Zeile das a_{ij}-fache der ersten Zeile, um Nullen in der j-ten Spalte zu erzeugen. Wir erhalten eine Matrix der folgenden Form:
$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{pmatrix}$$

- Durch Zeilenoperationen auf den letzten $m-1$ Zeilen wird diese Form der Matrix nicht zerstört. Per Induktion können wir durch solche Operationen den rechten unteren Block der Matrix in Zeilenstufenform bringen. Wir erhalten insgesamt eine Matrix

$$\begin{pmatrix} 1 & * & \cdots & * & ? & * & \cdots & * & \cdots & ? & * & \cdots & * \\ & & & & 1 & * & \cdots & * & \cdots & ? & * & \cdots & * \\ & & & & & & & & \cdots & & \cdots & & \\ & & & & & & & & & 1 & * & \cdots & * \\ & & & & & & & & & & & & \\ & & & & & & & & & & & & \end{pmatrix}$$

- Um alle noch verbleibenden Einträge "?" zu Null zu machen, subtrahieren wir sukzessive für $i = 2, 3, \ldots$ Vielfache der i-ten Zeile von den vorigen Zeilen. □

Wir können nun für die Lösung eines linearen Gleichungssystems elementare Zeilenoperationen auf seine erweiterte Koeffizientenmatrix $(A \mid b)$ anwenden, um eine Matrix $(A' \mid b')$ mit A' in reduzierter Zeilenstufenform zu erhalten. Dann liest man die Lösungsmenge sofort ab:

Bemerkung. Wenn die Koeffizientenmatrix A' reduzierte Zeilenstufenform hat, so hat die erweiterte Koeffizientenmatrix die Form

$$(A' \mid b') = \left(\begin{array}{ccccccccccc|c} 1 & * & \cdots & * & 0 & * & \cdots & * & \cdots & 0 & * & \cdots & * & b'_1 \\ & & & & 1 & * & \cdots & * & \cdots & 0 & * & \cdots & * & b'_2 \\ & & & & & & & & \cdots & & \cdots & & & \vdots \\ & & & & & & & & & 1 & * & \cdots & * & b'_r \\ & & & & & & & & & & & & & b'_{r+1} \\ & & & & & & & & & & & & & \vdots \\ & & & & & & & & & & & & & b'_m \end{array}\right).$$

Für das zugehörige lineares Gleichungssystem gilt dann:

a) Das lineare Gleichungssystem ist lösbar genau für $b'_{r+1} = \cdots = b'_m = 0$.

b) In diesem Fall erhält man alle Lösungen wie folgt:

- Seien $j_1, \ldots, j_r$ die Pivotindices der Matrix.
- Für die $n-r$ Variablen x_j mit $j \notin \{j_1, \ldots, j_r\}$ kann man beliebige reelle Werte einsetzen, diese Variablen bezeichnet man daher auch als *freie* Variablen.
- Die Werte der übrigen r Variablen x_j mit $j \in \{j_1, \ldots, j_r\}$ sind dann anhand der Gleichungen in der reduzierten Zeilenstufenform direkt ablesbar, sie werden daher auch als *gebundene* oder *abhängige* Variablen bezeichnet.

Wir erhalten eine Bijektion

$$\varphi\colon \quad \mathbb{R}^{n-r} \longrightarrow \mathscr{L}(A \mid b),$$

wobei die freien Variablen genau den Koordinaten von $\mathbb{R}^{n-r}$ entsprechen.

Beispiel. Für festes $a \in \mathbb{R}$ betrachte man das lineare Gleichungssystem:

$$\begin{aligned} x_1 + x_2 + 5x_3 + x_4 &= 1 \\ x_1 + x_2 + 6x_3 + 2x_4 &= 1 \\ x_1 + x_2 + 7x_3 + 3x_4 &= a \end{aligned}$$

Wir wenden den Gauß-Algorithmus auf die erweiterte Koeffizientenmatrix $(A \mid b)$ dieses linearen Gleichungssystems an:

$$\left(\begin{array}{cccc|c} 1 & 1 & 5 & 1 & 1 \\ 1 & 1 & 6 & 2 & 1 \\ 1 & 1 & 7 & 3 & a \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 1 & 5 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 2 & a-1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 1 & 5 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & a-1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 1 & 0 & -4 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & a-1 \end{array}\right)$$

Somit ist das lineare Gleichungssystem lösbar genau für $a = 1$, und in diesem Fall ist seine Lösungsmenge

$$\mathscr{L}(A \mid b) \;=\; \{(c_1, c_2, c_3, c_4) \in \mathbb{R}^4 \mid c_1 = 1 - c_2 + 4c_4 \text{ und } c_3 = -c_4\}.$$

Wir erhalten eine Bijektion $\varphi\colon \mathbb{R}^2 \to \mathscr{L}(A \mid b)$, $(u, v) \mapsto (1 - u + 4v, u, -v, v)$.

Ausblick

Die lineare Algebra hört mit dem Gauß-Algorithmus nicht auf, sondern zielt auf ein tieferes konzeptionelleres Verständnis:

- Ist die reduzierte Zeilenstufenform einer Matrix eindeutig bestimmt?
- Was ist die geometrische Bedeutung der Zahl r in der Zeilenstufenform?
- Kann man lineare Abbildungen wie die Bijektion φ in koordinatenfreier Form verstehen, oder durch Koordinatenwechsel in besonders einfache Form bringen?
- Was passiert, wenn man die reelle Zahlen ersetzt durch andere Zahlbereiche wie z.B. den Körper $\mathbb{F}_2 = \{0, 1\}$, mit dem Computer rechnen?

Das Ziel der linearen Algebra ist es, eine allgemeine Sprache für den Umgang mit linearen Strukturen zu entwickeln, mit der sich diese und viele weitere Fragen sehr einfach beantworten lassen. Diese Vorlesung wird also über weite Strecken auch ein Sprachkurs sein — und wie beim Erlernen jeder anderen Sprache benötigen Sie am Anfang etwas Geduld, werden aber am Ende (hoffentlich) reich belohnt.

Kapitel I
Gruppen, Ringe, Körper

Zusammenfassung In diesem Kapitel führen wir die grundlegenden algebraischen Strukturen ein, auf denen die lineare Algebra aufbaut. Eine Gruppe ist eine Menge mit einer assoziativen Verknüpfung, die ein neutrales Element hat und in der jedes Element invertierbar ist. Ein Ring ist eine additiv geschriebene abelsche Gruppe mit einer weiteren assoziativen Verknüpfung, der Multiplikation, die bezüglich der Addition distributiv ist. Ein Körper ist ein kommutativer Ring, in dem jedes von Null verschiedene Element ein multiplikatives Inverses hat. Wichtige Beispiele für Körper sind die Körper der rationalen, reellen und komplexen Zahlen.

1 Gruppen

Als Kind lernt man zunächst, wie man natürliche Zahlen addiert, später lernt man Subtraktion, Multiplikation und Division, wobei $\mathbb{N}$ erweitert wird zu $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$... Alle diese Strukturen beruhen auf dem Begriff einer Verknüpfung:

Definition 1.1. Ein *Monoid* ist ein Paar $(M, \circ)$ bestehend aus einer Menge M und einer Verknüpfung

$$\circ\colon \quad M \times M \longrightarrow M, \quad (a,b) \mapsto a \circ b,$$

sodass die folgenden beiden Eigenschaften erfüllt sind:

a) Die Verknüpfung ist *assoziativ*: Es ist $(a \circ b) \circ c = a \circ (b \circ c)$ für alle $a, b, c \in M$.

b) Es gibt ein *neutrales Element* für die Verknüpfung, d.h. ein Element $e \in M$ mit der Eigenschaft

$$a \circ e = e \circ a = a \quad \text{für alle} \quad a \in M.$$

Statt $a \circ b$ schreiben wir auch $a \bullet b$, $a \cdot b$, ab etc. Diese Flexibilität ist nötig, da wir später mit mehreren Verknüpfungen zugleich zu tun haben werden.

Definition 1.2. Ein Monoid $(M, \circ)$ heißt *kommutativ* oder *abelsch*, falls gilt:

$$a \circ b \;=\; b \circ a \quad \text{für alle} \quad a, b \in M.$$

Nur im kommutativen Fall verwenden wir auch die additive Notation $+$ statt $\circ$.

Beispiel 1.3. Es gilt:

a) $(\mathbb{Z}, +)$ ist ein kommutatives Monoid mit neutralem Element $e = 0$.

b) $(\mathbb{Z}, \cdot\,)$ ist ein kommutatives Monoid mit neutralem Element $e = 1$.

Aus diesem Grund bezeichnet man neutrale Elemente in Monoiden manchmal auch als *Einselemente*, aber in additiver Notation als *Nullelemente*. Caveat emptor!

Beispiel 1.4. Sei X eine beliebige Menge. Dann ist

$$M \;=\; \mathrm{Abb}(X, X) \;=\; \{\text{Abbildungen } f : X \to X\}$$

ein Monoid bezüglich der Verkettung $\circ$ von Abbildungen, mit $e = id_X$ als neutralem Element. Wenn X mindestens zwei Elemente enthält, dann ist dieses Monoid nicht kommutativ: Denn seien $a, b \in X$ zwei verschiedene Elemente. Für die konstanten Abbildungen

$$f\colon\, X \longrightarrow X,\, x \mapsto a \quad \text{und} \quad g\colon\, X \longrightarrow X,\, x \mapsto b$$

gilt dann offenbar $f \circ g \neq g \circ f$. Wenn X mindestens drei Elemente enthält, können wir auch ein Beispiel mit nicht-konstanten Abbildungen f und g konstruieren: Denn seien $a, b, c \in X$ paarweise verschieden und $f, g \in \mathrm{Abb}(X, X)$ definiert durch

$$f(x) \;:=\; \begin{cases} b & \text{für } x = a, \\ a & \text{für } x = b, \\ x & \text{sonst.} \end{cases} \qquad g(x) \;:=\; \begin{cases} c & \text{für } x = a, \\ a & \text{für } x = c, \\ x & \text{sonst.} \end{cases}$$

Dann ist $f \circ g \neq g \circ f$, denn

$$\begin{aligned} (f \circ g)(a) \;&=\; f(g(a)) \;=\; f(c) = c, \\ (g \circ f)(a) \;&=\; g(f(a)) \;=\; g(b) = b. \end{aligned}$$

Bemerkung 1.5. In der Definition von Monoiden hatten wir nur die Existenz eines neutralen Elementes gefordert, aber kein solches ausgewählt. Der Grund ist, dass es höchstens ein solches geben kann: Sei $(M, \circ)$ ein Monoid und $e_1, e_2 \in M$ seien neutrale Elemente, dann gilt

$$\begin{aligned} e_1 \;&=\; e_1 \circ e_2 && \text{(wegen } a \circ e_2 = a \text{ für alle } a \in M) \\ &=\; e_2 && \text{(wegen } e_1 \circ a = a \text{ für alle } a \in M) \end{aligned}$$

Wir bezeichnen das eindeutige neutrale Element des Monoids auch mit $e_M \in M$.

Verknüpfungen $\circ$ auf einer endlichen Menge M werden oft angegeben durch eine Tabelle, die in Zeile x und Spalte y das Element $x \circ y$ enthält. Eine solche Tabelle nennt man *Verknüpfungstafel*. Das folgende Beispiel zeigt, warum wir für neutrale Elemente zwei Bedingungen gefordert haben:

Beispiel 1.6. Sei $M = \{x, y\}$ eine Menge mit zwei Elementen, und $\circ$ sei durch die folgende Verknüpfungstafel definiert:

$\circ$	x	y
x	x	y
y	x	y

Man rechnet leicht nach, dass diese Verknüpfung assoziativ ist. Jedoch ist $(M, \circ)$ kein Monoid, denn es gibt kein neutrales Element: Zwar ist x *linksneutral* in dem Sinne, dass $x \circ a = a$ für alle $a \in M$ gilt, aber x ist nicht *rechtsneutral*, denn $a \circ x \neq a$ für $a = y$. Also ist x kein neutrales Element, und analog ist auch y kein solches.

Zum Lösen von Gleichungen wollen wir nicht nur Elemente verknüpfen, wir wollen dies auch umkehren — also beispielsweise statt addieren subtrahieren und statt multiplizieren dividieren:

Definition 1.7. Eine *Gruppe* ist ein Monoid $(G, \circ)$, sodass zu jedem Element $a \in G$ ein Element $b \in G$ existiert mit

$$a \circ b = b \circ a = e,$$

wobei e das neutrale Element des Monoids sei. Wir nennen dann b ein zu a *inverses* Element. Eine Gruppe heißt *abelsch*, wenn sie als Monoid abelsch ist.

Beispiel 1.8. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, … sind abelsche Gruppen.

Für Gruppen, die nicht abelsch sind, stellt die Definition inverser Elemente zwei Bedingungen: Ein $b \in G$ heißt *linksinvers* zu a, wenn $b \circ a = e$ ist, und *rechtsinvers*, wenn $a \circ b = e$ ist. Per Definition ist ein inverses Element also sowohl links- als auch rechtsinvers. Ähnlich wie für das neutrale Element folgt, dass auch das Inverse zu einem Gruppenelement eindeutig bestimmt ist:

Lemma 1.9. *Sei G eine Gruppe. Dann hat jedes Element $a \in G$ genau ein Inverses.*

Beweis. Für je zwei Inverse b_1, b_2 von $a \in G$ gilt

$$\begin{aligned}
b_1 &= e \cdot b_1 && \text{(weil } e \text{ linksneutrales Element)} \\
&= (b_2 \cdot a) \cdot b_1 && \text{(weil } b_2 \text{ linksinvers zu } a\text{)} \\
&= b_2 \cdot (a \cdot b_1) && \text{(wegen Assoziativität)} \\
&= b_2 \cdot e && \text{(weil } b_1 \text{ rechtsinvers zu } a\text{)} \\
&= b_2 && \text{(weil } e \text{ rechtsneutrales Element)}
\end{aligned}$$

und somit folgt die Behauptung. □

Wir bezeichnen das eindeutig bestimmte Inverse Element zu $a \in G$ mit $b = a^{-1}$ bzw. für additiv geschriebene abelsche Gruppen mit $b = -a$. Aus den Definitionen erhält man folgende nützliche Rechenregeln:

Lemma 1.10. *Sei $(G, \cdot)$ eine Gruppe. Dann gilt:*

a) Inversion: Für alle $a, b \in G$ ist $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$ und $(a^{-1})^{-1} = a$.

b) Kürzungsregel: Für alle $x, y, c \in G$ gilt $x = y \iff c \cdot x = c \cdot y \iff x \cdot c = y \cdot c$.

c) Lösbarkeit linearer Gleichungen: Für alle $a, b \in G$ gibt es eindeutige $x, y \in G$ mit

$$a \cdot x \;=\; y \cdot a \;=\; b.$$

Beweis. Wir beweisen hier nur die Inversionsregeln und überlassen den Rest als Übungsaufgabe. Per Definition gilt für beliebige $a, b \in G$:

$$b \text{ invers zu } a \iff a \cdot b = b \cdot a = e \iff a \text{ invers zu } b$$

Wenn wir hier $b = a^{-1}$ einsetzen, steht auf der linken Seite eine wahre Aussage und die rechte Seite liefert somit $a = (a^{-1})^{-1}$. Aus der Assoziativität folgt ferner

$$\begin{aligned}(b^{-1} \cdot a^{-1}) \cdot (a \cdot b) \;&=\; b^{-1} \cdot ((a^{-1} \cdot a) \cdot b) \\ &=\; b^{-1} \cdot (e \cdot b) \;=\; b^{-1} \cdot b \;=\; e,\end{aligned}$$

und analog $(a \cdot b) \cdot (b^{-1} \cdot a^{-1}) = e$, also ist $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$ wie gewünscht. □

Erfreulicherweise müssen wir zum Nachweis der Gruppeneigenschaft nur die Hälfte der Axiome prüfen, nämlich die Existenz eines linksneutralen Elements und die von linksinversen Elementen. Die andere Hälfte folgt dann automatisch:

Lemma 1.11. *Sei G eine Menge und $\cdot : G \times G \to G$ eine assoziative Verknüpfung mit folgenden beiden Eigenschaften:*

a) Es existiert ein $e \in G$ mit $e \cdot a = a$ für alle $a \in G$.

b) Zu jedem $a \in G$ existiert ein $b \in G$ mit $b \cdot a = e$.

Dann ist $(G, \cdot)$ eine Gruppe und e ist ihr neutrales Element.

Beweis. Wir zeigen zunächst, dass jedes zu a Linksinverse auch ein Rechtsinverses ist. Sei $b \in G$ mit $ba = e$. Nach Annahme hat auch b ein Linksinverses $c \in G$. Also ist $cb = e$, und es folgt:

$$\begin{aligned}
ab &= (ea)b && \text{(weil } e \text{ linksneutral)} \\
&= ((cb)a)b && \text{(wegen } cb = e) \\
&= (c(ba))b && \text{(Assoziativität)} \\
&= (ce)b && \text{(wegen } ba = e) \\
&= c(eb) && \text{(Assoziativität)} \\
&= cb && \text{(weil } e \text{ linksneutral)} \\
&= e && \text{(wegen } cb = e)
\end{aligned}$$

Also ist jedes Linksinverse auch ein Rechtsinverses. Zu zeigen bleibt nur noch, dass das gegebene linksneutrale Element $e \in G$ auch rechtsneutral ist: Sei $a \in G$. Nach Annahme existiert $b \in G$ mit $ba = e$. Aus dem vorigen Schritt wissen wir $ab = e$, und es folgt

$$\begin{aligned} ae &= a(ba) && \text{(weil } ba = e\text{)} \\ &= (ab)a && \text{(Assoziativität)} \\ &= ea && \text{(weil } ab = e\text{)} \\ &= a && \text{(weil } e \text{ linksneutral)} \end{aligned}$$

wie gewünscht. □

Korollar 1.12. *Sei $(M, \cdot)$ ein Monoid. Dann ist die Menge*

$$G = \{a \in M \mid \exists b \in M : ba = ab = e\}$$

seiner invertierbaren Elemente eine Gruppe bezüglich der Verknüpfung $\cdot$ auf M.

Beweis. Nach dem Lemma ist nur zu zeigen, dass $G \subseteq M$ abgeschlossen unter der Verknüpfung ist. Aber das ist klar: Für alle $a_1, a_2 \in G$ gibt es Inverse $b_1, b_2 \in M$, dann ist $(b_2b_1)(a_1a_2) = b_2(b_1a_1)a_2 = b_2a_2 = e = (a_1a_2)(b_2b_1)$, also $a_1a_2 \in G$. □

Beispiel 1.13. Die invertierbaren Elemente

a) in $M = (\mathbb{Z}, \cdot)$ bilden die Gruppe $G = \{\pm 1\}$ bezüglich der Multiplikation.

b) in $M = (\mathbb{Q}, \cdot)$ bilden die Gruppe $G = \mathbb{Q} \setminus \{0\}$ bezüglich der Multiplikation.

c) in $M = (\mathrm{Abb}(X,X), \circ)$ bilden bezüglich der Verkettung $\circ$ von Abbildungen die für die Beschreibung von Symmetrien wichtige Gruppe

$$\mathrm{Sym}(X) := \{ f \in \mathrm{Abb}(X,X) \mid f \text{ ist bijektiv} \}.$$

Definition 1.14. Für $n \in \mathbb{N}$ definieren wir die *symmetrische Gruppe* auf n Elementen durch

$$\mathfrak{S}_n := \mathrm{Sym}(X) \quad \textit{für} \quad X = \{1, 2, \ldots, n\}.$$

Die Elemente dieser Gruppe bezeichnet man auch als *Permutationen*, wir werden uns damit später im Kapitel über Determinanten ausführlicher beschäftigen.

Beispiel 1.15. Es gilt:

a) $\mathfrak{S}_1 = \{id\}$ ist die triviale Gruppe, die nur aus dem neutralen Element besteht.

b) $\mathfrak{S}_2 = \{id, \sigma\}$ für die Permutation $\sigma\colon \{1,2\} \to \{1,2\}$, welche die Zahlen 1 und 2 vertauscht. Die Verknüpfungstafel dieser Gruppe sieht so aus:

$\circ$	id	σ
id	id	σ
σ	σ	id

c) $\mathfrak{S}_3 = \{id, \sigma_1, \sigma_2, \sigma_3, \rho, \rho^2\}$ für die sechs wie folgt definierten Permutationen:

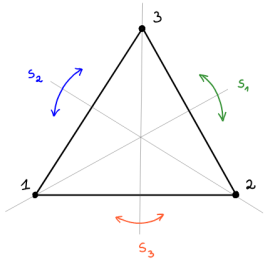
n	$id(n)$	$\sigma_1(n)$	$\sigma_2(n)$	$\sigma_3(n)$	$\rho(n)$	$\rho^2(n)$
1	1	1	3	2	2	3
2	2	3	2	1	3	1
3	3	2	1	3	1	2

Als Übung überlege man sich, wie die Verknüpfungstafel von $\mathfrak{S}_3$ aussieht. Man kann diese Gruppe als Symmetriegruppe eines gleichseitigen Dreiecks verstehen:

Beispiel 1.16. Sei $(D_3, \cdot)$ die Gruppe aller Kongruenzabbildungen der Ebene, die ein gleichseitiges Dreieck auf sich abbilden, wobei die Verkettung von Abbildungen zur Unterscheidung vom vorigen Beispiel mit $\bullet$ bezeichnet werde. Es ist

$$D_3 \;=\; \{id, s_1, s_2, s_3, r, r^2\}$$

wobei r eine Drehung um den Schwerpunkt mit Drehwinkel $2\pi/3$ ist und s_i die Spiegelung an der i-ten Seitenhalbierenden bezeichnet:



Wir haben eine Bijektion

$$f\colon\ D_3 \longrightarrow \mathfrak{S}_3 \quad \text{definiert durch} \quad f(s_i) = \sigma_i, \quad f(r^i) = \rho^i \quad \text{für} \quad i = 1,2,3.$$

Man rechnet sofort nach, dass diese Bijektion mit der Gruppenstruktur kompatibel ist in dem Sinne, dass $f(a \cdot b) = f(a) \circ f(b)$ für alle $a, b \in D_3$ gilt. Allgemeiner definieren wir:

Definition 1.17. Ein *Homomorphismus* zwischen zwei Gruppen $(G, \cdot)$ und $(H, \circ)$ ist eine Abbildung

$$f\colon \quad G \longrightarrow H \quad \text{mit} \quad f(a \cdot b) \;=\; f(a) \circ f(b) \quad \text{für alle} \quad a, b \;\in\; G.$$

Wenn f zudem bijektiv ist, nennen wir f einen *Isomorphismus*. Wir schreiben dann auch kurz

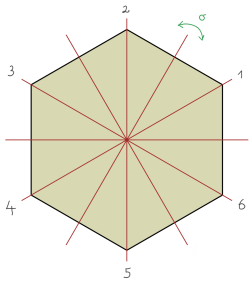
$$f\colon \quad (G, \cdot) \xrightarrow{\sim} (H, \circ).$$

und sagen, die beiden Gruppen $(H, \circ)$ und $(G, \cdot)$ seien zueinander *isomorph*.

Beispiel 1.18. Für $n \geq 3$ definieren wir die *Diedergruppe* D_n als die Gruppe aller Kongruenzabbildungen, welche ein regelmäßiges n-Eck auf sich abbilden. Diese Gruppe hat genau $2n$ Elemente: Die n Spiegelungen an den Mittelachsen und die n Drehungen um Vielfache von $2\pi/n$. Indem wir die Ecken mit $1,\ldots,n$ numerieren, können wir jedem Element von D_n eine Permutation $\sigma \in \mathfrak{S}_n$ zuordnen. Wir erhalten einen injektiven Gruppenhomomorphismus

$$D_n \longrightarrow \mathfrak{S}_n.$$

Wegen $|D_n| = 2n$ und $|\mathfrak{S}_n| = n!$ ist dieser nur für $n = 3$ ein Isomorphismus. Die folgende Abbildung illustriert die Permutation σ zu einer Achsenspiegelung eines regulären Sechsecks:



$$\begin{aligned}\sigma(1) &= 2\\ \sigma(2) &= 1\\ \sigma(3) &= 6\\ \sigma(4) &= 5\\ \sigma(5) &= 4\\ \sigma(6) &= 3\end{aligned}$$

Beispiel 1.19. Die Drehungen eines regelmäßigen n-Ecks um Vielfache von $2\pi/n$ werden beschrieben durch den Homomorphismus

$$f\colon \quad (\mathbb{Z},+) \longrightarrow (D_n,\cdot), \quad n \mapsto \text{Drehung um } 2\pi/n.$$

Homomorphismus zu sein, bedeutet hier $f(m+n) = f(m)\cdot f(n)$ für alle $m,n \in \mathbb{Z}$.

Wir haben in der Definition von Homomorphismen nur die Kompatibilität mit der Verknüpfung gefordert. Daraus folgt schon die Kompatibilität mit neutralen und inversen Elementen:

Lemma 1.20. *Sei $f : G \to H$ ein Gruppenhomomorphismus.*

a) Für die neutralen Elemente e_G und e_H gilt $f(e_G) = e_H$.

b) Für alle $a \in G$ gilt $f(a^{-1}) = (f(a))^{-1}$.

Beweis. Es ist $f(e_G) = f(e_G \cdot e_G) = f(e_G)\cdot f(e_G)$ und Multiplikation mit $(f(e_G))^{-1}$ liefert sofort a). Teil b) folgt analog. □

2 Untergruppen

Häufig hat man es mit Teilmengen einer Gruppe zu tun, die stabil sind unter der Verknüpfung in der Gruppe und bezüglich dieser selber eine Gruppe bilden, wie z.B. die Menge aller Drehungen in der Diedergruppe D_{2n}:

Definition 2.1. Eine *Untergruppe* einer Gruppe $(G,\cdot)$ ist eine Teilmenge $H \subseteq G$, sodass folgende Bedingungen gelten, wobei e das neutrale Element von G sei:

a) Es ist $e \in H$.

b) Für alle $a \in H$ ist auch $a^{-1} \in H$.

c) Für alle $a,b \in H$ ist auch $a \cdot b \in H$.

Die erste Bedingung besagt insbesondere, dass $H \neq \varnothing$ ist. Nimmt man dies an, so lassen sich die obigen drei Axiome auch etwas eleganter zusammenfassen:

Lemma 2.2. *Sei $(G,\cdot)$ eine Gruppe. Eine Teilmenge $H \subseteq G$ ist eine Untergruppe genau dann, wenn sie nichtleer ist und wenn gilt:*

$$\alpha \cdot \beta^{-1} \in H \quad \textit{für alle} \quad \alpha,\beta \in H.$$

Beweis. Jede Untergruppe erfüllt offenbar diese Bedingung. Umgekehrt folgen aus der Bedingung alle Untergruppenaxiome:

a) $e \in H$ (wähle $\alpha = \beta \in H$ beliebig),

b) Für jedes $a \in H$ ist $a^{-1} \in H$ (wähle $\alpha = e, \beta = a$),

c) Für alle $a,b \in H$ ist $a \cdot b \in H$ (wähle $\alpha = a, \beta = b^{-1}$). □

Jede Gruppe G besitzt offenbar die Teilmengen $H = \{e\} \subseteq G$ und $H = G$ als Untergruppen. Diese werden auch als die *trivialen Untergruppen* bezeichnet. Ein etwas interessanteres Beispiel:

Beispiel 2.3. Für $n \in \mathbb{Z}$ ist die Teilmenge $n\mathbb{Z} := \{nk \mid k \in \mathbb{Z}\} \subseteq \mathbb{Z}$ eine Untergruppe in der additiven Gruppe $(\mathbb{Z},+)$:

- Es ist $0 = n \cdot 0 \in n\mathbb{Z}$ und somit $n\mathbb{Z} \neq \varnothing$.
- Für $\alpha,\beta \in n\mathbb{Z}$ ist $\alpha = nk, \beta = nl$ mit $k,l \in \mathbb{Z}$ und somit

$$\alpha - \beta \;=\; nk - nl \;=\; n(k-l) \;\in\; n\mathbb{Z}.$$

Wir haben damit bereits *alle* Untergruppen von $(\mathbb{Z},+)$ gefunden:

Lemma 2.4. *Jede Untergruppe der additiven Gruppe $(\mathbb{Z},+)$ hat die Form $H = n\mathbb{Z}$ für ein eindeutiges $n \in \mathbb{N}_0$.*

Beweis. Im Fall $H \cap \mathbb{N} = \varnothing$ ist $H = \{0\}$, da additive Untergruppen unter $x \mapsto -x$ stabil sind. Wir können dann $n = 0$ wählen. Im Fall $H \cap \mathbb{N} \neq \varnothing$ sei $n := \min H \cap \mathbb{N}$. Für jedes $h \in H$ gibt Division mit Rest eine Darstellung

$$\begin{aligned} & h = kn + r \quad \text{mit} \quad r \in \{0,1,\ldots,n-1\}, k \in \mathbb{Z} \\ \Longrightarrow \quad & r = h - kn \in H \cap \{0,1,\ldots,n-1\} \\ \Longrightarrow \quad & r = 0 \quad \text{nach Wahl von } n = \min H \cap \mathbb{N} \end{aligned}$$

Also gilt $H = n\mathbb{Z}$. Die Eindeutigkeit folgt analog. □

Für kompliziertere Gruppen ist es meist nicht so einfach, alle ihre Untergruppen zu bestimmen. Für *endliche* Gruppen G erhalten wir immerhin einige Information aus der Anzahl ihrer Elemente, der sogenannten *Ordnung* $|G|$:

Satz 2.5 (Satz von Lagrange). *Sei G eine endliche Gruppe, und sei $H \subseteq G$ eine beliebige Untergruppe. Dann ist die Ordnung $|H|$ ein Teiler der Ordnung $|G|$.*

Beweis. Wir definieren auf G eine Relation $\sim$ durch $a \sim b \iff ab^{-1} \in H$. Dies ist eine Äquivalenzrelation:

- Reflexivität: Für alle $a \in G$ ist $a \sim a$ wegen $aa^{-1} = e \in H$.
- Symmetrie: Aus $a \sim b$ folgt $b \sim a$ wegen $ba^{-1} = (ab^{-1})^{-1} \in H$.
- Transitivität: Aus $a \sim b$ und $b \sim c$ folgt $a \sim c$ wegen $ac^{-1} = ab^{-1} \cdot bc^{-1} \in H$.

Wir haben hier alle Untergruppenaxiome benutzt! Die Äquivalenzklassen bzgl. $\sim$ haben die Form

$$[b] \;=\; \{a \in G \mid ab^{-1} \in H\} \;=\; \{hb \in G \mid h \in H\} \;=\; Hb$$

wobei wir für die zweite Gleichung $h = ab^{-1}$ substituiert haben. Jede Äquivalenzklasse enthält genau $|H|$ Elemente, da die Abbildung

$$Hb \longrightarrow H, \quad a \mapsto ab^{-1}$$

bijektiv ist mit der Umkehrabbildung $H \longrightarrow Hb, h \mapsto hb$. Seien $b_1, \ldots, b_r \in G$ ein vollständiges Repräsentantensystem für die endlich vielen Äquivalenzklassen, es sei also $G = Hb_1 \cup \cdots \cup Hb_r$ und $Hb_i \cap Hb_j = \varnothing$ für alle $i \neq j$. Zählen der Elemente zeigt

$$|G| \;=\; \sum_{i=1}^{r} |Hb_i| \;=\; \sum_{i=1}^{r} |H| \;=\; r \cdot |H|$$

und somit ist $|G|$ teilbar durch $|H|$. □

Beispiel 2.6. Die einzigen nichttrivialen Untergruppen von $D_3 = \{id, s_1, s_2, s_3, r, r^2\}$ sind

$$H = \{id, s_i\} \quad \text{für } i = 1,2,3, \quad \text{und} \quad H = \{id, r, r^2\}.$$

Beweis. Man sieht leicht, dass die angegebenen Teilmengen Untergruppen sind. Sei umgekehrt eine beliebige nichttriviale Untergruppe $H \subseteq \mathfrak{S}_3$ gegeben. Nach dem Satz von Lagrange ist $|H| \in \{2,3\}$. Wir unterscheiden nun folgende Fälle:

a) Falls $s_i \in H$ ist, so enthält H die Untergruppe $\{id, s_i\}$ der Ordnung zwei. Wieder nach dem Satz von Lagrange muß also die Gruppenordnung $|H|$ eine gerade Zahl sein. Wegen $|H| \in \{2,3\}$ folgt $|H| = 2$ und somit $H = \{id, s_i\}$.

b) Falls $r \in H$ ist, gilt $H \supseteq \{id, r, r^2\}$, und wegen $|H| \leq 3$ folgt $H = \{id, r, r^2\}$.

c) Falls $r^2 \in H$ ist, gilt $r = r \circ id = r \circ r^3 = (r^2)^2 \in H$ und wir sind in Fall b). □

3 Bild, Kern und Quotienten

Für einen Gruppenhomomorphismus $f\colon G \to H$ sind das *Bild* und der *Kern* definiert als

$$\operatorname{im}(f) \;=\; \{f(g) \mid g \in G\} \;\subseteq\; H \quad \text{und} \quad \ker(f) \;=\; \{g \in G \mid f(g) = e\} \;\subseteq\; G,$$

wobei $e \in H$ das neutrale Element bezeichne. Es gilt:

Lemma 3.1. *Sei $f\colon G \to H$ ein Gruppenhomomorphismus.*

a) Das Bild $\operatorname{im}(f)$ *ist eine Untergruppe von H,*

b) Der Kern $\ker(f)$ *ist eine Untergruppe von G.*

Beweis. Es ist $\ker(f) \neq \varnothing$, da das neutrale Element im Kern liegt. Für $\alpha, \beta \in \ker(f)$ ist ferner

$$\begin{aligned} f(\alpha\beta^{-1}) \;&=\; f(\alpha) f(\beta^{-1}) && (f \text{ Homomorphismus}) \\ &=\; f(\alpha) \cdot (f(\beta))^{-1} && (\text{nach Lemma 1.20}) \\ &=\; e \cdot e^{-1} && (\text{wegen } \alpha, \beta \in \ker(f)) \\ &=\; e \end{aligned}$$

und somit $\alpha\beta^{-1} \in \ker(f)$. Für $\operatorname{im}(f)$ argumentiert man analog. □

Beispiel 3.2. Für $f\colon \mathbb{Z} \longrightarrow D_n, a \mapsto$ (Drehung um $2\pi a/n$) erhalten wir die bereits bekannten Untergruppen

$$\begin{aligned} \ker(f) &= n\mathbb{Z} \;\subseteq\; \mathbb{Z}, \\ \operatorname{im}(f) &= \{\text{Drehungen um Vielfache von } 2\pi/n\} \;\subseteq\; D_n. \end{aligned}$$

Anschaulich mißt der Kern, welche Information bei Anwenden von f verloren geht: Beispielsweise lässt sich aus einer Drehung um den Winkel $2\pi a/n$ die ganze Zahl $a \in \mathbb{Z}$ nur bis auf Addition von Elementen von $n\mathbb{Z}$ rekonstruieren. Allgemein hat der Kern die folgende Interpretation:

Lemma 3.3. *Sei $f\colon G \to H$ ein Gruppenhomomorphismus. Für $a, b \in G$ gilt dann:*

$$f(a) \;=\; f(b) \quad \Longleftrightarrow \quad ab^{-1} \;\in\; \ker(f).$$

Insbesondere ist f injektiv genau dann, wenn $\ker(f) = \{e\}$ *ist.*

Beweis. Es ist $f(a) = f(b) \Leftrightarrow f(ab^{-1}) = f(a)f(b)^{-1} = e \Leftrightarrow ab^{-1} \in \ker(f)$. Somit folgt

$$f^{-1}(c) \;=\; \begin{cases} \ker(f) \cdot b & \text{falls } c = f(b), \\ \varnothing & \text{falls } c \notin \operatorname{im}(f) \end{cases}$$

und daher ist f injektiv genau dann, wenn $|\ker(f)| = 1$ ist. □

Der 'Informationsverlust' bei Anwenden von f wird also beschrieben durch die Untergruppe $\ker(f)$. Für *abelsche* Gruppen gibt es umgekehrt zu jeder Untergruppe einen Homomorphismus mit genau dieser Untergruppe als Kern:

Satz 3.4. *Sei G eine abelsche Gruppe. Dann gibt es für jede Untergruppe $K \subseteq G$ eine abelsche Gruppe G/K und einen surjektiven Homomorphismus*

$$p: \quad G \longrightarrow G/K \quad \textit{mit} \quad \ker(p) = K.$$

Beweis. Wir definieren eine Relation $\sim$ auf G durch $a \sim b \Longleftrightarrow a - b \in K$. Diese Relation $\sim$ ist, wie wir aus dem Beweis des Satzes von Lagrange wissen,

a) reflexiv: Für alle $a \in G$ ist $a \sim a$ wegen $a - a = 0 \in K$.

b) symmetrisch: Aus $a \sim b$ folgt $b \sim a$ wegen $b - a = -(a-b) \in K$.

c) transitiv: Aus $a \sim b$ und $b \sim c$ folgt $a \sim c$ wegen $a - c = (a-b) + (b-c) \in K$.

Also ist $\sim$ eine Äquivalenzrelation. Sei $G/K := G/\sim$ der Quotient, d.h. die Menge der Äquivalenzklassen. Wir wollen diesen Quotient zu einer Gruppe machen mit der Verknüpfung

$$+: \quad G/K \times G/K \longrightarrow G/K, \quad [a] + [b] := [a+b].$$

Diese Verknüpfung ist wohldefiniert:

- Sei $[a] = [a']$ und $[b] = [b']$.
- Dann ist $a \sim a'$ und $b \sim b'$, also ist $a - a' \in K$ und $b - b' \in K$.
- Somit folgt $(a+b) - (a'+b') = (a-a') + (b-b') \in K$.
- Folglich ist $a + b \sim a' + b'$, d.h. $[a+b] = [a'+b']$ wie gewünscht.

Dass G/K mit der soeben definierten Verknüpfung eine Gruppe bildet und dass die Abbildung $p : G \to G/K, a \mapsto [a]$ ein Gruppenhomomorphismus ist, rechnet man sofort nach. Zudem ist p surjektiv und $\ker(p) = \{g \in G \mid g - 0 \in K\} = K$. □

Triviale Extremfälle dieser Konstruktion sind $G/\{0\} \simeq G$ und $G/G \simeq \{0\}$. Etwas interessanter ist das folgende Beispiel:

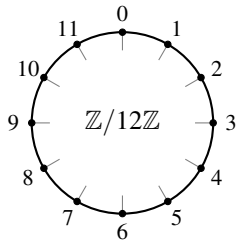
Beispiel 3.5. Für die additive Gruppe $G = \mathbb{Z}$ und $K = m\mathbb{Z}$ mit $m > 0$ ist die soeben konstruierte Äquivalenzrelation die Kongruenz modulo m: Die Äquivalenzklassen sind gegeben durch

$$\begin{aligned} [a] = [b] \ \text{ in } \mathbb{Z}/m\mathbb{Z} \ &\Longleftrightarrow \ a \equiv b \pmod m \\ &\Longleftrightarrow \ a - b \text{ ist durch } m \text{ teilbar} \end{aligned}$$

Wir haben aus der Menge $\mathbb{Z}/m\mathbb{Z} = \big\{[0],[1],[2],\ldots,[m-1]\big\}$ eine abelsche Gruppe gemacht mit

$$[a] + [b] = \big[\text{Rest bei Division von } a+b \text{ durch } m\big]$$

Für $m = 12$ kennen wir diese Addition vom Rechnen mit Uhrzeiten:



$[9]+[4]=[1]$ in $\mathbb{Z}/12\mathbb{Z}$

4 Ringe und Körper

Bisher haben wir immer nur eine Verknüpfung auf einmal betrachtet. Ein Ring ist eine Menge mit zwei Verknüpfungen, die sich ähnlich wie die Addition und die Multiplikation ganzer Zahlen verhalten:

Definition 4.1. Ein *Ring* ist ein Tripel $(R,+,\cdot)$, bestehend aus einer Menge R mit zwei Verknüpfungen

$$+\colon R\times R \longrightarrow R \quad \text{und} \quad \cdot\colon R\times R \longrightarrow R$$

sodass folgende drei Eigenschaften gelten:

a) $(R,+)$ ist eine abelsche Gruppe,

b) $(R,\cdot)$ ist ein Monoid,

c) Für alle $a,b,c \in R$ gelten die *Distributivgesetze*

$$(a+b)\cdot c = a\cdot c + b\cdot c, \quad c\cdot(a+b) = c\cdot a + c\cdot b.$$

Wir nennen das neutrale Element von $+$ das *Nullelement* $0 \in R$, und das neutrale Element von $\cdot$ das *Einselement* $1 \in R$. Falls $a\cdot b = b\cdot a$ für alle Elemente $a,b \in R$ ist, nennen wir R einen *kommutativen* Ring.

Beispiel 4.2. Es gilt:

a) $(\mathbb{Z},+,\cdot)$ und $(\mathbb{Q},+,\cdot)$ sind kommutative Ringe.

b) $(\mathbb{N},+,\cdot)$ ist kein Ring, da $(\mathbb{N},+)$ keine Gruppe ist.

c) $(2\mathbb{Z},+,\cdot)$ ist kein Ring, da $(2\mathbb{Z},\cdot)$ kein Monoid ist.

In Ringen kann man mittels Assoziativität und Distributivität wie in den ganzen Zahlen rechnen, z.B. gilt

- $0\cdot a = 0$, denn $0\cdot a = 0\cdot a + 0\cdot a - 0\cdot a = (0+0)\cdot a - 0\cdot a = 0\cdot a - 0\cdot a = 0$.
- $(-1)\cdot a = -a$, denn $a + (-1)\cdot a = 1\cdot a + (-1)\cdot a = (1+(-1))\cdot a = 0\cdot a = 0$.

Bemerkung 4.3. Wir haben in der Definition von Ringen nicht gefordert, dass das Einselement vom Nullelement verschieden ist. Wenn in einem Ring R aber $1 = 0$ gilt, folgt

$$a \;=\; a \cdot 1 \;=\; a \cdot 0 \;=\; 0$$

für alle $a \in R$ und somit ist R der (nicht sehr interessante) *Nullring* $R = \{0\}$.

Wir können nun addieren, subtrahieren und Multiplizieren. Was noch fehlt, ist die Division. Die Frage nach multiplikativen Inversen führt auf die folgende

Definition 4.4. Die *Einheitengruppe* eines Ringes R ist die Gruppe

$$R^\times \;:=\; \big\{\, r \in R \mid \exists s \in R :\; s \cdot r = r \cdot s = 1 \,\big\},$$

die von den invertierbaren Elemente des Monoids $(R, \cdot\,)$ gebildet wird; dass diese eine Gruppe bezüglich der Multiplikation bilden, folgt aus Korollar 1.12.

Beispiel 4.5. Es ist $\mathbb{Z}^\times = \{\pm 1\}$, $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}$, und $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$.

Kommutative Ringe, in denen wie in $\mathbb{Q}$ oder in $\mathbb{R}$ jedes von Null verschiedene Element multiplikativ invertierbar ist, haben einen eigenen Namen:

Definition 4.6. Ein *Körper* ist ein kommutativer Ring K mit $K^\times = K \setminus \{0\}$.

Man beachte, dass der Nullring kein Körper ist, denn für den Nullring $R = \{0\}$ gilt $R^\times = R$. Körper sind die grundlegenden Zahlbereiche, auf denen die lineare Algebra aufbaut. Man kann darin prima rechnen, z.B. gilt die Kürzungsregel:

Lemma 4.7. *Sei K ein Körper. Dann gilt:*

a) Für alle $a, b \in K \setminus \{0\}$ ist auch $a \cdot b \neq 0$.

b) Seien $a \in K \setminus \{0\}$ und $b_1, b_2 \in K$ mit $a \cdot b_1 = a \cdot b_2$, dann ist $b_1 = b_2$.

Beweis. Teil *a)* folgt aus der Tatsache, dass für Körper die Menge $K \setminus \{0\} = K^\times$ eine Gruppe bezüglich der Multiplikation bildet und somit mit je zwei Elementen auch ihr Produkt enthält. Teil *b)* folgt aus *a)* durch Betrachten von $b := b_1 - b_2$. □

Die Kürzungsregel gilt natürlich nicht nur in Körpern, sondern beispielsweise auch im Ring $R = \mathbb{Z}$. Kommutative Ringe, in denen die Kürzungsregel gilt, haben einen eigenen Namen:

Definition 4.8. Ein *Integritätsring* ist ein kommutativer Ring $R \neq \{0\}$ mit $ab \neq 0$ für alle von Null verschiedenen $a, b \in R \setminus \{0\}$.

Wichtige Beispiele für kommutative Ringe, die wir häufiger antreffen werden, sind die folgenden endlichen Ringe:

Lemma 4.9. *Für $n \in \mathbb{N}$ ist die Gruppe $R = \mathbb{Z}/n\mathbb{Z}$ ein kommutativer Ring bezüglich der repräsentantenweise definierten Multiplikation*

$$[a] \cdot [b] \;:=\; [a \cdot b].$$

Beweis. Als additive Gruppe kennen wir $\mathbb{Z}/n\mathbb{Z}$ schon, die Wohldefiniertheit der Multiplikation folgt analog:

- Sei $[a_1] = [a_2]$ und $[b_1] = [b_2]$ in $\mathbb{Z}/n\mathbb{Z}$.
- Es folgt $a_2 = a_1 + kn$ und $b_2 = b_1 + ln$ mit $k, l \in \mathbb{Z}$.
- Dann ist $a_2 b_2 = a_1 b_1 + (a_1 l + k b_1 + kln)n \equiv a_1 b_1 \pmod{n}$, also $[a_2 b_2] = [a_1 b_1]$.

Mit der so definierten Multiplikation wird $\mathbb{Z}/n\mathbb{Z}$ ein kommutativer Ring, da $\mathbb{Z}$ ein solcher ist. □

Beispiel 4.10. Der Ring $\mathbb{Z}/2\mathbb{Z}$ ist ein Körper, dieser Körper mit zwei Elementen ist die Basis für alle Computer. Seine Addition und Multiplikation sehen so aus, wobei wir der Einfachheit halber die eckigen Klammern um Ringelemente weglassen:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Andererseits ist $\mathbb{Z}/4\mathbb{Z}$ *kein* Körper, nicht einmal ein Integritätsring: Hier gilt die Kürzungsregel nicht, denn

$$[2] \cdot [2] \;=\; [0], \quad \text{aber} \quad [2] \neq [0] \quad \text{in} \quad \mathbb{Z}/4\mathbb{Z}.$$

Die Additions- und Multiplikationstafel des Ringes $\mathbb{Z}/4\mathbb{Z}$ sieht so aus, wobei wir wieder die eckigen Klammern um Ringelemente weglassen:

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\cdot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Allgemein gilt:

Lemma 4.11. *Für $p \in \mathbb{N}$ sind äquivalent:*

a) $\mathbb{Z}/p\mathbb{Z}$ ist ein Integritätsring.

b) p ist eine Primzahl.

Beweis. Es gilt:

$$\begin{aligned} p > 1 \text{ prim} &\iff \text{Für } a, b \in \mathbb{Z} \text{ gilt: } p \mid ab \text{ impliziert } p \mid a \text{ oder } p \mid b \\ &\iff \text{In } \mathbb{Z}/p\mathbb{Z} \text{ gilt: } [a] \cdot [b] = 0 \text{ impliziert } [a] = 0 \text{ oder } [b] = 0. \end{aligned}$$

Die letzte Bedingung besagt genau, dass $\mathbb{Z}/p\mathbb{Z}$ ein Integritätsring ist. □

Tatsächlich ist für Primzahlen p der Integritätsring $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ sogar ein Körper aus folgendem Grund:

Lemma 4.12. *Jeder endliche Integritätsring ist ein Körper.*

Beweis. Sei R ein endlicher Integritätsring und $a \in R \setminus \{0\}$. Nach der Kürzungsregel ist die Abbildung $R \to R, b \mapsto ab$ injektiv. Also ist diese Abbildung auch surjektiv, da R endlich ist. Folglich existiert ein $b \in R$ mit $ab = 1$. □

Beispiel 4.13. In dem Körper $\mathbb{F}_7 = \mathbb{Z}/7\mathbb{Z}$ sind multiplikative Inverse gegeben durch

$$\begin{aligned}
[1]^{-1} &= [1] && \textit{wie in jedem Körper},\\
[2]^{-1} &= [4] && \textit{wegen } 2\cdot 4 \equiv 1 (\mathrm{mod}\, 7),\\
[3]^{-1} &= [5] && \textit{wegen } 3\cdot 5 \equiv 1 (\mathrm{mod}\, 7),\\
[4]^{-1} &= [2] && \textit{wegen } 4\cdot 2 \equiv 1 (\mathrm{mod}\, 7),\\
[5]^{-1} &= [3] && \textit{wegen } 5\cdot 3 \equiv 1 (\mathrm{mod}\, 7),\\
[6]^{-1} &= [6] && \textit{wegen } 6\cdot 6 \equiv 1 (\mathrm{mod}\, 7).
\end{aligned}$$

Wie für Gruppen, so betrachtet man auch für Ringe gern Abbildungen, welche mit der Ringstruktur kompatibel sind:

Definition 4.14. Ein *Ringhomomorphismus* ist eine Abbildung $f\colon R \to S$ zwischen zwei Ringen, sodass die folgenden Eigenschaften gelten:

a) Für alle $a, b \in R$ ist $f(a+b) = f(a) + f(b)$ und $f(a \cdot b) = f(a) \cdot f(b)$,

b) Für die Einselemente $1_R \in R$ und $1_S \in S$ der Ringe gilt $f(1_R) = 1_S$.

Man beachte, dass jeder Ringhomomorphismus nach Bedingung a) insbesondere ein Homomorphismus additiver Gruppen ist, sodass für die Nullelemente der beiden Ringe automatisch $f(0_R) = 0_S$ gilt.

Beispiel 4.15. Für $n \in \mathbb{N}$ ist die Restklassenabbildung $f\colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}, a \mapsto [a]$ ein Ringhomomorphismus. Wegen

$$[a] \;=\; [\text{Rest von } a \text{ bei Division durch } n]$$

lässt sich daher mit Resten bei Division durch n prima rechnen: Was ist z.B. die letzte Dezimalziffer von

$$3^{1000} \;\approx\; 1.3221 \cdot 10^{477}?$$

Durch Rechnen in $\mathbb{Z}/10\mathbb{Z}$ erhalten wir:

$$\begin{aligned}
& [3^2] = [9] = [-1]\\
\Longrightarrow\quad & [3^4] = [3^2]^2 = [-1]^2 = [1]\\
\Longrightarrow\quad & [3^{1000}[= [3^4]^{250} = [1]^{250} = [1]\\
\Longrightarrow\quad & \text{Die letzte Ziffer von } 3^{1000} \text{ ist eine Eins.}
\end{aligned}$$

5 Die komplexen Zahlen

Als wichtiges Beispiel für einen Körper wollen wir hier noch die komplexen Zahlen betrachten, die eine Rolle in vielen Anwendungen spielen. Um ihre Konstruktion zu verstehen, sollte man sich klar machen, dass der Begriff von *Zahlen* historisch immer wieder erweitert wurde, um neue Probleme zu lösen:

Beispiel 5.1. Sei $m = 2$. Die Gleichung $z^2 = m$ hat keine rationale Lösung $z \in \mathbb{Q}$, aber wenn wir von den rationalen zu reellen Zahlen übergehen, finden wir die beiden reellen Lösungen $z = \pm\sqrt{m} \in \mathbb{R}$. Für diese eine Gleichung hätten wir natürlich nicht den gesamten Körper der reellen Zahlen benötigt: Man rechnet leicht nach, dass schon die Teilmenge

$$K := \{x + y\sqrt{m} \in \mathbb{R} \mid x, y \, \mathbb{Q}\} \subset \mathbb{R}$$

einen Körper bezüglich der Addition und Multiplikation reeller Zahlen bildet. Um den Körper K ohne Benutzung reeller Zahlen zu beschreiben, beachte man, dass die Abbildung

$$f\colon \quad M := \mathbb{Q}^2 \longrightarrow K, \quad (x,y) \mapsto x + y\sqrt{m}$$

bijektiv ist. Mittels dieser Bijektion übersetzt sich die Addition und Multiplikation des Körpers K zu entsprechenden Verknüpfungen auf der Menge M. Explizit führt dies wegen

$$\begin{aligned}(x_1 + y_1\sqrt{m}) + (x_2 + y_2\sqrt{m}) &= (x_1 + x_2) + (y_1 + y_2)\sqrt{m} \\ (x_1 + y_1\sqrt{m}) \cdot (x_2 + y_2\sqrt{m}) &= (x_1x_2 + 2y_1y_2) + (x_1y_2 + x_2y_1)\sqrt{m}\end{aligned}$$

auf die Definition

$$\begin{aligned}(x_1, y_1) + (x_2, y_2) &:= (x_1 + x_2,\ y_1 + y_2), \\ (x_1, y_1) \cdot (x_2, y_2) &:= (x_1x_2 + my_1y_2,\ x_1y_2 + x_2y_1)\end{aligned}$$

für $(x_1, y_1), (x_2, y_2) \in M$. Diese Definition kann man verstehen, ohne $\mathbb{R}$ zu kennen!

Nach dieser Übung sollte man keine Angst mehr vor quadratischen Gleichungen haben, die keine reellen Lösungen besitzen. Um einen Körper zu finden, der den Körper der reellen Zahlen enthält und in dem die Gleichung $z^2 = -1$ lösbar wird, ersetzen wir im obigen Beispiel einfach m durch -1:

Definition 5.2. Auf der Menge $\mathbb{C} := \mathbb{R}^2$ definieren wir Verknüpfungen $+$ und $\cdot$ durch

$$\begin{aligned}(x_1, y_1) + (x_2, y_2) &:= (x_1 + x_2,\ y_1 + y_2), \\ (x_1, y_1) \cdot (x_2, y_2) &:= (x_1x_2 - y_1y_2,\ x_1y_2 + y_1x_2).\end{aligned}$$

Man rechnet leicht nach, dass diese Addition und Multiplikation die Menge $\mathbb{C}$ zu einem Körper macht, wir nennen diesen den Körper der *komplexen Zahlen*.

Wir betrachten die reellen Zahlen als Teilmenge der komplexen Zahlen via der injektiven Abbildung $\mathbb{R} \to \mathbb{C}, x \mapsto (x,0)$. Wegen $(x_1,0)+(x_2,0) = (x_1+x_2,0)$ und $(x_1,0)\cdot(x_2,0) = (x_1\cdot x_2,0)$ muß man hier die Notation für die Addition und Multiplikation reeller Zahlen nicht von der für komplexe Zahlen unterscheiden. Wir schreiben kurz

- $0 := (0,0) \in \mathbb{C}$ für das Nullelement,
- $1 := (1,0) \in \mathbb{C}$ für das Einselement,
- $i := (0,1) \in \mathbb{C}$ für die sogenannte *imaginäre Einheit*.

Letztere löst die quadratische Gleichung: $i^2 = (0,1)\cdot(0,1) = (-1,0) = -1$. Jede komplexe Zahl lässt sich schreiben als $z = x+iy$ mit eindeutigen $x,y \in \mathbb{R}$. Dabei heißt

- $\mathrm{Re}(z) := x$ der *Realteil* von z,
- $\mathrm{Im}(z) := y$ der *Imaginärteil* von z.

Für $z = x+iy$ mit $x,y \in \mathbb{R}$ definieren wir das *komplex Konjugierte* $\overline{z} := x-iy$. Man rechnet sofort nach, dass

$$\overline{z+w} \;=\; \overline{z}+\overline{w} \quad \text{und} \quad \overline{z\cdot w} \;=\; \overline{z}\cdot\overline{w} \quad \text{für alle} \quad z,w \;\in\; \mathbb{C}$$

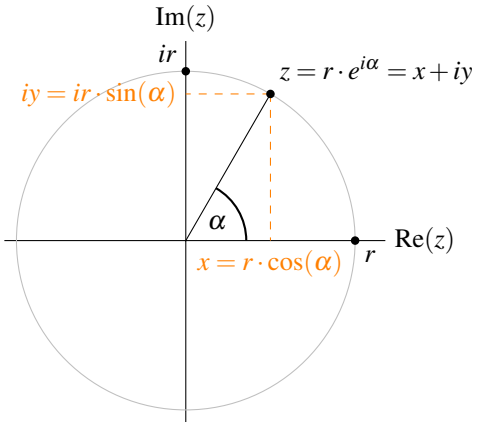
ist. Wir definieren ferner den *Absolutbetrag* $|z| = \sqrt{x^2+y^2} \in \mathbb{R}_{\geq 0}$. Wegen $|z|^2 = z\cdot\overline{z}$ gilt

$$\frac{1}{z} \;=\; \frac{1}{|z|^2}\cdot\overline{z} \;=\; \frac{x}{x^2+y^2} - i\cdot\frac{y}{x^2+y^2} \quad \text{für} \quad z \;=\; x+iy \;\neq\; 0.$$

Die Multiplikation von komplexen Zahlen lässt sich besonders gut verstehen, wenn wir $(x,y) \in \mathbb{R}^2$ als Punkt in der reellen Ebene auffassen und dann in dieser Ebene zu Polarkoordinaten übergehen, d.h.

$$x \;=\; r\cdot\cos(\alpha) \quad \text{und} \quad y \;=\; r\cdot\sin(\alpha)$$

mit einer reellen Zahl $r \geq 0$ und einem Winkel $\alpha \in \mathbb{R}$ schreiben:



In dieser *Polardarstellung* wird $z = r \cdot (\cos(\alpha) + i \cdot \sin(\alpha))$ mit $r = |z|$. Um mit den Winkeln für Punkte auf dem Einheitskreis bequem zu rechnen, verwenden wir die Kurznotation

$$e^{i\alpha} := \cos(\alpha) + i \cdot \sin(\alpha) \quad \text{für} \quad \alpha \in \mathbb{R}.$$

Diese Notation erklärt sich daraus, dass man diesen Ausdruck in der Analysis als Wert der Exponentialfunktion im Punkt $i\alpha \in \mathbb{C}$ interpretieren kann. Wir können nun die Polardarstellung von komplexen Zahlen schreiben als

$$z = r \cdot e^{i\alpha} \quad \text{mit} \quad r = |z| \quad \text{und} \quad \alpha \in \mathbb{R}.$$

Das ist für die Multiplikation komplexer Zahlen sehr nützlich:

Korollar 5.3. *Für $z = r \cdot e^{i\alpha}$ und $w = s \cdot e^{i\alpha}$ ist $z \cdot w = rs \cdot e^{i(\alpha+\beta)}$.*

Beweis. Für $\alpha, \beta \in \mathbb{R}$ ist $e^{i\alpha} \cdot e^{i\beta} = e^{i(\alpha+\beta)}$ wegen der aus der Analysis bekannten Additionsformeln

$$\begin{aligned} \cos(\alpha+\beta) &= \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \\ \sin(\alpha+\beta) &= \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) \end{aligned}$$

für die Sinus- und Cosinusfunktion. □

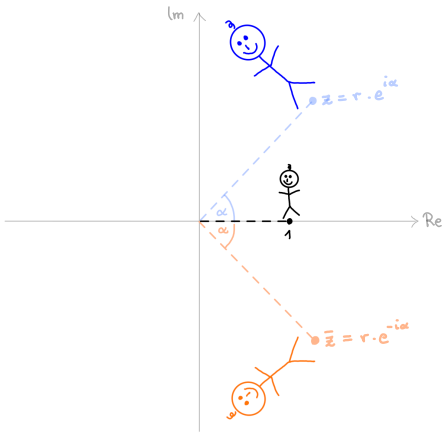
Wenn wir $\mathbb{C} = \mathbb{R}^2$ mit der Ebene identifizieren, ist somit für festes $z = r \cdot e^{i\alpha} \neq 0$ die Abbildung

$$\mathbb{C} \longrightarrow \mathbb{C}, \quad w \mapsto z \cdot w$$

eine Drehstreckung mit Streckfaktor $r > 0$ und Drehwinkel $\alpha \in \mathbb{R}$. Ebenso zeigt die Formel

$$\overline{r \cdot e^{i\alpha}} = r \cdot (\cos(\alpha) - i\sin(\alpha)) = r \cdot (\cos(-\alpha) + i\sin(-\alpha)) = r \cdot e^{-i\alpha}$$

dass die komplexe Konjugation einer Spiegelung an der reellen Achse entspricht:



Wir hatten die komplexen Zahlen eingeführt, um die Gleichung $z^2+1=0$ lösen zu können. Tatsächlich haben wir damit viel mehr erreicht, auch wenn wir das hier nicht beweisen wollen:

Satz 5.4 (Fundamentalsatz der Algebra). *Jedes Polynom*

$$f(z) \;=\; z^d + a_{d-1}z^{d-1} + \cdots + a_1 z + a_0$$

vom Grad $d > 0$ mit Koeffizienten $a_0, \ldots, a_{d-1} \in \mathbb{C}$ hat mindestens eine komplexe Nullstelle, d.h. es existiert mindestens eine komplexe Zahl $z_1 \in \mathbb{C}$ mit $f(z_1) = 0$.

Durch Ausklammern des durch eine Nullstelle gegebenen Linearfaktors $z - z_1$ erhalten wir dann eine Faktorisierung $f(z) = (z-z_1)\cdot g(z)$ für ein Polynom $g(z)$ vom Grad $d-1$. Induktiv folgt

$$f(z) \;=\; \prod_{i=1}^{d} (z - z_i),$$

also eine Faktorisierung in Linearfaktoren mit Nullstellen $z_1, \ldots, z_d \in \mathbb{C}$. Es gibt viele elegante Beweise für den Fundamentalsatz — aber sie gehören nicht in die lineare Algebra, sondern eher in die Analysis oder Topologie.

6 Polynome

In der linearen Algebra werden wir später nicht nur mit Polynomen über $\mathbb{C}$, sondern auch über anderen Körpern zu tun haben:

Definition 6.1. Sei K ein Körper. Ein *Polynom* in einer Variable x über K ist ein Ausdruck der Form

$$P \;=\; a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \quad \text{mit } n \in \mathbb{N}_0,\, a_0, \ldots, a_n \in K.$$

Formal ist ein Polynom nichts anderes als die Folge seiner Koeffizienten; wenn wir die Zahl n nicht explizit nennen wollen, fassen wir die Koeffizienten auf als eine unendliche Folge $a_0, a_1, a_2, \ldots$ in K mit $a_i = 0$ für alle bis auf endlich viele i und schreiben dann

$$P \;=\; \sum_{i \geq 0} a_i x^i.$$

Im Fall von Polynomen ist aber auch mit dieser saloppen Notation immer gemeint, dass auf der rechten Seite nur endlich viele Summanden stehen — anders als bei Potenzreihen in der Analysis! Wir bezeichnen die Menge aller Polynome über K mit

$$K[x] \;=\; \left\{ \sum_{i \geq 0} a_i x^i \;\middle|\; a_0, a_1, \cdots \in K \text{ mit } a_i = 0 \text{ für alle bis auf endlich viele } i \right\}$$

Definition 6.2. Für Polynome $P = \sum_{i\geq 0} a_i x^i \in K[x]$ und $Q = \sum_{j\geq 0} b_j x^j \in K[x]$ setzen wir

$$\begin{aligned} P+Q &:= \sum_{k\geq 0} c_k x^k \qquad \text{mit} \quad c_k := a_k + b_k, \\ P\cdot Q &:= \sum_{k\geq 0} d_k x^k \qquad \text{mit} \quad d_k := \sum_{i=0}^{k} a_i b_{k-i}. \end{aligned}$$

Man sieht leicht, dass $K[x]$ mit dieser Addition und Multiplikation ein kommutativer Ring wird. Sein Null- und Einselement sind die Polynome

$$\begin{aligned} 0 &:= 0 + 0\cdot x + 0\cdot x^2 + \cdots \quad \text{(das sog. 'Nullpolynom')}, \\ 1 &:= 1 + 0\cdot x + 0\cdot x^2 + \cdots \quad \text{(das sog. 'Einspolynom')}. \end{aligned}$$

Beispielsweise gilt

$$(a_1 x + a_0)\cdot(b_1 x + b_0) \;=\; a_1 b_1 \cdot x^2 + (a_1 b_0 + a_0 b_1)\cdot x + a_0 b_0.$$

Wir können für die formale Variable x auch konkrete Werte einsetzen:

Definition 6.3. Der *Wert* eines Polynoms $P = \sum_{i=0}^{n} a_i x^i \in K[x]$ an einer Stelle $x_0 \in K$ ist definiert durch

$$P(x_0) \;:=\; \sum_{i=0}^{n} a_i x_0^i \;\in\; K.$$

Indem wir die Stelle x_0 variieren, erhalten wir eine Abbildung $K \to K, x_0 \mapsto P(x_0)$, wir nennen diese die *Polynomfunktion* zu dem gegebenen Polynom.

Über endlichen Körpern sollte man Polynome sorgfältig von Polynomfunktionen unterscheiden. Beispielsweise betrachte man über dem Körper $K = \mathbb{F}_2$ mit zwei Elementen das Polynom $P = x^2 + x \in \mathbb{F}_2[x]$: Dieses ist nicht das Nullpolynom, aber die zugehörige Polynomfunktion ist identisch Null.

Definition 6.4. Der *Grad* von $P = \sum_{i\geq 0} a_i x^i \in K[x]$ ist definiert als

$$\deg(P) \;:=\; \begin{cases} -\infty & \text{falls } P \text{ das Nullpolynom ist,} \\ \max\{i \in \mathbb{N}_0 \mid a_i \neq 0\} & \text{sonst.} \end{cases}$$

Lemma 6.5. *Für alle $P, Q \in K[x]$ ist* $\deg(P\cdot Q) = \deg(P) + \deg(Q)$.

Beweis. Falls P oder Q das Nullpolynom ist, gilt die Behauptung mit der üblichen Konvention, dass $-\infty + n := -\infty$ für alle n ist. Sei nun $m = \deg(P), n = \deg(Q) \in \mathbb{N}_0$, also

$$\begin{aligned} P &= a_m x^m + \cdots + a_1 x + a_0 && \text{mit } a_m \neq 0, \\ Q &= b_n x^n + \cdots + b_1 x + b_0 && \text{mit } b_n \neq 0. \end{aligned}$$

Es folgt

$$P \cdot Q \;=\; \sum_{k=0}^{m+n} c_k x^k \quad \text{mit} \quad c_k \;:=\; \sum_{i \geq 0} a_i b_{k-i}.$$

Dabei ist $c_{m+n} = a_m b_n \neq 0$, also $\deg(P \cdot Q) = m+n = \deg(P) + \deg(Q)$. □

Korollar 6.6. *Für jeden Körper K ist der Polynomring $K[x]$ ein Integritätsring.*

Beweis. Für $P, Q \in R[x] \setminus \{0\}$ ist $\deg(P), \deg(Q) \geq 0$. Nach dem vorigen Lemma folgt $\deg(P \cdot Q) = \deg(P) + \deg(Q) \geq 0$. Also ist insbesondere $P \cdot Q \neq 0$. □

Die obigen Resultate gelten allgemeiner auch für Polynome mit Koeffizienten in einem Integritätsring. Mit Polynomen über Körpern kann man aber besonders gut rechnen, z.B. kann man hier ähnlich wie in $\mathbb{Z}$ mit Rest dividieren:

Satz 6.7. *Sei K ein Körper. Dann gibt es für alle $F, G \in K[x]$ mit $G \neq 0$ eindeutige Polynome $Q, R \in K[x]$ mit*

$$F \;=\; G \cdot Q + R \quad \textit{und} \quad \deg(R) \;<\; \deg(G).$$

Beweis. Wir zeigen zuerst die *Eindeutigkeit* der Division mit Rest: Gegeben seien zwei Darstellungen

$$\begin{aligned} F \;&=\; G \cdot Q_1 + R_1 \qquad \text{mit} \quad \deg(R_1) < \deg(G) \\ &=\; G \cdot Q_2 + R_2 \qquad \text{mit} \quad \deg(R_2) < \deg(G) \end{aligned}$$

Dann ist $G \cdot (Q_1 - Q_2) = R_2 - R_1$, also

$$\begin{aligned} \deg(G) + \deg(Q_1 - Q_2) \;=\; \deg(G \cdot (Q_1 - Q_2)) \;&=\; \deg(R_2 - R_1) \\ &\leq\; \max\{\deg(R_1), \deg(R_2)\} \\ &<\; \deg(G) \end{aligned}$$

Also ist $\deg(Q_1 - Q_2) < 0$, d.h. $Q_1 = Q_2$. Es folgt $R_1 = F - G \cdot Q_1 = F - G \cdot Q_2 = R_2$.

Für die *Existenz* fixieren wir ein Polynom $G = \sum_{j=0}^{n} b_j x^j \in K[x]$ mit $b_n \neq 0$ und zeigen per Induktion über $m \in \mathbb{N}$ die Aussage:

$$A(m): \quad \begin{cases} \text{für alle } F \in K[x] \text{ mit } \deg(F) \leq m \text{ gibt es } Q, R \in K[x], \\ \text{sodass gilt: } F = G \cdot Q + R \text{ und } \deg(R) < n = \deg(G). \end{cases}$$

Für alle $m < n$ ist die Aussage $A(m)$ erfüllt, man kann dazu einfach $R = F$ und $Q = 0$ wählen. Im Folgenden sei daher $m \geq n$, und wir machen die Induktionsannahme, dass $A(m-1)$ bereits bewiesen sei. Wir wollen hieraus $A(m)$ folgern: Sei dazu ein Polynom $F = \sum_{i=0}^{m} a_i x^i$ mit $a_m \neq 0$ gegeben. Wir definieren $\tilde{F} \in K[x]$ mit $\deg(\tilde{F}) < m$ durch

$$\tilde{F} \;:=\; F - \tfrac{a_m}{b_n} \cdot G \cdot x^{m-n}$$

Wegen der Annahme $A(m-1)$ existieren Polynome $\tilde{Q}, R \in K[x]$ mit $\tilde{F} = \tilde{Q} \cdot G + R$ und $\deg(R) < n$. Es folgt

$$F \;=\; \tilde{F} + \tfrac{a_m}{b_n} \cdot G \cdot x^{m-n} \;=\; \tilde{Q} \cdot G + R + \tfrac{a_m}{b_n} \cdot G \cdot x^{m-n} \;=\; \left(\tilde{Q} + \tfrac{a_m}{b_n} \cdot x^{m-n}\right) \cdot G + R$$

Dabei gilt noch immer $\deg(R) < \deg(G)$. Damit folgt die Aussage $A(m)$. □

Der obige Beweis per Induktion ist konstruktiv und führt auf das Verfahren der Polynomdivision, das hier nur mit einem Beispiel illustriert sei:

Beispiel 6.8. Für die Polynome

$$F = x^3 + 3x^2 + 2x + 1, \;\; G = x^2 + 2x + 3 \;\in\; \mathbb{Q}[x]$$

berechnet man durch Polynomdivision:

$$\begin{array}{rl} x^3 + 3x^2 + 2x + 1 = \big(x^2 + 2x + 3\big)\big(x+1\big) - 3x - 2 \\ \underline{-x^3 - 2x^2 - 3x} \\ \qquad\quad x^2 \;\; -x + 1 \\ \qquad\quad \underline{-x^2 - 2x - 3} \\ \qquad\qquad\quad -3x - 2 \end{array}$$

Wir erhalten hier also $Q = x + 1$ und $R = -3x - 2$.

Ringe, in denen wie im obigen Satz eine Division mit Rest möglich ist, verdienen einen eigenen Namen:

Definition 6.9. Ein *Euklidischer Ring* ist ein Integritätsring R mit folgender weiterer Eigenschaft: Es gibt eine Funktion

$$\delta: \quad R \setminus \{0\} \;\longrightarrow\; \mathbb{N}_0,$$

sodass für alle $a \in R$, $b \in R \setminus \{0\}$ Elemente $q, r \in R$ existieren mit

a) $a = qb + r$, und

b) $\delta(r) < \delta(b)$ im Fall $r \neq 0$.

Wir nennen dann δ auch eine *Gradfunktion* für R.

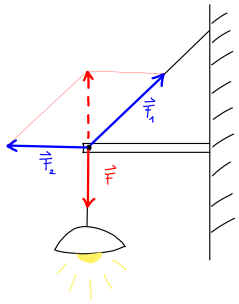
Kapitel II
Vektorräume

Zusammenfassung Lineare Strukturen spielen eine zentrale Rolle in allen Teilen der Mathematik und ihren Anwendungen. In diesem Kapitel werden wir den für diese zentralen Begriff eines Vektorraumes über einem Körper einführen. Elemente eines Vektorraumes kann man durch Tupel von Körperelementen angeben, wenn man eine Basis, d.h. ein linear unabhängiges Erzeugendensystem wählt. Die Anzahl der Vektoren in einer Basis hängt nicht von der Basis ab und heißt die Dimension des Vektorraumes. In Verallgemeinerung der Konstruktion von Vektorräumen aus Tupeln werden wir schließlich die direkte Summe von Vektorräumen betrachten.

1 Definition und Beispiele

Den Begriff eines Vektors kennen viele aus der Physik:

Beispiel 1.1. Das Kabel einer Hängelampe werde eine senkrecht aus einer Mauer ragende Schiene geführt wie in der folgenden Abbildung. Welche Zugkraft muß das Kabel aushalten?



Sei F die Gewichtskraft der Lampe. Dann ergeben sich die Zugkraft F_1 am Seil und die auf die Schiene wirkende Druckkraft F_2 aus dem in der obigen Abbildung

skizzierten *Kräfteparallelogramm*. Physiker schreiben

$$\vec{F}_1 + \vec{F}_2 \;=\; -\vec{F}$$

und sagen, dass Kräfte *Vektoren* seien, also neben einem Betrag auch eine Richtung besitzen. Dabei gilt:

- Vektoren kann man addieren: Die Summe von zwei Kräften ist gegeben durch ein Kräfteparallelogramm wie in der obigen Skizze.
- Vektoren kann man mit einer reellen Zahl multiplizieren, dabei wird ihr Betrag reskaliert und die Richtung dreht sich im Fall negativer Vorzeichen um.

Durch Kombination beider Operationen kann man beliebige *Linearkombinationen* von Vektoren bilden. Solche Linearkombinationen treten in zahlreichen weiteren Situationen auf:

Beispiel 1.2. In *Interpolationsproblemen* sucht man ein Polynom, das an gegebenen Stellen vorgegebene Werte annimmt. Gesucht sei z.B. für $a,b,c \in \mathbb{R}$ ein $f \in \mathbb{R}[x]$ mit der Eigenschaft

$$f(1) \;=\; a, \quad f(2) \;=\; b, \quad f(3) \;=\; c.$$

Eine elegante Lösung dieses Interpolationsproblems erhält man durch Betrachten der drei Polynome

$$f_1 \;=\; \tfrac{1}{2}(x-2)(x-3), \quad f_2 \;=\; -(x-1)(x-3), \quad f_3 \;=\; \tfrac{1}{2}(x-1)(x-2).$$

Per Konstruktion gilt

$$f_i(j) \;=\; \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j, \end{cases}$$

somit erhalten wir eine Lösung des Interpolationsproblems als Linearkombination der Form

$$f \;=\; af_1 + bf_2 + cf_3 \;\in\; \mathbb{R}[x].$$

Das ist optimal: Man sieht leicht, dass dies die einzige Lösung mit $\deg(f) \leq 2$ ist.

Aus mathematischer Sicht machen wir in beiden Beispielen dasselbe: Sowohl Kräfte als auch Polynome kann man addieren und mit sogenannten *Skalaren*, d.h. Elementen aus einem Körper, multiplizieren. Es gibt eine nahezu unendliche Zahl weiterer Beispiele. Dabei kann sich die Natur der betrachteten Objekte von Fall zu Fall sehr unterscheiden und wir wollen auch über anderen Zahlbereichen wie $\mathbb{C}$, $\mathbb{Q}$ oder endlichen Körpern rechnen (etwa in der Codierungstheorie). Daher fassen wir die in allen Fällen benötigte Struktur wie folgt zusammen:

Definition 1.3. Ein *Vektorraum über einem Körper K* oder kurz ein *K-Vektorraum* ist ein Tupel $(V,+,\cdot)$ bestehend aus einer abelschen Gruppe $(V,+)$ und aus einer Verknüpfung

$$\cdot : \quad K \times V \;\to\; V, \quad (\alpha, v) \;\mapsto\; \alpha \cdot v,$$

der sogenannten *Skalarmultiplikation*, sodass für alle $\alpha,\beta \in K$, $v,w \in V$ gilt:

a) Assoziativität: $\alpha \cdot (\beta \cdot v) = (\alpha\beta) \cdot v$.

b) Distributivität: $(\alpha+\beta) \cdot v = \alpha \cdot v + \beta \cdot v$ und $\alpha \cdot (v+w) = \alpha \cdot v + \alpha \cdot w$.

c) Kompatibilität mit dem Einselement: $1 \cdot v = v$.

Die Elemente eines Vektorraumes bezeichnen wir als *Vektoren*.

Beispiel 1.4. Jeder K-Vektorraum V enthält mindestens ein Element, das neutrale Element der Gruppe $(V,+)$. Dieses Element heißt der *Nullvektor* $0 \in V$. Umgekehrt bildet die triviale additive Gruppe $V = \{0\}$ in trivialer Weise einen Vektorraum über jedem Körper. Man nennt diesen Vektorraum $V = \{0\}$ den *Nullraum*.

Beispiel 1.5. Der *Standard-Vektorraum* $V = K^n$ für $n \in \mathbb{N}$ ist die Menge der n-Tupel von Elementen aus K. Wir schreiben derartige Tupel manchmal platzsparend wie in der Mengenlehre als Zeilenvektoren $(a_1, \ldots, a_n)$. Meist werden wir aber ab jetzt Spaltenvektoren

$$v = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \quad \text{mit} \quad a_1, \ldots, a_n \in K$$

nutzen. Die Addition und Skalarmultiplikation sind komponentenweise definiert durch

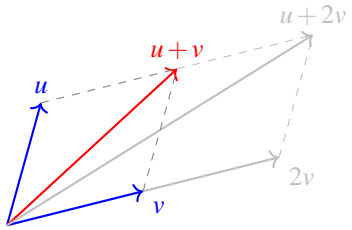
$$\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := \begin{pmatrix} u_1+v_1 \\ \vdots \\ u_n+v_n \end{pmatrix} \quad \text{und} \quad \alpha \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} := \begin{pmatrix} \alpha \cdot w_1 \\ \vdots \\ \alpha \cdot w_n \end{pmatrix}$$

Damit wird $V = K^n$ ein K-Vektorraum. Sein Nullvektor ist der Spaltenvektor, dessen Einträge alle Null sind, und additive Inverse von Vektoren sind gegeben durch

$$-\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} := \begin{pmatrix} -a_1 \\ \vdots \\ -a_n \end{pmatrix} \quad \text{für } a_1, \ldots, a_n \in K.$$

Dass die Skalarmultiplikation assoziativ, distributiv und mit $1 \in K$ kompatibel ist, folgt komponentenweise aus den entsprechenden Eigenschaften von K.

Anschaulich kann man den $\mathbb{R}$-Vektorraum $\mathbb{R}^3$ betrachten als Modell für den uns umgebenden Raum, die soeben definierte komponentenweise Summe von Vektoren entspricht dann geometrisch der Diagonalen in einem Parallelogramm:



Beispiel 1.6. Nicht in jedem Vektorraum lassen sich Vektoren durch endliche Tupel von Körperelementen beschreiben, es gibt auch sehr viel größere Vektorräume:

- Die Menge aller Folgen $(a_1, a_2, \dots)$ von Elementen $a_i \in K$ ist ein Vektorraum über K mit der gliedweisen Addition und Skalarmultiplikation

$$\begin{aligned}(a_1, a_2, \dots) + (b_1, b_2, \dots) &:= (a_1 + b_1, a_2 + b_2, \dots),\\ \alpha \cdot (a_1, a_2, \dots) &:= (\alpha a_1, \alpha a_2, \dots).\end{aligned}$$

- Die Menge $V = \mathrm{Abb}(\mathbb{R}, \mathbb{R}) = \big\{\text{Abbildungen } f : \mathbb{R} \to \mathbb{R}\big\}$ ist ein Vektorraum über den reellen Zahlen mit der *punktweisen Vektorraumstruktur*, die definiert ist durch

$$\begin{aligned}(f+g)(x) &:= f(x) + g(x) && \text{für } f, g \in V\\ (\alpha \cdot f)(x) &:= \alpha \cdot f(x) && \text{und } \alpha \in \mathbb{R}.\end{aligned}$$

  Dasselbe gilt für

$$\begin{aligned}V &= \big\{\text{stetige Abbildungen } f : \mathbb{R} \to \mathbb{R}\big\},\\ V &= \big\{\text{differenzierbare Abbildungen } f : \mathbb{R} \to \mathbb{R}\big\}, \text{ usw.}\end{aligned}$$

Für die Addition und Multiplikation mit Skalaren gelten in Vektorräumen die üblichen Rechenregeln, wobei wir der Klarheit halber ausnahmsweise mit $0_K \in K$ das Nullelement des Körpers und mit $0_V \in V$ das Nullelement des Vektorraumes bezeichnen:

Lemma 1.7. *Sei V ein Vektorraum über K. Für alle $\alpha \in K$ und $v \in V$ gelten dann die Identitäten*

$$0_K \cdot v \;=\; \alpha \cdot 0_V \;=\; 0_V \quad \textit{und} \quad (-\alpha) \cdot v \;=\; \alpha \cdot (-v) \;=\; -(\alpha \cdot v).$$

Es gilt außerdem die Kürzungsregel*: Aus $\alpha \cdot v = 0_V$ folgt $\alpha = 0_K$ oder $v = 0_V$.*

Beweis. Die ersten beiden Rechenregeln beweist man genauso wie die analogen Rechenregeln in Ringen. Wir zeigen daher nur die Kürzungsregel: Für $\alpha \cdot v = 0$ mit $\alpha \in K \setminus \{0_K\}$ ist $v = 1_K \cdot v = (\alpha^{-1} \cdot \alpha) \cdot v = \alpha^{-1} \cdot (\alpha \cdot v) = \alpha^{-1} \cdot 0_V = 0_V$ nach den Axiomen für Vektorräume. □

Die Kürzungsregel ist übrigens der Grund, warum wir die lineare Algebra hier über Körpern und nicht über allgemeineren Integritätsringen entwickeln: Sie macht das Leben deutlich einfacher und wird im Folgenden immer wieder verwendet!

2 Untervektorräume

Häufig interessiert man sich für Teilmengen eines Vektorraumes, die stabil unter der Addition und Skalarmultiplikation sind:

Definition 2.1. Sei V ein Vektorraum über K. Wir bezeichnen eine Teilmenge $U \subseteq V$ als einen *K-Untervektorraum* oder auch kurz als einen *Unterraum* oder *Teilraum* von V, wenn gilt:

a) Es ist $U \neq \varnothing$.

b) Für alle $u_1, u_2 \in U$ ist auch $u_1 + u_2 \in U$.

c) Für alle $\alpha \in K$, $u \in U$ ist auch $\alpha \cdot u \in U$.

Lemma 2.2. *Sei V ein Vektorraum über K. Dann ist auch jeder Unterraum $U \subseteq V$ ein Vektorraum über K mit der Addition und Skalarmultiplikation*

$$+ : U \times U \longrightarrow U \quad \textit{und} \quad \cdot : K \times U \longrightarrow U$$

welche die Einschränkung der Addition und Skalarmultiplikation des Vektorraums V auf die Teilmengen $U \times U \subseteq V \times V$ bzw. $K \times U \subseteq K \times V$ sind.

Beweis. Folgt direkt aus den Definitionen. Die Existenz additiver Inverse folgt dabei aus der Eigenschaft $c)$ in der obigen Definition mit $\alpha = -1$, denn $-u = (-1) \cdot u$. □

Beispiel 2.3. Jeder K-Vektorraum V enthält die sogenannten *trivialen Unterräume*, nämlich den Nullraum $U = \{0\}$ und den ganzen Vektorraum $U = V$. Für jeden Vektor $v \in V \setminus \{0\}$ ist ferner

$$\langle v \rangle_K \; := \; \{\alpha v \in V \mid \alpha \in K\} \; \subseteq \; V$$

ein Untervektorraum, wir nennen ihn die durch v aufgespannte *Gerade*.

Beispiel 2.4. Seien $a_1, a_2, a_3 \in K$. Im Vektorraum $V = K^3$ ist dann

$$U \; := \; \left\{ (v_1, v_2, v_3) \in K^3 \;\middle|\; a_1 v_1 + a_2 v_2 + a_3 v_3 = 0 \right\}$$

ein Untervektorraum:

a) Es ist $(0,0,0) \in U$, also $U \neq \varnothing$.

b) Für $u = (u_1, u_2, u_3), v = (v_1, v_2, v_3) \in U$ ist $u + v \in U$ wegen

$$\sum_{i=1}^{3} a_i (u_i + v_i) \; = \; \sum_{i=1}^{3} a_i u_i + \sum_{i=1}^{3} a_i v_i \; = \; 0 + 0 \; = \; 0.$$

c) Ebenso prüft man direkt $\alpha \cdot u \in U$ für alle $u \in U$ und $\alpha \in K$ nach.

Wenn die Koeffizienten a_1, a_2, a_3 nicht alle drei Null sind, kann man sich $U \subset K^3$ anschaulich als eine *Ebene* vorstellen. Für $a_1 \neq 0$ besteht diese beispielsweise genau aus den Vektoren

$$v \; = \; \alpha \cdot \begin{pmatrix} -a_2 \\ a_1 \\ 0 \end{pmatrix} \; + \; \beta \cdot \begin{pmatrix} -a_3 \\ 0 \\ a_1 \end{pmatrix} \quad \text{mit} \quad \alpha, \beta \in K.$$

Beispiel 2.5. Die Menge

$$\mathscr{D}(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} \text{ differenzierbar}\} \subseteq \text{Abb}(\mathbb{R}, \mathbb{R})$$

aller differenzierbaren Funktionen bildet einen Untervektorraum im $\mathbb{R}$-Vektorraum aller reeller Funktionen aus Beispiel 1.6:

- $\mathscr{D}(\mathbb{R}) \neq \varnothing$, da die Nullfunktion differenzierbar ist.
- Sind $f, g : \mathbb{R} \to \mathbb{R}$ differenzierbar, dann auch $f + g$.
- Ist $f : \mathbb{R} \to \mathbb{R}$ differenzierbar, dann auch αf für $\alpha \in \mathbb{R}$.

Beispiel 2.6. Die Teilmenge

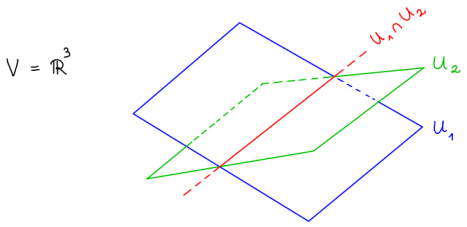
$$U = \{f \in \mathscr{D}(\mathbb{R}) \mid f'(x) = f(x) \text{ für alle } x \in \mathbb{R}\} \subseteq \mathscr{D}(\mathbb{R})$$

ist ein $\mathbb{R}$-Untervektorraum, denn:

- Es ist $U \neq \varnothing$, da die Nullfunktion in U liegt.
- Für $f, g \in U$ ist auch $f + g \in U$, denn $(f+g)' = f' + g' = f + g$.
- Für $\alpha \in \mathbb{R}$ und $f \in U$ ist auch $\alpha \cdot f \in U$, denn $(\alpha \cdot f)' = \alpha \cdot (f') = \alpha \cdot f$.

In der Analysis werden Sie beweisen, dass U genau aus den reellen Vielfachen der Exponentialfunktion besteht: Man könnte ihn als eine "Gerade"bezeichnen.

Im $\mathbb{R}^3$ schneiden sich je zwei *verschiedene* Ebenen durch den Ursprung entlang einer Gerade:



Allgemein ist der Durchschnitt von Unterräumen wieder ein Untervektorraum:

Lemma 2.7. *Sei V ein Vektorraum über einem Körper K. Sei $(U_i)_{i \in I}$ eine Kollektion von Untervektorräumen $U_i \subseteq V$, wobei I eine beliebige Indexmenge sei. Dann ist auch der Durchschnitt*

$$U := \bigcap_{i \in I} U_i \subseteq V \quad \textit{ein Untervektorraum.}$$

Beweis. Per Definition von Unterräumen ist $0 \in U_i$ für alle i. Also ist $0 \in U$ und somit $U \neq \varnothing$. Seien jetzt $u, v \in U$ und $\alpha \in K$. Dann ist $u, v \in U_i$ für alle $i \in I$. Somit folgt $u + v$, $\alpha v \in U_i$ für alle $i \in I$, also $u + v$, $\alpha v \in U$ wie gewünscht. □

Für jeden von Null verschiedenen Vektor eines K-Vektorraumes haben wir die von diesem aufgespannte Gerade betrachtet. Allgemeiner kann man für nichtleere Teilmengen $A \subseteq V$ die Teilmenge

$$\langle A \rangle_K := \left\{ \sum_{i=1}^{n} \alpha_i v_i \;\middle|\; n \in \mathbb{N}, \alpha_i \in K, v_i \in A \right\} \subseteq V$$

bilden. Wir bezeichnen diese Teilmenge als den *K-Aufspann* von A, ihre Elemente heißen *K-Linearkombinationen* von Elementen aus A. Den Körper K lassen wir gern weg, wenn er aus dem Kontext klar ist. Für endliche Mengen $A = \{v_1, \ldots, v_n\}$ sparen wir uns in der Notation die Mengenklammern und schreiben kurz

$$\langle v_1, \ldots, v_n \rangle_K := \langle \{v_1, \ldots, v_n\} \rangle_K.$$

Den Aufspann der leeren Menge $A = \varnothing$ definieren wir formal durch $\langle \varnothing \rangle_K := \{0\}$, was zur Vermeidung von Fallunterscheidungen nützlich ist.

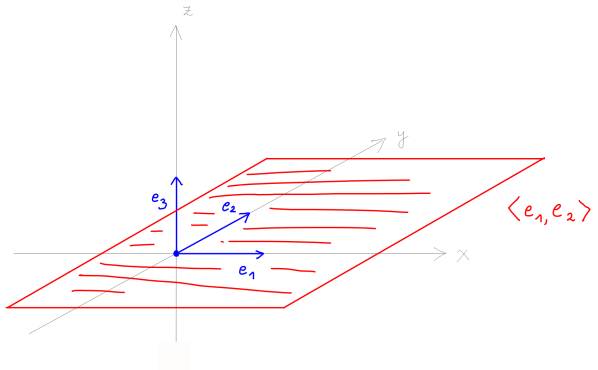
Beispiel 2.8. Im $\mathbb{R}$-Vektorraum $V = \mathbb{R}^3$ seien

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

gegeben. Dann ist

$$\langle e_1, e_2 \rangle_{\mathbb{R}} = \{ \alpha_1 e_1 + \alpha_2 e_2 \mid \alpha_1, \alpha_2 \in \mathbb{R} \} = \left\{ \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ 0 \end{pmatrix} \mid \alpha_1, \alpha_2 \in \mathbb{R} \right\}$$

die Koordinatenebene durch die ersten beiden Koordinatenachsen:



Man bezeichnet den Aufspann einer Teilmenge $A \subseteq V$ eines Vektorraumes V auch als den von A *aufgespannten Untervektorraum*, aus gutem Grund:

Lemma 2.9. *Es sei V ein K-Vektorraum. Für jede Teilmenge $A \subseteq V$ ist dann ihr Aufspann*

$$\langle A \rangle_K \subseteq V \quad \textit{ein Untervektorraum.}$$

Beweis. Für $A = \varnothing$ ist der Aufspann per Definition der Nullraum und dieser bildet einen Untervektorraum, wir dürfen also $A \neq \varnothing$ annehmen. Es ist $v = \alpha \cdot v$ für $\alpha = 1$ und alle $v \in A$. Somit folgt per Definition des Aufspanns $A \subseteq \langle A \rangle_K$. Insbesondere ist der Aufspann also nicht leer. Zu zeigen bleibt die Stabilität unter Addition und Skalarmultiplikation. Seien dazu $v, w \in \langle A \rangle_K$ beliebig vorgegeben. Wir schreiben diese in der Form

$$v = \sum_{i=1}^{m} \alpha_i v_i \quad \text{mit } \alpha_i \in K,\, v_i \in A \quad \text{und} \quad w = \sum_{j=1}^{n} \beta_j w_j \quad \text{mit } \beta_j \in K,\, w_j \in A.$$

Dann ist $v + w = \gamma_1 u_1 + \cdots + \gamma_{m+n} u_{m+n}$ mit

$$\gamma_i = \begin{cases} \alpha_i \\ \beta_{i-m} \end{cases} \quad \text{und} \quad u_i = \begin{cases} v_i & \text{für} \quad \begin{cases} i \leq m, \\ i > m. \end{cases} \\ w_{i-m} \end{cases}$$

Also ist $v + w \in \langle A \rangle_K$. Analog sieht man $\alpha \cdot u \in \langle A \rangle_K$ für alle $\alpha \in K, u \in \langle A \rangle_K$. □

Satz 2.10. *Es sei V ein K-Vektorraum. Der Aufspann einer Teilmenge $A \subseteq V$ ist der kleinste sie enthaltende Untervektorraum, d.h. es gilt:*

a) Es ist $\langle A \rangle_K \subseteq V$ ein Untervektorraum, der A enthält.

b) Jeder andere solche Untervektorraum enthält $\langle A \rangle_K$.

Beweis. Ist $U \subseteq V$ ein beliebiger Untervektorraum mit $A \subseteq U$, dann folgt aus der Abgeschlossenheit von Unterräumen unter der Addition und Skalarmultiplikation, dass

$$\sum_{i=1}^{n} \alpha_i u_i \in U \quad \text{für alle } \alpha_i \in K,\ u_i \in A \subseteq U$$

gilt. Also ist $\langle A \rangle_K \subseteq U$ per Definition des Aufspanns. Umgekehrt ist $\langle A \rangle_K$ nach dem vorigen Lemma selber bereits ein A enthaltender Unterraum. □

Bemerkung 2.11. Um die Aussage des obigen Satzes anschaulich zu machen, nennt man den Aufspann einer Teilmenge $A \subseteq V$ auch die *lineare Hülle* von A.

3 Erzeuger und lineare Unabhängigkeit

Wir wollen uns nun überlegen, wie man die Elemente eines Vektorraumes konkret angeben kann. Wir beginnen dazu mit einem Erzeugendensystem:

Definition 3.1. Ein Vektorraum V über dem Körper K heißt *endlich erzeugt*, wenn er von endlich vielen Vektoren aufgespannt werden kann, d.h. wenn es ein $n \in \mathbb{N}$ und $v_1, \ldots, v_n \in V$ gibt mit

$$V = \langle v_1, \ldots, v_n \rangle_K.$$

Wir nennen dann das Tupel $(v_1,\dots,v_n)$ ein *Erzeugendensystem* von V über K. Man beachte, dass wir hier ein Tupel betrachten, keine Menge: Die Indices $i = 1,\dots,n$ sind für das Hinschreiben von Vektoren als Linearkombinationen

$$v = \sum_{i=1}^{n} \alpha_i v_i \quad \text{mit } \alpha_1,\dots,\alpha_n \in K$$

nützlich. Allgemeiner nennen wir für eine beliebige Indexmenge I ein Tupel $(v_i)_{i\in I}$ von Vektoren $v_i \in V$ ein *Erzeugendensystem* von V über K, wenn gilt:

$$V = \langle v_i \mid i \in I \rangle_K.$$

Dabei muß die Menge I nicht endlich sein, z.B. hat jeder Vektorraum trivialerweise ein Erzeugendensystem, das aus allen Vektoren in V besteht. Aber in der Praxis wollen wir natürlich möglichst kleine Erzeugendensysteme benutzen.

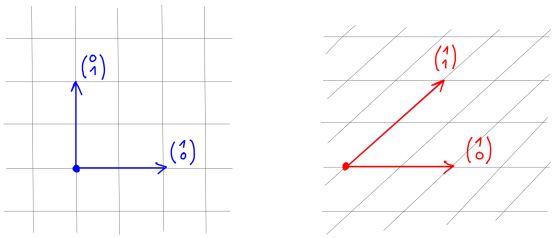
Beispiel 3.2. Für den Standard-Vektorraum $V = K^2$ ist

$$V = \left\langle \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \right\rangle_K = \left\langle \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix} \right\rangle_K$$

wegen

$$\begin{pmatrix}x\\y\end{pmatrix} = x\cdot\begin{pmatrix}1\\0\end{pmatrix} + y\cdot\begin{pmatrix}0\\1\end{pmatrix} = (x-y)\cdot\begin{pmatrix}1\\0\end{pmatrix} + y\cdot\begin{pmatrix}1\\1\end{pmatrix} \quad \text{für alle } x,y \in K.$$

Wir haben somit zwei Erzeugendensysteme von V gefunden. Für $K = \mathbb{R}$ führt die Beschreibung von Vektoren mithilfe dieser Erzeugendensysteme auf die folgenden beiden Koordinatensysteme in der reellen Ebene:



Allgemein gilt: Für jeden Körper K und $n \in \mathbb{N}$ ist der Standard-Vektorraum $V = K^n$ endlich erzeugt über K, mit einem Erzeugendensystem $(e_1,\dots,e_n)$ bestehend aus den sog. *Standard-Basisvektoren*

$$e_i = \begin{pmatrix}0\\\vdots\\0\\1\\0\\\vdots\\0\end{pmatrix} \leftarrow i\text{-te Stelle} \qquad (1 \le i \le n)$$

Beispiel 3.3. Der K-Vektorraum $V = K[t]$ ist nicht endlich erzeugt.

Beweis. Wäre $V = K[t]$ endlich erzeugt über K, so gäbe es Polynome $p_1, \dots, p_n \in V$ mit der Eigenschaft

$$V \;=\; \langle p_1, \dots, p_n \rangle_K.$$

Nach Definition des Aufspanns könnten wir dann *jedes* Polynom $q \in V = K[t]$ schreiben als $q = \alpha_1 p_1 + \cdots + \alpha_n p_n$ mit $\alpha_1, \dots, \alpha_n \in K$. Der Grad *jedes* Polynoms würde dann

$$\deg(q) \;\leq\; d \;:=\; \max\{\deg(p_i) \mid i = 1, \dots, n\}$$

erfüllen, was z.B. für das Polynom $q := t^{d+1}$ einen Widerspruch liefert. □

Wir interessieren uns im Folgenden vor allem für endlich erzeugte Vektorräume und möchten diese möglichst effizient beschreiben: Wie kann man ein möglichst kleines Erzeugendensystem finden? Wenn wir dazu mit einem beliebigen endlichen Erzeugendensystem beginnen und versuchen, daraus Vektoren zu entfernen, stellt sich die Frage: Wann kann man aus einem gegebenen System von Vektoren einen Vektor weglassen, ohne dabei den Aufspann des Systems zu verändern? Im Satz 3.8 werden wir sehen, dass dies eng mit dem folgenden Begriff zusammenhängt:

Definition 3.4. Es sei V ein Vektorraum über K. Vektoren $v_1, \dots, v_n \in V$ heißen

a) *linear abhängig* über K, wenn es $\alpha_1, \dots, \alpha_n \in K$ gibt, die nicht alle Null sind, sodass gilt:

$$\alpha_1 v_1 + \cdots + \alpha_n v_n \;=\; 0.$$

b) *linear unabhängig* über K, wenn sie nicht linear abhängig sind, d.h. wenn die folgende Implikation gilt:

$$\Big(\alpha_1 v_1 + \cdots + \alpha_n v_n \;=\; 0 \quad \text{mit} \quad \alpha_1, \dots, \alpha_n \in K \Big) \quad \Longrightarrow \quad \alpha_1 = \cdots = \alpha_n = 0.$$

Beispiel 3.5. Es gilt:

a) Das aus einem einzigen Vektor $v \in V$ bestehende System ist linear unabhängig genau dann, wenn $v \neq 0$ ist: Denn genau dann folgt aus $\alpha \cdot v = 0$ bereits $\alpha = 0$.

b) Jedes System von Vektoren $v_1, \dots, v_n \in V$ mit $v_{i_0} = 0$ für ein i_0 ist linear abhängig, denn

$$\sum_{i=1}^{n} \alpha_i v_i \;=\; 0 \quad \text{mit} \quad \alpha_i \;=\; \begin{cases} 1 & \text{für } i = i_0, \\ 0 & \text{für } i \neq i_0. \end{cases}$$

c) Analog sieht man, dass auch jedes System von Vektoren $v_1, \dots, v_n \in V$, das einen Vektor mehrfach enthält, linear abhängig sein muß: Denn ist $v_{i_0} = v_{j_0}$ für zwei Indizes $i_0 \neq j_0$, so folgt

$$\sum_{i=1}^{n} \alpha_i v_i \;=\; 0 \quad \text{mit} \quad \alpha_i \;=\; \begin{cases} +1 & \text{für } i = i_0, \\ -1 & \text{für } i = j_0, \\ 0 & \text{sonst.} \end{cases}$$

d) In $V = K^n$ bilden die Standard-Basisvektoren ein linear unabhängiges System, denn eine Linearkombination

$$\alpha_1 e_1 + \cdots + \alpha_n e_n \;=\; \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

ist der Nullvektor genau dann, wenn alle Einträge des Vektors auf der rechten Seite verschwinden, also genau für $\alpha_1 = \cdots = \alpha_n = 0$.

e) In $V = K^2$ sind die drei Vektoren

$$v_0 \;=\; \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_1 \;=\; \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_2 \;=\; \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

linear abhängig, denn

$$v_0 - 2v_1 + v_2 \;=\; \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} \;=\; \begin{pmatrix} 1-2+1 \\ 0-2+2 \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Der obige Begriff der linearen Unabhängigkeit lässt sich wie folgt auf Familien von unendlich vielen Vektoren ausdehnen:

Definition 3.6. Eine Familie $(v_i)_{i \in I}$ von Vektoren $v_i \in V$ heißt *linear unabhängig*, wenn jede endliche Teilfamilie linear unabhängig ist, d.h. wenn für jede endliche Teilmenge $I_0 \subseteq I$ gilt:

$$\left(\sum_{i \in I_0} \alpha_i v_i \;=\; 0 \quad \text{mit} \quad \alpha_i \in K \right) \quad \Longrightarrow \quad \forall i \;\in\; I_0 \colon \quad \alpha_i \;=\; 0.$$

Beispiel 3.7. In $V = K[t]$ ist die Familie der Monome $v_i = t^i$ linear unabhängig, denn ein Polynom ist das Nullpolynom genau dann, wenn alle seine Koeffizienten verschwinden:

$$\sum_{i=0}^{n} \alpha_i t^i \;=\; 0 \text{ in } K[t] \quad \Longleftrightarrow \quad \alpha_0 = \cdots = \alpha_n = 0.$$

Satz 3.8. *Für Familien $(v_i)_{i \in I}$ von Vektoren in einem K-Vektorraum V sind folgende Aussagen zueinander äquivalent:*

a) Es ist $(v_i)_{i \in I}$ linear unabhängig über K.

b) Es ist $v_{i_0} \notin \langle v_i \mid i \in I \setminus \{i_0\} \rangle_K$ für alle $i_0 \in I$.

c) Es ist $\langle v_i \mid i \in I \setminus \{i_0\} \rangle_K \neq \langle v_i \mid i \in I \rangle_K$ für alle $i_0 \in I$.

d) Linearkombinationen sind eindeutig, d.h. für jede endliche Teilmenge $I_0 \subseteq I$ gilt:

$$\left(\sum_{i \in I_0} \alpha_i v_i \;=\; \sum_{i \in I_0} \beta_i v_i \quad \textit{mit} \quad \alpha_i, \beta_i \;\in\; K \right) \qquad \Longrightarrow \qquad \forall i \in I_0 \colon \;\; \alpha_i \;=\; \beta_i$$

Beweis. Wir zeigen $\neg a \Longrightarrow \neg b \Longrightarrow \neg c \Longrightarrow \neg d \Longrightarrow \neg a$:

Wenn $a)$ nicht gilt, dann gibt es per Definition eine endliche Teilmenge $I_0 \subseteq I$ und $\alpha_i \in K$ mit

$$\sum_{i \in I_0} \alpha_i v_i \;=\; 0 \quad \text{und} \quad \alpha_{i_0} \neq 0 \quad \text{für ein } i_0 \in I_0.$$

Dann ist

$$v_{i_0} \;=\; -\sum_{i \in I_0 \setminus \{i_0\}} \frac{\alpha_i}{\alpha_{i_0}} \cdot v_i \;\in\; \big\langle v_i \mid i \in I \setminus \{i_0\} \big\rangle_K$$

und somit gilt $b)$ nicht. Angenommen, es gelte nun $b)$ nicht. Dann gibt es ein $i_0 \in I$ mit der Eigenschaft

$$v_{i_0} \;\in\; H \;:=\; \langle v_i \mid i \in I \setminus \{i_0\} \rangle_K.$$

Da jede Teilmenge eines Vektorraumes in ihrem Aufspann enthalten ist, gilt aber auch $v_i \in H$ für alle $i \in I \setminus \{i_0\}$. Somit folgt $v_i \in H$ für ausnahmslos alle $i \in I$. Dann ist aber

$$\langle v_i \mid i \in I \rangle_K \;\subseteq\; H.$$

Die umgekehrte Inklusion ist trivial per Definition von H als Aufspann der v_i, also folgt $\langle v_i \mid i \in I \rangle = H$ und damit gilt $c)$ nicht. Angenommen, es gelte nun $c)$ nicht, sei also

$$\langle v_i \mid i \in I \setminus \{i_0\} \rangle_K = \langle v_i \mid i \in I \rangle_K$$

für ein $i_0 \in I$. Dann ist insbesondere $v_{i_0} \in \langle v_i \mid i \in I \setminus \{i_0\} \rangle$. Also existieren $\alpha_i \in K$ mit

$$v_{i_0} \;=\; \sum_{i \in I \setminus \{i_0\}} \alpha_i v_i \quad \text{für eine endliche Teilmenge } I_0 \subseteq I \text{ mit } i_0 \in I_0.$$

Andererseits gilt

$$v_{i_0} \;=\; \sum_{i \in I_0} \beta_i v_i \quad \text{mit} \quad \beta_i \;=\; \begin{cases} 1 & \text{für } i = i_0, \\ 0 & \text{sonst.} \end{cases}$$

Somit gilt $d)$ nicht, denn $\beta_{i_0} = 1 \neq 0 = \alpha_{i_0}$. Angenommen, es gelte nun $d)$ nicht, es sei also

$$\sum_{i \in I_0} \alpha_i v_i \;=\; \sum_{i \in I_0} \beta_i v_i$$

für eine endliche Teilmenge $I_0 \subseteq I$ und $\alpha_i, \beta_i \in K$, wobei $\alpha_{i_0} \neq \beta_{i_0}$ für mindestens ein $i_0 \in I_0$ gelte. Dann folgt

$$\sum_{i \in I_0} \gamma_i v_i \;=\; 0 \quad \text{mit} \quad \gamma_i \;=\; \alpha_i - \beta_i.$$

Dabei ist $\gamma_{i_0} \neq 0$ und somit ist $(v_i)_{i \in I}$ linear abhängig. Also gilt $a)$ nicht. □

4 Basen von Vektorräumen

Erzeugendensysteme eines Vektorraumes enthalten genügend viele Vektoren, um den ganzen Vektorraum aufzuspannen. Linear unabhängige Familien enthalten nur so wenigen Vektoren, dass man keinen davon weglassen kann, ohne ihren Aufspann zu verkleinern. Jetzt soll es um die goldene Mitte zwischen beidem gehen:

Definition 4.1. Eine *Basis* eines Vektorraums V über K ist ein linear unabhängiges Erzeugendensystem des Vektorraums.

Beispiel 4.2. Im $\mathbb{R}$-Vektorraum $V = \mathbb{R}^2$ betrachte man

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

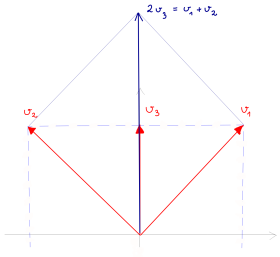
a) (v_1) ist linear unabhängig, aber kein Erzeugendensystem von V über $\mathbb{R}$:

$$\langle v_1 \rangle_{\mathbb{R}} \;=\; \left\{ \begin{pmatrix} x \\ x \end{pmatrix} \mid x \in \mathbb{R} \right\} \;\neq\; V.$$

b) (v_1, v_2, v_3) ist ein Erzeugendensystem von V, aber nicht $\mathbb{R}$-linear unabhängig:

$$v_1 + v_2 - 2v_3 \;=\; \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

c) (v_1, v_2) ist eine Basis des $\mathbb{R}$-Vektorraumes $V = \mathbb{R}^2$. Das ist anschaulich klar, wenn man ein um 45° gedrehtes Koordinatensystem in der Ebene betrachtet:



In der Tat ist hier für beliebige $x, y \in \mathbb{R}$ die Gleichung

$$\begin{pmatrix} x \\ y \end{pmatrix} \overset{??}{=} \alpha_1 v_1 + \alpha_2 v_2 \;=\; \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} \;=\; \begin{pmatrix} \alpha_1 - \alpha_2 \\ \alpha_1 + \alpha_2 \end{pmatrix}$$

für eindeutige Koeffizienten $\alpha_1, \alpha_2 \in \mathbb{R}$ erfüllt, nämlich genau für

$$\begin{aligned} \alpha_1 &= (y+x)/2, \\ \alpha_2 &= (y-x)/2. \end{aligned}$$

Also ist (v_1, v_2) eine Basis von $\mathbb{R}^2$ nach dem folgendem Lemma:

Lemma 4.3. *Sei V ein Vektorraum über K. Für $v_1, \dots, v_n \in V$ sind äquivalent:*

a) $(v_1, \dots, v_n)$ *ist eine Basis von V über K.*

b) Für jedes $v \in V$ gibt es eindeutige $\alpha_1, \dots, \alpha_n \in K$ mit $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$.

Beweis. Die Existenz von α_i wie in *b)* besagt per Definition, dass $v_1, \dots, v_n$ ein Erzeugendensystem bilden. Die Eindeutigkeit der α_i ist nach Satz 3.8 äquivalent dazu, dass $v_1, \dots, v_n$ linear unabhängig sind. □

Teil *b)* des Lemmas erklärt, dass man die Wahl einer Basis auffassen kann als Wahl eines linearen Koordinatensystems auf dem Vektorraum. Das Paradebeispiel hierfür ist $V = K^n$ mit der Standardbasis:

Beispiel 4.4. Im Vektorraum $V = K^n$ bilden die Standard-Basisvektoren $e_1, \dots, e_n$ mit

$$e_i \;=\; \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-te Stelle}$$

eine Basis, diese wird als die *Standard-Basis* von K^n bezeichnet. Die Eigenschaft *b)* im vorigen Lemma besagt hier einfach, dass jeder Vektor aus K^n eine Darstellung als Linearkombination

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \;=\; \alpha_1 e_1 + \cdots + \alpha_n e_n$$

der Standard-Basisvektoren mit eindeutigen Koeffizienten $\alpha_1, \dots, \alpha_n \in K$ besitzt.

Beispiel 4.5. Sei $V = \langle v_1, v_2, v_3 \rangle_{\mathbb{R}} \subseteq \mathbb{R}^3$ der von

$$v_1 \;=\; \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},\; v_2 \;=\; \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix},\; v_3 \;=\; \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

aufgespannte Untervektorraum. Dann gilt:

a) (v_1) ist linear unabhängig, aber kein Erzeugendensystem für V:

$$\langle v_1 \rangle_{\mathbb{R}} \;=\; \left\{ \begin{pmatrix} x \\ x \\ x \end{pmatrix} \mid x \in \mathbb{R} \right\} \;\neq\; V$$

b) (v_1, v_2, v_3) ist ein Erzeugendensystem für V, aber nicht linear unabhängig:

$$v_1 + v_2 - 2v_3 \;=\; \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

c) (v_1, v_2) ist eine Basis für V nach dem obigen Lemma 4.3:

- Wir bemerken zunächst, dass

$$v_1, v_2, v_3 \in W := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid z = x \right\}$$

ist. Für den Aufspann dieser drei Vektoren folgt $V = \langle v_1, v_2, v_3 \rangle \subseteq W$.

- Andererseits hat jeder Vektor aus W die Gestalt

$$w = \begin{pmatrix} x \\ y \\ x \end{pmatrix} \in W$$

mit $x, y \in K$, und eine kurze Rechnung zeigt, dass für jeden solchen Vektor die Gleichung

$$w = \alpha_1 v_1 + \alpha_2 v_2$$

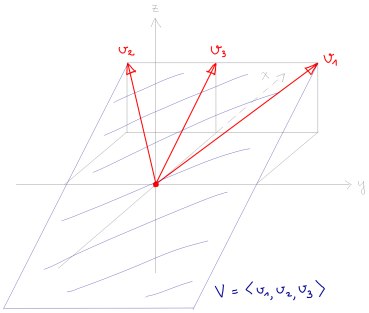
durch eindeutige $\alpha_1, \alpha_2 \in K$ gelöst wird. Explizit ist die eindeutige Lösung gegeben durch

$$\alpha_1 = \frac{x+y}{2} \quad \text{und} \quad \alpha_2 = \frac{x-y}{2}.$$

- Nach dem Lemma 4.3 bilden somit die zwei Vektoren v_1, v_2 eine Basis von W und wegen $v_1, v_2 \in V \subset W$ folgt dann zugleich

$$V = W$$

Somit bilden v_1, v_2 auch eine Basis des Vektorraumes V wie behauptet. Es handelt sich hierbei um eine Ebene:



Als nächstes wollen wir sehen, dass jeder Vektorraum eine Basis hat und wie man eine solche finden kann. Hierzu gibt es zwei Ansätze:

a) top-down: Verkleinere ein gegebenes Erzeugendensystem.

b) bottom-up: Vergrößere ein gegebenes linear unabhängiges System.

Im Folgenden fixieren wir einen K-Vektorraum $V \neq \{0\}$ und beziehen Begriffe wie Basis, Erzeugendensystem usw. stets auf diesen Vektorraum über K.

Satz 4.6. *Für Familien $B = (v_i)_{i\in I}$ von Vektoren in V sind äquivalent:*

a) B ist eine Basis.

b) B ist ein minimales Erzeugendensystem, d.h.

- *B ist ein Erzeugendensystem, aber*
- *für kein $i_0 \in I$ ist $(v_i)_{i\in I\setminus\{i_0\}}$ ein solches.*

c) B ist ein maximales linear unabhängiges System, d.h.

- *B ist ein linear unabhängiges System, aber*
- *wenn man zu B einen beliebigen weiteren Vektor $v \in V$ hinzufügt, wird das so erhaltene System linear abhängig.*

Beweis. Wir zeigen $a \Rightarrow b \Rightarrow c \Rightarrow a$.

Zu $a \Rightarrow b$: Sei $B = (v_i)_{i\in I}$ eine Basis. Nach unserer Definition einer Basis ist dann B insbesondere ein Erzeugendensystem. Wäre B als Erzeugendensystem nicht minimal, dann gäbe es ein $i_0 \in I$ mit

$$\big\langle v_i \mid i \in I \setminus \{i_0\} \big\rangle \;=\; V \;=\; \big\langle v_i \mid i \in I \big\rangle.$$

Dann wäre $B = (v_i)_{i\in I}$ nach Satz 3.8 linear abhängig im Widerspruch zur Annahme.

Zu $b \Rightarrow c$: Sei $B = (v_i)_{i\in I}$ ein minimales Erzeugendensystem. Wäre nun B linear abhängig, so gäbe es nach Satz 3.8 ein $i_0 \in I$ mit

$$\big\langle v_i \mid i \in I \setminus \{i_0\} \big\rangle \;=\; \big\langle v_i \mid i \in I \big\rangle$$

im Widerspruch dazu, dass B ein minimales Erzeugendensystem ist. Also ist B ein linear unabhängiges System. Nehmen wir einen beliebigen Vektor $v \in V$ und einen Index $i_0 \notin I$ hinzu, so wird

$$(v_i)_{i\in I'} \quad \text{mit} \quad I' \;:=\; I \sqcup \{i_0\}, \quad v_{i_0} \;:=\; v$$

linear abhängig nach Satz 3.8: Denn

$$\big\langle v_i \mid i \in I' \setminus \{i_0\} \big\rangle \;=\; \langle v_i \mid i \in I \rangle \;=\; V \;=\; \big\langle v_i \mid i \in I' \big\rangle.$$

Zu $c \Rightarrow a$: Sei $B = (v_i)_{i\in I}$ maximal linear unabhängig. Nach Annahme ist B linear unabhängig. Wäre das System von Vektoren B kein Erzeugendensystem, so gäbe es ein $v \in V$ mit $v \notin \langle v_i \mid i \in I \rangle$. Da B maximal linear unabhängig ist, müßte aus B durch Hinzufügen von v ein linear abhängiges System werden, somit gäbe es $\alpha, \alpha_i \in K$, nicht alle Null, mit

$$\alpha v + \sum_{i\in I} \alpha_i v_i \;=\; 0.$$

Wegen $v \notin \langle v_i \mid i \in I \rangle$ wäre $\alpha = 0$. Aber dann wäre das System B linear abhängig im Widerspruch zur Annahme. □

Korollar 4.7. *Ein K-Vektorraum V ist endlich erzeugt genau dann, wenn er eine endliche Basis besitzt. In diesem Fall können wir aus jedem Erzeugendensystem durch Weglassen von Vektoren eine Basis auswählen.*

Beweis. Jeder Vektorraum mit einer endlichen Basis ist insbesondere auch endlich erzeugt. Sei umgekehrt V endlich erzeugt. Ist ein Erzeugendensystem nicht minimal, so enthält es ein kleineres ES. Somit enthält jedes endliche ES ein minimales ES, und nach Satz 4.6 ist dies eine Basis. □

Insbesondere hat jeder endlich erzeugte Vektorraum eine Basis! Unter Benutzung des Auswahlaxioms aus der Mengenlehre kann man dasselbe sogar für nicht endlich erzeugte Vektorräume beweisen. Zur Erinnerung:

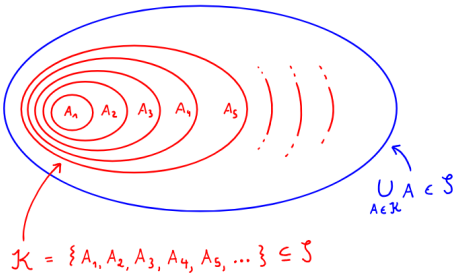
- Sei $\mathscr{S} \subseteq \mathscr{P}(M)$ eine Menge von Teilmengen einer Menge M.
- Wir nennen $\mathscr{S}$ eine *Kette*, wenn sie bezüglich Inklusion total geordnet ist, wenn also für alle $A, B \in \mathscr{S}$ gilt:

$$A \subseteq B \quad \text{oder} \quad B \subseteq A.$$

- Ein Element $A \in \mathscr{S}$ heißt *maximal*, wenn es in $\mathscr{S}$ kein größeres Element gibt, d.h. wenn für alle $B \in \mathscr{S}$ gilt:

$$A \subseteq B \quad \Longrightarrow \quad B = A.$$

Die folgende Abbildung illustriert eine Kette, die abzählbare Kollektion der roten ineinander verschachtelten Mengen. Die blau angedeutete Vereinigungsmenge muß selber nicht Teil der Kette sein:



Wir zitieren hier ohne Beweis die folgende Konsequenz aus dem Auswahlaxiom:

Zorn's Lemma. *Sei M eine Menge, und sei $\mathscr{S} \subseteq \mathscr{P}(M)$ eine nichtleere Menge von Teilmengen von M mit der Eigenschaft, dass für jede Kette $\mathscr{K} \subseteq \mathscr{S}$ auch ihre Vereinigungsmenge in $\mathscr{S}$ liegt:*

$$\bigcup_{A \in \mathscr{K}} A \in \mathscr{S}.$$

Dann besitzt $\mathscr{S}$ ein maximales Element.

Bei der Anwendung von Zorn's Lemma in der obigen Form hilft oft die folgende einfache Beobachtung:

Lemma 4.8. *Sei $\mathscr{K} \subseteq \mathscr{P}(M)$ eine Kette. Zu je endlich vielen $A_1, \ldots, A_n \in \mathscr{K}$ gibt es ein $i_0 \in \{1, \ldots, n\}$ mit*

$$A_i \subseteq A_{i_0} \quad \textit{für alle} \quad i \in \{1, \ldots, n\}.$$

Beweis. Wir nutzen vollständige Induktion über n. Für $n = 1$ ist nichts zu zeigen. Sei also $j_0 \in \{1, \ldots, n-1\}$ gegeben mit $A_j \subseteq A_{j_0}$ für alle $j \in \{1, \ldots, n-1\}$. Dann folgt die gewünschte Aussage mit

$$i_0 := \begin{cases} j_0 & \text{für } A_n \subseteq A_{j_0}, \\ n & \text{für } A_{j_0} \subseteq A_n. \end{cases}$$ □

Satz 4.9. *Jeder Vektorraum V besitzt eine Basis.*

Beweis. Wir wenden das Zorn'sche Lemma an auf $M = V$ und

$$\mathscr{S} = \big\{ A \subseteq V \mid \text{die Familie } (v)_{v \in A} \text{ ist linear unabhängig} \big\}.$$

Wenn die Voraussetzung des Zorn'schen Lemmas erfüllt ist, dann besitzt $\mathscr{S}$ ein bezüglich Inklusion maximales Element. Da maximale linear unabhängige Systeme in V genau die Basen des Vektorraumes V sind, sind wir dann fertig.

Sei also $\mathscr{K} \subseteq \mathscr{S}$ eine Kette. Zu zeigen ist, dass dann auch $B := \bigcup_{A \in \mathscr{K}} A$ in $\mathscr{S}$ liegt, dass also die Elemente von B ein linear unabhängiges System von Vektoren bilden. Seien dazu endlich viele Vektoren $v_1, \ldots, v_n \in B$ beliebig vorgegeben. Per Definition von B existiert in der gegebenen Kette zu jedem $i \in \{1, \ldots, n\}$ ein $A_i \in \mathscr{K}$ mit $v_i \in A_i$. Nach Lemma 4.8 über endliche Teilmengen von Ketten gibt es dann sogar einen Index i_0 mit $A_i \subseteq A_{i_0}$ für alle i. Dann liegen die Vektoren $v_1, \ldots, v_n$ also alle in A_{i_0}. Wegen $A_{i_0} \in \mathscr{S}$ sind sie somit linear unabhängig. □

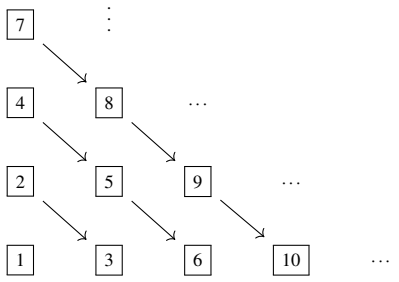
5 Dimension von Vektorräumen

Die Interpretation von Basen als lineare Koordinatensysteme zeigt, dass diese nicht eindeutig sind. Das ist sehr nützlich, weil wir später zu jedem Problem die passende Basis wählen können! Intuitiv scheint aber klar, dass die Anzahl der Koordinaten nicht von der gewählten Basis abhängt: Wir beschreiben

- Punkte auf der Geraden $V = \mathbb{R}$ durch *eine* reelle Zahl,
- Punkte in der Ebene $V = \mathbb{R}^2$ durch *zwei* reelle Zahlen,
- Punkte im Raum $V = \mathbb{R}^3$ durch *drei* reelle Zahlen, etc.

Tatsächlich werden wir sehen, dass alle Basen eines gegebenen Vektorraum gleich viele Elemente haben. Das ist durchaus nicht offensichtlich — um zu sehen, dass es hier etwas zu beweisen gilt, betrachten wir ein Analogon aus der Mengenlehre:

Beispiel 5.1. Es gibt bijektive Abbildungen $f : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$ von Mengen, auch wenn die rechte Seite größer als die linke aussieht. Cantors Diagonal-Abzählung ist ein Beispiel:



Durch Tricksen mit Dezimalentwicklungen reeller Zahlen kann man in ähnlicher Weise auch Bijektionen $f\colon \mathbb{R} \longrightarrow \mathbb{R}^2$ konstruieren (Übungsaufgabe).

In der linearen Algebra begegnet uns so etwas nicht! Um zu zeigen, dass alle Basen eines Vektorraum gleich viele Vektoren enthalten, schauen wir zuerst, wie man ein linear unabhängiges System zu einer Basis ergänzen kann.

Beispiel 5.2. In $V = \mathbb{R}^2$ sei ein Vektor $u = \alpha_1 e_1 + \alpha_2 e_2$ gegeben. Dann gilt:

a) Für $\alpha_2 \neq 0$ ist (u, e_1) eine Basis.

b) Für $\alpha_1 \neq 0$ ist (u, e_2) eine Basis.

Allgemein gilt:

Satz 5.3 (Basisergänzungssatz). *Sei V ein K-Vektorraum, und es seien*

- *ein linear unabhängiges System $(u_i)_{i\in I}$*
- *und ein Erzeugendensystem $(v_j)_{j\in J}$ gegeben.*

Dann gibt es eine Teilmenge $J_0 \subseteq J$ mit der Eigenschaft, dass das System

$$B := (w_k)_{k\in I\sqcup J_0} \quad \text{mit} \quad w_k := \begin{cases} u_i & \text{für } k = i \in I \\ v_j & \text{für } k = j \in J_0 \end{cases}$$

eine Basis des K-Vektorraumes V bildet.

Beweis. Sei $J_0 \subseteq J$ eine bezüglich Inklusion maximal gewählte Teilmenge mit der Eigenschaft, dass

$$B := (w_k)_{i\in I\sqcup J_0} \quad \text{mit} \quad w_k := \begin{cases} u_i & \text{für } k = i \in I \\ v_j & \text{für } k = j \in J_0 \end{cases}$$

ein linear unabhängiges System ist. Dass es so eine Teilmenge gibt, ist für endliche Indexmengen J klar. Für unendliche J folgt es aus Zorn's Lemma.

Wegen der Maximalität ist für jeden Index $j \in J \setminus J_0$ das System, das aus B durch Hinzufügen von v_j entsteht, linear abhängig. Es gibt also Koeffizienten $\alpha, \alpha_i \in K$, nicht alle Null, mit

$$\alpha \cdot v_j + \sum_{i \in I \sqcup J_0} \alpha_i \cdot w_i \;=\; 0.$$

Genauer gilt $\alpha \neq 0$, denn die w_i mit $i \in I \sqcup J_0$ bilden per Konstruktion ein linear unabhängiges System. Somit folgt

$$v_j \;=\; -\sum_{i \in I \sqcup J_0} \frac{\alpha_i}{\alpha} \cdot w_i \;\in\; \langle w_i \mid i \in I \sqcup J_0 \rangle \;=\; \langle B \rangle.$$

Es ist also $v_j \in \langle B \rangle$ für alle $j \in J \setminus J_0$. Für $j \in J_0$ gilt per Definition sogar $v_j \in B \subseteq \langle B \rangle$, insgesamt also

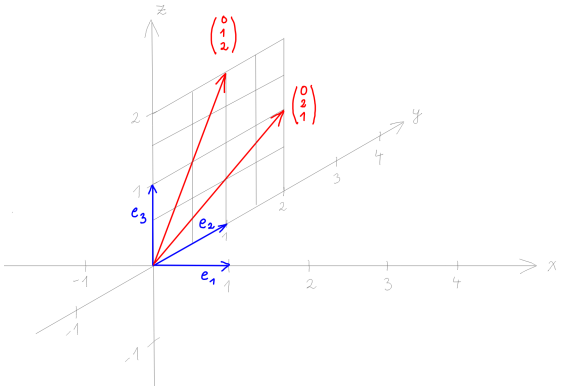
$$v_j \in \langle B \rangle \quad \text{für alle } j \in J.$$

Mit $(v_j)_{j \in J}$ ist dann auch B ein Erzeugendensystem, und dieses ist per Konstruktion linear unabhängig. □

Der obige Satz wird es uns erlauben, einige Vektoren aus einer Basis durch solche aus einem anderen linear unabhängigen System auszutauschen. Ein Beispiel:

Beispiel 5.4. In $V = \mathbb{R}^3$ betrachte man die Standardbasis (e_1, e_2, e_3) und das linear unabhängige System (u_1, u_2) mit

$$u_1 \;=\; \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad u_2 \;=\; \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}.$$



Ein Austausch der ersten zwei Vektoren in der Standardbasis liefert $\widetilde{B} = (u_1, u_2, e_3)$, dies ist keine Basis. Für die umnumerierte Basis (e_2, e_3, e_1) führt ein Austausch der ersten beiden Vektoren jedoch zu der Basis

$$\widetilde{B} \;=\; (u_1, u_2, e_1).$$

Satz 5.5 (Basisaustauschsatz). *Sei V ein K-Vektorraum und*

- $B = (v_1, \ldots, v_n)$ *eine endliche Basis,*
- $C = (u_1, \ldots, u_m)$ *ein linear unabhängiges System.*

Dann ist $m \leq n$, und nach eventuellem Umnumerieren der Vektoren in B ist das durch Austausch der ersten m Vektoren erhaltene System

$$\tilde{B} \;=\; (u_1, \ldots, u_m, v_{m+1}, \ldots, v_n) \quad \textit{eine Basis von } V \textit{ über } K.$$

Beweis. Nach dem Basisergänzungssatz können wir $(u_1, \ldots, u_m)$ ergänzen zu einer Basis

$$(u_1,\ \ldots,\ u_m,\ v_{j_1},\ v_{j_2},\ \ldots,\ v_{j_{k_0}})$$

für ein $k_0 \geq 0$ und geeignete Indices $j_\alpha \in \{1, \ldots, n\}$. Wenn wir wüßten, dass je zwei Basen eines Vektorraums gleich viele Elemente haben, wäre $m + k_0 = n$. Nach Umnumerieren können wir dann $j_\alpha = m + \alpha$ für $\alpha = 1, 2, \ldots, n - m$ annehmen und wären fertig. Aber wir wissen noch nicht, dass je zwei Basen gleich viele Elemente haben! Wir argumentieren daher anders und lassen sukzessive Vektoren u_i weg:

Das System $(u_1, \ldots, u_{m-1}, v_{j_1}, \ldots, v_{j_{k_0}})$ ist noch immer linear unabhängig, aber bildet keine Basis mehr. Nach dem Basisergänzungssatz können wir es ergänzen zu einer Basis

$$(u_1,\ \ldots,\ u_{m-1},\ v_{j_1},\ \ldots,\ v_{j_{k_0}},\ v_{j_{k_0+1}},\ \ldots,\ v_{j_{k_1}})$$

mit $k_1 > k_0$ und weiteren $j_\alpha \in \{1, \ldots, n\}$. Induktiv fortfahrend, erhalten wir im r-ten Schritt eine Basis

$$(u_1,\ \ldots,\ u_{m-r},\ v_{j_1},\ \ldots,\ v_{j_{k_{r-1}}},\ v_{j_{k_{r-1}+1}},\ \ldots, v_{j_{k_r}})$$

mit $k_r > k_{r-1}$ und weiteren $j_\alpha \in \{1, \ldots, n\}$. Nach m Schritten sind alle u_i ersetzt und wir erhalten eine neue Basis

$$(v_{j_1}, \ldots, v_{j_{k_m}}),$$

die ausschließlich Vektoren aus B enthält. Da B ein *minimales* Erzeugendensystem ist, muß diese neue Basis bis auf Umordnen mit der Basis B übereinstimmen, für die Mengen der Indices gilt also:

$$\{v_{j_1}, \ldots, v_{j_{k_m}}\} \;=\; \{v_1, \ldots, v_n\}$$

Es folgt $n = k_m$, da die Vektoren jeder Basis wegen der linearen Unabhängigkeit paarweise verschieden sind. Aber $k_m \geq m$, da wir ab dem ersten Schritt bei jeder Anwendung des Basisergänzungssatzes mindestens einen Vektor ergänzt haben und somit $k_m > k_{m-1} > \cdots > k_1 > 0$ ist. Insgesamt haben wir damit $n \geq m$ gezeigt für

- jede Basis $B = (v_1, \ldots, v_n)$,
- jedes linear unabhängige System $C = (u_1, \ldots, u_m)$.

Wenn C auch eine Basis ist, folgt per Symmetrie $n = m$. Also bestehen je zwei Basen des Vektorraumes V aus gleich vielen Vektoren und wir sind fertig. □

Wir halten die wichtigste Einsicht aus dem soeben gegebenen Beweis an dieser Stelle nochmals explizit fest:

Korollar 5.6. *Sei V ein endlich erzeugter K-Vektorraum. Dann bestehen je zwei Basen von V aus gleich vielen Vektoren.* □

Definition 5.7. Die *Dimension* eines Vektorraumes V über K ist definiert als

$$\dim_K(V) = \begin{cases} n & \text{falls } V \text{ eine Basis der Länge } n \in \mathbb{N}_0 \text{ hat,} \\ \infty & \text{falls } V \text{ nicht endlich erzeugt ist.} \end{cases}$$

Wenn K aus dem Kontext klar ist, schreiben wir auch kurz $\dim(V)$ statt $\dim_K(V)$.

Bemerkung 5.8. Mit etwas mehr Mengenlehre kann man zeigen, dass auch für nicht endlich erzeugte Vektorräume V gilt: Für je zwei Basen $(u_i)_{i\in I}$ und $(v_j)_{j\in J}$ von V existiert eine Bijektion

$$\varphi: \quad I \longrightarrow J$$

Wir werden diese präzisere Aussage in dieser Vorlesung nicht benötigen, für uns genügt es, im nicht endlich erzeugten Fall $\dim_K(V) = \infty$ zu schreiben.

Beispiel 5.9. Es gilt:

a) $\dim_K(V) = 0$ genau dann, wenn $V = \{0\}$ ist.

b) Für $V = K^n$ zeigt die Standardbasis, dass $\dim_K(V) = n$ ist.

c) Die Dimension eines Vektorraumes hängt davon ab, über welchem Körper wir arbeiten: Z.B. ist $V = \mathbb{C}$ ein komplexer Vektorraum der Dimension $\dim_{\mathbb{C}}(V) = 1$, aber ein reeller Vektorraum der Dimension $\dim_{\mathbb{R}}(V) = 2$.

d) Die Untervektorräume $V \subseteq \mathbb{R}^3$ mit $\dim_{\mathbb{R}}(V) = 1$ sind genau die Geraden durch den Ursprung, d.h. die linearen Hüllen von je einem Vektor $v \in \mathbb{R}^3 \setminus \{0\}$.

e) Die Untervektorräume $V \subseteq \mathbb{R}^3$ mit $\dim_{\mathbb{R}}(V) = 2$ sind genau die Ebenen durch den Ursprung, d.h. die linearen Hüllen von zwei linear unabhängigen Vektoren. Auch mehr als zwei Vektoren können natürlich eine Ebene aufspannen: Für die Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

ist $V = \langle v_1, v_2, v_3 \rangle_{\mathbb{R}} = \langle v_1, v_2 \rangle_{\mathbb{R}} \subset \mathbb{R}^3$ ein Untervektorraum mit $\dim_{\mathbb{R}}(V) = 2$.

Lemma 5.10. *Für K-Vektorräume V sind äquivalent:*

a) $\dim_K(V) = \infty$

b) Für jedes $n \in \mathbb{N}$ gibt es in V ein linear unabhängiges System aus n Vektoren.

c) Es gibt eine Folge von $v_1, v_2, \cdots \in V$, sodass $(v_n)_{n\in\mathbb{N}}$ linear unabhängig ist.

Beweis. Im Fall $a)$ gibt es in V keine endlichen maximalen linear unabhängigen Systeme. Man kann also induktiv $v_i \in V$ finden, sodass $(v_1, \ldots, v_n)$ für jedes $n \in \mathbb{N}$ ein linear unabhängiges System bildet. Per Definition der linearen Unabhängigkeit für unendliche Systeme von Vektoren gilt dann $c)$. Aus $c)$ folgt trivialerweise $b)$, und aus $b)$ folgt $a)$: Denn nach dem Basisaustauschsatz kann man jedes der linear unabhängigen Systeme aus $b)$ ergänzen zu einer Basis. Es existiert somit in V für jedes $n \in \mathbb{N}$ eine Basis, die mehr als n Vektoren enthält. Nach Korollar 5.6 muß dann $\dim_K(V) = \infty$ sein. □

Lemma 5.11. *Sei V ein K-Vektorraum mit $n = \dim_K(V) < \infty$, und sei $B = (v_1, \ldots, v_n)$ eine Familie von genau n Vektoren darin. Dann sind äquivalent:*

a) B ist eine Basis.
b) B ist linear unabhängig.
c) B ist ein Erzeugendensystem von V.

Beweis. Aus $a)$ folgt $b)$ und $c)$ per Definition einer Basis. Aus $b)$ erhält man $a)$, indem man $m = n$ im Basisaustauschsatz 5.5 wählt. Auch aus $c)$ folgt $a)$, denn jedes Erzeugendensystem lässt sich zu einer Basis verkleinern und jede Basis besteht nach Korollar 5.6 aus genau n Vektoren. □

Beispiel 5.12. Um zu sehen, ob in $V = K^n$ ein System von n gegebenen Vektoren eine Basis ist, müssen wir nur lineare Unabhängigkeit prüfen. In $V = \mathbb{R}^3$ betrachte man z.B.

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Um die lineare Unabhängigkeit von v_1, v_2, v_3 nachzuweisen, müssen wir zeigen, dass die Gleichung

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$$

keine Lösung $(\alpha_1, \alpha_2, \alpha_3) \neq (0,0,0)$ hat. Diese Gleichung ist ein homogenes LGS:

$$\begin{aligned} \alpha_1 - \alpha_2 &= 0 \\ \alpha_1 + \alpha_2 - \alpha_3 &= 0 \\ \alpha_1 + \alpha_2 + \alpha_3 &= 0 \end{aligned}$$

Dieses besitzt als einzige Lösung $(\alpha_1, \alpha_2, \alpha_3) = (0,0,0)$. Somit sind v_1, v_2, v_3 linear unabhängig. Nach Lemma 5.11 bilden sie dann eine Basis: Wir bekommen gratis dazu, dass das inhomogene LGS

$$\begin{aligned} \alpha_1 - \alpha_2 &= c_1 \\ \alpha_1 + \alpha_2 - \alpha_3 &= c_2 \\ \alpha_1 + \alpha_2 + \alpha_3 &= c_3 \end{aligned}$$

für alle $c_i \in \mathbb{R}$ eine eindeutige Lösung $(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$ besitzt!

Ein Dimensionsargument kann uns also beim Nachweis der Basiseigenschaft die Hälfte der Arbeit sparen. Zum Schluß dieses Abschnitts wollen wir noch folgende nützliche Abschätzung für Untervektorräume festhalte:

Lemma 5.13. *Ist V ein endlich erzeugter K-Vektorraum, dann gilt dasselbe auch für jeden Untervektorraum $W \subseteq V$. Für die Dimension gilt dabei* $\dim(W) \leq \dim(V)$*, und Gleichheit gilt genau dann, wenn $W = V$ ist.*

Beweis. Jede linear unabhängige Familie in W ist auch eine solche in V und besteht somit nach dem Basisaustauschsatz aus höchstens $\dim(V)$ Elementen. Also ist W endlich erzeugt mit $\dim(W) \leq \dim(V)$. Im Fall von Gleichheit ist nach dem vorigen Lemma 5.11 jede Basis von W bereits eine Basis von V. □

Für Vektorräume unendlicher Dimension gilt die letzte Aussage nicht, selbst wenn wir unseren mengentheoretisch naiven Begriff der Dimension ∞ ersetzen durch die Kardinalität einer Basis: Der K-Vektorraum $V = K[x]$ hat eine Basis aus den Monomen x^n mit $n \in \mathbb{N}_0$. Die Menge aller Polynome mit konstantem Term Null ist ein Untervektorraum $U \subseteq V$ mit einer Basis aus den Monomen x^n mit $n \in \mathbb{N}$; beide Basen sind abzählbar unendlich, aber $U \neq V$.

6 Direkte Summen

Bei der Diskussion linearer Unabhängigkeit haben wir den Aufspann von Familien von Vektoren betrachtet. Dasselbe geht für Familien von Untervektorräumen:

Definition 6.1. Sei V ein Vektorraum über K. Für Unterräume $U_1, \ldots, U_r \subseteq V$ ist ihre *Summe in V* definiert durch

$$U_1 + \cdots + U_r := \langle U_1 \cup \cdots \cup U_r \rangle_K.$$

Aus der Definition als Aufspann ist klar, dass die Summe von Untervektorräumen wieder ein Untervektorraum ist. Da die $U_i \subseteq V$ Untervektorräume sind, folgt aus der Definition zudem

$$U_1 + \cdots + U_r = \{u_1 + \cdots + u_r \mid u_i \in U_i \text{ für alle } i\}.$$

Die Darstellung von Vektoren als Summe $u_1 + \cdots + u_r$ muß nicht eindeutig sein:

Beispiel 6.2. In $V = \mathbb{R}^3$ betrachte man die Untervektorräume

$$U_1 = \{(x,y,z) \in \mathbb{R}^3 \mid x = 0\},$$
$$U_2 = \{(x,y,z) \in \mathbb{R}^3 \mid z = 0\}.$$

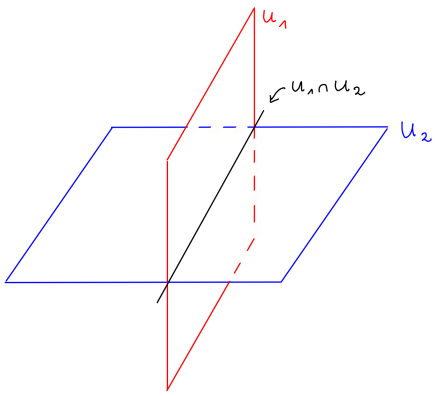
In $U_1 + U_2 = \mathbb{R}^3$ ist

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

Umstellen dieser Gleichung zeigt

$$\begin{pmatrix}0\\1\\1\end{pmatrix} - \begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}1\\1\\0\end{pmatrix} - \begin{pmatrix}1\\0\\0\end{pmatrix} \in U_1 \cap U_2,$$

die Mehrdeutigkeit der Zerlegung hat also damit zu tun, dass $U_1 \cap U_2 \neq \{0\}$ ist:



Tatsächlich sind in jedem Vektorraum V für je zwei Untervektorräume $U_1, U_2 \subseteq V$ folgende Eigenschaften zueinander äquivalent:

a) Es ist $U_1 \cap U_2 = \{0\}$.

b) Aus $u_1 + u_2 = 0$ mit $u_i \in U_i$ folgt $u_1 = u_2 = 0$.

c) Jedes $u \in U_1 + U_2$ hat eine eindeutige Zerlegung $u = u_1 + u_2$ mit $u_i \in U_i$.

Um Arbeit zu sparen, wollen wir die analoge Aussage direkt für die Summe von beliebig vielen Untervektorräumen $U_1, \ldots, U_r \subseteq V$ zeigen. Die Verallgemeinerung der Bedingung a) auf $r > 2$ Summanden erfordert allerdings etwas Vorsicht: Hier genügt es nicht, lediglich $U_i \cap U_j = \{0\}$ für alle $i \neq j$ zu fordern! In $V = \mathbb{R}^2$ betrachte man z.B.

$$\begin{aligned} U_1 &= \langle e_1 \rangle_{\mathbb{R}}, \\ U_2 &= \langle e_2 \rangle_{\mathbb{R}}, \\ U_3 &= \langle e_1 + e_2 \rangle_{\mathbb{R}}. \end{aligned}$$

Dann ist $U_i \cap U_j = \{0\}$ für alle $i \neq j$, aber Vektoren aus $U_1 + U_2 + U_3$ lassen sich auf mehr als eine Weise in eine Summe von Vektoren aus den drei Unterräumen zerlegen:

$$0 + 0 + (e_1 + e_2) = e_1 + e_2 + 0.$$

Die richtige Verallgemeinerung sieht so aus:

Lemma 6.3. *Für Untervektorräume $U_1, \dots, U_r \subseteq V$ sind äquivalent:*

a) $U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_r) = \{0\}$ *für alle* i.

b) Aus $u_1 + \cdots + u_r = 0$ *mit Vektoren* $u_i \in U_i$ *folgt* $u_1 = \cdots = u_r = 0$.

c) Jedes $u \in U_1 + \cdots + U_r$ *hat eine eindeutige Zerlegung* $u = u_1 + \cdots + u_r$ *mit* $u_i \in U_i$.

Beweis. Wir zeigen $c) \Rightarrow b) \Rightarrow a) \Rightarrow c)$.

Aus $c)$ folgt sofort $b)$. Um von Eigenschaft $b)$ auf $a)$ zu schließen, seien $v_i \in U_i$ gegeben mit

$$v_{i_0} \;=\; v_1 + \cdots + v_{i_0-1} + v_{i_0+1} + \cdots + v_r \;\in\; U_{i_0} \cap (U_1 + \cdots + U_{i_0-1} + U_{i_0+1} \cdots + U_r)$$

für ein i_0. Dann folgt

$$u_1 + \cdots + u_r \;=\; 0 \quad \text{mit} \quad u_i \;=\; \begin{cases} v_i & \text{für } i \neq i_0, \\ -v_{i_0} & \text{für } i = i_0. \end{cases}$$

Wenn $b)$ gilt, so folgt hieraus $u_i = 0$ für alle i und somit folgt $a)$.

Zu zeigen bleibt noch, dass aus $a)$ auch $c)$ folgt. Die Existenz einer Darstellung wie in $c)$ folgt aus der Definition der Summe von Untervektorräumen, es geht also nur um die Eindeutigkeit. Dazu seien $v_i, w_i \in U_i$ mit $v_1 + \cdots + v_r \;=\; w_1 + \cdots + w_r$ gegeben. Dann folgt $u_1 + \cdots + u_r = 0$ für $u_i = v_i - w_i$. Somit ist

$$u_i = -u_1 - \cdots - u_{i-1} - u_{i+1} - \cdots - u_r \in U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_r)$$

und nach der Voraussetzung $a)$ folgt hieraus $u_i = 0$, also $v_i = w_i$. □

Definition 6.4. Wenn die äquivalenten Bedingungen aus dem Lemma 6.3 gelten, bezeichnen wir die Summe $U = U_1 + \cdots + U_r$ als *direkt* und schreiben in diesem Fall auch

$$U \;=\; U_1 \oplus \cdots \oplus U_r \;=\; \bigoplus_{i=1}^{r} U_i.$$

Diese Definition verallgemeinert den Begriff der linearen Unabhängigkeit:

Lemma 6.5. *Für* $U_i = \langle u_i \rangle \subseteq V$ *mit* $u_i \in V \setminus \{0\}$ *sind äquivalent:*

a) Die Summe $U_1 + \cdots + U_r \subseteq V$ *ist direkt.*

b) Die Vektoren $u_1, \dots, u_r$ *sind linear unabhängig.*

Beweis. Nach Lemma 6.3 ist die Summe direkt genau dann, wenn für alle $v_i \in U_i$ gilt:

$$v_1 + \cdots + v_r \;=\; 0 \quad \Longrightarrow \quad v_1 = \cdots = v_r = 0.$$

Mit $U_i = \{\alpha_i u_i \mid \alpha_i \in K\}$ wird diese Bedingung zu

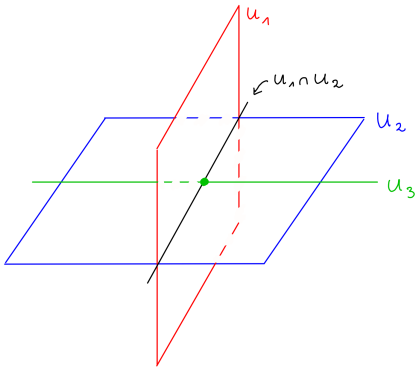
$$\alpha_1 u_1 + \cdots + \alpha_r v_r \;=\; 0 \quad \Longrightarrow \quad \alpha_1 = \cdots = \alpha_r = 0.$$

Das ist genau die Bedingung für lineare Unabhängigkeit. □

Beispiel 6.6. In $V = \mathbb{R}^3$ betrachte man

$$\begin{aligned} U_1 &= \{(x,y,z) \in \mathbb{R}^3 \mid x = 0\}, \\ U_2 &= \{(x,y,z) \in \mathbb{R}^3 \mid z = 0\}, \\ U_3 &= \langle v \rangle_{\mathbb{R}} \quad \text{für ein } v = (x,y,z) \in \mathbb{R}^3 \text{ mit } x \neq 0. \end{aligned}$$

Hier ist die Summe $\mathbb{R}^3 = U_1 + U_3$ direkt, aber die Summe $\mathbb{R}^3 = U_1 + U_2$ nicht:



Satz 6.7. *Sei V ein Vektorraum über K und $U \subseteq V$ ein Untervektorraum. Dann gibt es einen Untervektorraum $U' \subseteq V$ mit der Eigenschaft $V = U \oplus U'$ und für jeden solchen gilt*

$$\dim_K(V) = \dim_K(U) + \dim_K(U').$$

Beweis. Sei $(u_i)_{i \in I}$ eine Basis von U. Der Basisergänzungssatz besagt, dass wir durch Hinzunahme von Vektoren eines Erzeugendensystems $(v_j)_{j \in J}$ eine Basis von V der Gestalt

$$(w_k)_{k \in I \sqcup J_0} \quad \text{mit} \quad w_k = \begin{cases} u_k & \text{für } k \in I \\ v_k & \text{für } k \in J_0 \subseteq J \end{cases}$$

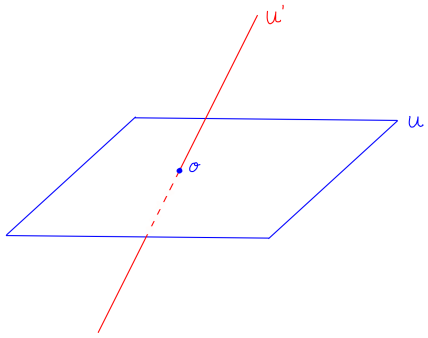
erhalten. Dann ist

$$V = \langle w_k \mid k \in I \sqcup J_0 \rangle = \langle w_k \mid k \in I \rangle \oplus \langle w_k \mid k \in J_0 \rangle,$$

wobei die Direktheit der Summe aus Lemma 6.3 und der linearen Unabhängigkeit von $(w_k)_{k \in I \sqcup J_0}$ folgt. Wir können also $U' = \langle w_k \mid k \in J_0 \rangle$ wählen. □

Einen Untervektorraum $U' \subseteq V$ mit $V = U \oplus U'$ nennt man auch ein *Komplement* von U in V. Man beachte, dass Komplemente nicht eindeutig sind: Beispielsweise

ist für eine Ebene $U \subseteq \mathbb{R}^3$ jede nicht in U enthaltene Gerade $U' \subseteq \mathbb{R}^3$ durch den Ursprung ein Komplement:



Als Anwendung der Formel für die Dimension im obigen Satz erhalten wir folgende sehr nützliche Formel:

Korollar 6.8 (Dimensionsformel für Unterräume). *Es sei V ein K-Vektorraum, und $U_1, U_2 \subseteq V$ seien zwei Untervektorräume endlicher Dimension. Dann ist*

$$\dim_K(U_1 + U_2) \;=\; \dim_K(U_1) + \dim_K(U_2) - \dim_K(U_1 \cap U_2).$$

Beweis. Sei $U := U_1 \cap U_2$, und sei W_i ein Komplement von $U_1 \cap U_2$ in U_i. Dann gilt also

$$U_1 \;=\; U \oplus W_1 \quad \text{und} \quad U_2 \;=\; U \oplus W_2$$

und somit insbesondere $U_1 + U_2 = U + W_1 + W_2$. Wir zeigen nun, dass die Summe auf der rechten Seite eine direkte Summe ist: Angenommen, es gilt $u + w_1 + w_2 = 0$ für Vektoren $u \in U$, $w_i \in W_i$. Dann ist

$$w_1 \;=\; -(u + w_2) \;\in\; W_1 \cap (U + W_2).$$

Aber

$$\begin{aligned} W_1 \cap (U + W_2) \;&=\; W_1 \cap U_2 && \text{wegen } U_2 = U \oplus W_2 \\ &\subseteq\; W_1 \cap U_1 \cap U_2 && \text{wegen } W_1 \subseteq U_1 \\ &=\; W_1 \cap U && \text{wegen } U = U_1 \cap U_2 \\ &=\; \{0\} && \text{wegen } U_1 = U \oplus W_1. \end{aligned}$$

Also folgt $w_1 = 0$. Analog sieht man auch $w_2 = 0$ und damit $u = 0$. Wir haben damit gezeigt:

$$U_1 + U_2 \;=\; U \oplus W_1 \oplus W_2.$$

Aus der Formel für die Dimension direkter Summen aus dem vorigen Satz erhalten wir somit

$$\dim(U_1 + U_2) \;=\; \dim(U) + \dim(W_1) + \dim(W_2).$$

Dieselbe Formel liefert

$$\dim(W_i) \;=\; \dim(U_i) - \dim(U) \quad \text{wegen} \quad U_i = U \oplus W_i.$$

Also folgt wie gewünscht $\dim(U_1+U_2) = \dim(U_1) + \dim(U_2) - \dim(U)$. □

Korollar 6.9. *Für Untervektorräume $U_1, U_2 \subseteq V$ sind äquivalent:*

a) $U_1 \cap U_2 = \{0\}$.
b) $\dim(U_1+U_2) = \dim(U_1) + \dim(U_2)$.
c) Die Summe $U_1+U_2 \subseteq V$ ist direkt.

Beweis. Wir haben bereits gesehen, dass die Direktheit der Summe äquivalent ist zu $U_1 \cap U_2 = \{0\}$. Die Behauptung folgt somit aus der Dimensionsformel. □

Wir haben bisher nur die Summe von Untervektorräumen in einem gegebenen Vektorraum betrachtet. Die Direktheit der Summe war dabei eine *Eigenschaft*, die erfüllt sein konnte oder auch nicht. Für Familien von Vektorräumen, die nicht als Untervektorräume eines festen Vektorraumes gegeben sind, ergibt es keinen Sinn, von ihrer Summe zu sprechen. Man kann jedoch einen sie enthaltenden Vektorraum formal wie folgt konstruieren, wobei die Direktheit der Summe in die *Definition* eingebaut ist:

Definition 6.10. Gegeben seien Vektorräume $U_1, \dots, U_r$ über K. Wir machen das Produkt

$$V \;=\; U_1 \times \cdots \times U_r$$

zu einem Vektorraum mit komponentenweiser Addition und Skalarmultiplikation, d.h. wir setzen

$$\begin{aligned} (u_1, \dots, u_r) + (v_1, \dots, v_r) \;&:=\; (u_1+v_1, \dots, u_n+v_r) && \text{für } u_i, v_i \in U_i \\ \alpha \cdot (u_1, \dots, u_r) \;&:=\; (\alpha u_1, \dots, \alpha u_r) && \text{und } \alpha \in K. \end{aligned}$$

Mittels der injektiven Abbildung

$$f_i\colon \quad U_i \;\hookrightarrow\; V \;=\; U_1 \times \cdots \times U_r, \quad u \;\mapsto\; (0, \dots, 0, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{u}, 0, \dots, 0)$$

können wir U_i mit einem Untervektorraum $f(U_i) \subseteq V$ identifizieren, und dabei ist offenbar

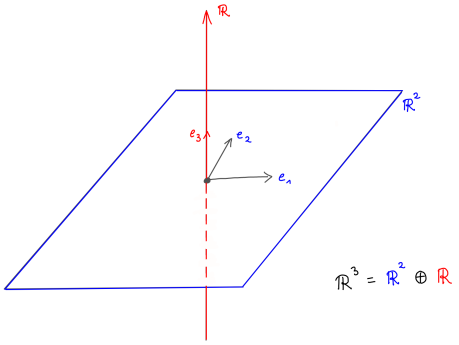
$$V \;=\; f_1(U_1) \oplus \cdots \oplus f_r(U_r).$$

In der Praxis unterdrückt man die injektiven Abbildungen f_i in der Notation: Man schreibt kurz

$$V \;=\; U_1 \oplus \cdots \oplus U_r$$

und nennt diesen Vektorraum die *externe (oder konstruierte) direkte Summe* der U_i.

Beispiel 6.11. Die externe direkte Summe der Standardvektorräume K^m und K^n ist der Standardvektorraum $K^{m+n} = K^m \oplus K^n$ (für $m,n \in \mathbb{N}$):



Bemerkung 6.12. Man sollte die beiden Konzepte von direkten Summen sorgfältig voneinander unterscheiden:

a) Die Direktheit einer Summe von Untervektorräumen ist eine *Eigenschaft*,

b) Die externe direkte Summe von Vektorräumen ist eine *Konstruktion*.

Leider wird für beides die Notation $\oplus$ verwendet. Dabei ist Vorsicht geboten: Da jeder Untervektorraum insbesondere selber ein Vektorraum ist, kann man auch für Untervektorräume $U_1, U_2 \subseteq V$ stets ihre externe direkte Summe bilden, unabhängig davon, ob die im umbegenden Vektorraum V gebildete Summe $U_1 + U_2 \subseteq V$ direkt ist oder nicht. Genauer sind äquivalent:

a) Die in V gebildete Summe $U_1 + U_2$ von Untervektorräumen ist direkt.

b) Es ist $\dim_K(U_1 + U_2) = \dim_K(U_1 \oplus U_2)$ für die externe direkte Summe $U_1 \oplus U_2$.

Kapitel III
Lineare Abbildungen und Matrizen

Zusammenfassung Das Studium linearer Abbildungen zwischen Vektorräumen ist die zentrale Aufgabe der linearen Algebra und ihrer Anwendungen. Wir werden in diesem Kapitel sehen, wie man lineare Abbildungen nach Wahl einer Basis im Definitions- und Zielraum explizit durch Matrizen beschreiben kann, mit denen sich wunderbar rechnen lässt, und wie sich diese Matrizen bei einem Wechsel der Basen ändern. Ganz nebenbei werden wir dabei die Struktur der Lösungsmenge linearer Gleichungssysteme verstehen und den Gauß-Algorithmus in neuem Licht sehen.

1 Lineare Abbildungen

Eine lineare Abbildung von Vektorräumen ist eine Abbildung, die kompatibel ist mit der Addition und Skalarmultiplikation:

Definition 1.1. Wir sagen, eine Abbildung $f : V \to W$ zwischen Vektorräumen über einem Körper K sei *K-linear* oder ein *Homomorphismus*, wenn gilt:

a) $f(u+v) = f(u) + f(v)$ für alle $u, v \in V$,

b) $f(\alpha \cdot v) = \alpha \cdot f(v)$ für alle $v \in V$ und alle $\alpha \in K$.

Wir schreiben

$$\mathrm{Hom}_K(V,W) \;=\; \big\{ f \colon V \to W \mid f \text{ ist } K\text{-linear} \big\}.$$

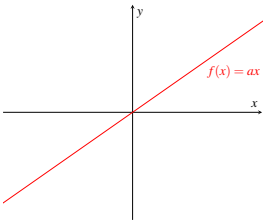
Eine lineare Abbildung $f \colon V \to W$ heißt

a) *Monomorphismus*, falls sie injektiv ist. Notation: $f \colon V \hookrightarrow W$.

b) *Epimorphismus*, falls sie surjektiv ist. Notation: $f \colon V \twoheadrightarrow W$.

c) *Isomorphismus*, falls sie bijektiv ist. Notation: $f \colon V \xrightarrow{\sim} W$.

Beispiel 1.2. Für $a \in \mathbb{R}$ ist $f : \mathbb{R} \longrightarrow \mathbb{R}, f(x) = ax$ eine lineare Abbildung. Ihr Graph ist eine Gerade durch den Ursprung:

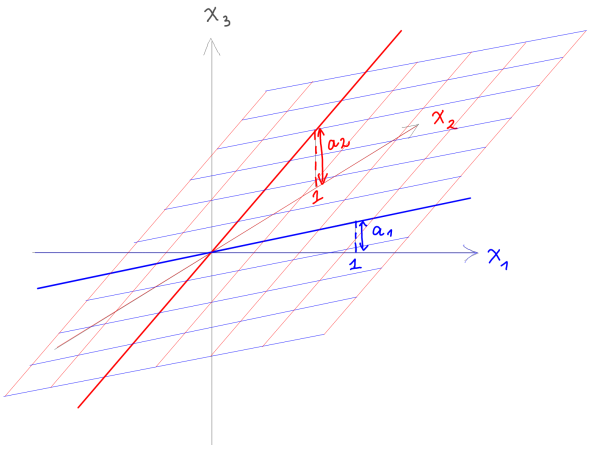


Man beachte, dass Abbildungen der Form $g : \mathbb{R} \to \mathbb{R}, x \mapsto ax + b$ für $b \neq 0$ nicht linear im Sinn unserer Definition sind, z.B. ist

$$g(x_1 + x_2) \;=\; a(x_1 + x_2) + b \;\neq\; (ax_1 + b) + (ax_2 + b) \;=\; g(x_1) + g(x_2) \quad \text{für } b \neq 0.$$

Umgangssprachlich werden solche Abbildungen zwar oft auch als linear bezeichnet, der korrekte Begriff hierfür lautet jedoch *affin-lineare Abbildung*.

Beispiel 1.3. Für $(a_1, a_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$ ist $f \colon \mathbb{R}^2 \longrightarrow \mathbb{R},\ f(x) = a_1x_1 + a_2x_2$ eine lineare Abbildung. Ihr Graph ist eine Ebene durch den Ursprung:



Allgemein bezeichnen wir lineare Abbildungen $f \colon V \to K$ von einem K-Vektorraum nach K auch als *Linearformen* auf diesem Vektorraum.

Beispiel 1.4. Lineare Abbildungen treten auch in der Analysis auf: Beispielsweise sei $V = C([0,1])$ der reelle Vektorraum aller stetigen Funktionen $g : [0,1] \to \mathbb{R}$ mit der punktweisen Addition und Skalarmultiplikation. Dann definiert das Integral eine Linearform

$$V \;\longrightarrow\; \mathbb{R}, \quad g \;\mapsto\; \int_0^1 g(x)dx,$$

denn es ist

$$\int_0^1 (g+h)(x)dx = \int_0^1 g(x)dx + \int_0^1 h(x)dx,$$
$$\int_0^1 (\alpha \cdot g)(x)dx = \alpha \cdot \int_0^1 g(x)dx.$$

für alle Funktionen $g,h \in C([0,1])$ und alle $\alpha \in \mathbb{R}$.

Beispiel 1.5. Lineare Abbildungen kann man für die Beschreibung geometrischer Transformationen benutzen:

a) Die Abbildung

$$f\colon \quad \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x \\ y \end{pmatrix}$$

ist linear. Sie beschreibt eine Spiegelung an der y-Achse in der Ebene.

b) Für $c > 0$ ist

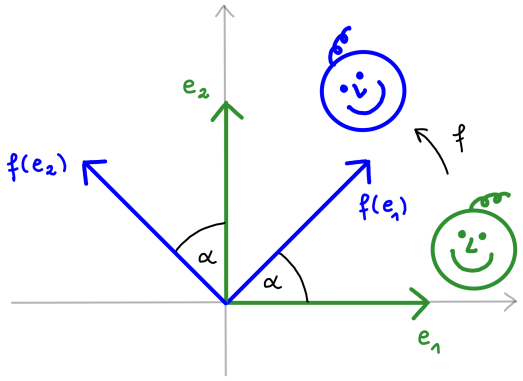
$$f\colon \quad \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ c \cdot y \end{pmatrix}$$

eine lineare Abbildung, eine Streckung entlang der y-Achse mit Streckfaktor c.

c) Für $\alpha \in \mathbb{R}$ ist

$$f\colon \quad \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x\cos\alpha - y\sin\alpha \\ x\sin\alpha + y\cos\alpha \end{pmatrix} = x\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} + y\begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$$

eine lineare Abbildung. Es handelt sich hier um eine Drehung mit Drehwinkel α um den Ursprung, wie man leicht durch Betrachten der Bilder $f(e_1)$ und $f(e_2)$ der Standardbasisvektoren e_1 und e_2 sieht:



Mit linearen Abbildungen kann man sehr einfach rechnen:

Lemma 1.6. *Seien V,W zwei K-Vektorräume und $f \in \mathrm{Hom}_K(V,W)$, dann gilt:*

a) Es ist $f(0) = 0$ und $f(-v) = -f(v)$.

b) Für $v_1,\dots,v_n \in V$ und $\alpha_1,\dots,\alpha_n \in K$ ist

$$f(\alpha_1 v_1 + \cdots + \alpha_n v_n) \;=\; \alpha_1 f(v_1) + \cdots + \alpha_n f(v_n).$$

Beweis. Die Eigenschaft $a)$ gilt für jeden Homomorphismus f additiver Gruppen, und $b)$ erhält man direkt aus der Definition von K-Linearität. □

Das folgende Kriterium spart Schreibarbeit, wenn wir prüfen wollen, ob eine gegebene Abbildung linear ist:

Lemma 1.7. *Für K-Vektorräume V,W und Abbildungen $f : V \to W$ sind äquivalent:*

a) f ist K-linear.

b) Für alle $u,v \in V$ und $\alpha \in K$ gilt: $f(u+\alpha v) \;=\; f(u) + \alpha \cdot f(v)$.

Beweis. Aus $a)$ folgt sofort $b)$. Umgekehrt folgt aus $b)$ auch

$$\begin{aligned} f(u+v) \;&=\; f(u)+f(v) && \text{durch Wahl von } \alpha = 1, \\ f(\alpha \cdot v) \;&=\; \alpha \cdot f(v) && \text{durch Wahl von } u = 0, \end{aligned}$$

denn für $u = 0$ ist $f(u) = 0$ nach Bemerkung 1.6. □

Prüfen wir damit beispielsweise nach, dass die Verkettung linearer Abbildungen ebenfalls eine lineare Abbildung ist:

Korollar 1.8. *Seien U,V,W Vektorräume über K. Dann ist*

$$g \circ f \;\in\; \mathrm{Hom}_K(U,W) \quad \textit{für alle} \quad f \in \mathrm{Hom}_K(U,V),\; g \in \mathrm{Hom}_K(V,W).$$

Beweis. Für $u,v \in U$ und $\alpha \in K$ gilt:

$$\begin{aligned} (g \circ f)(u+\alpha v) \;&=\; g(f(u+\alpha v)) && \text{via } \circ \\ &=\; g(f(u) + \alpha f(v)) && \text{da } f \text{ linear} \\ &=\; g(f(u)) + \alpha g(f(v)) && \text{da } g \text{ linear} \\ &=\; (g \circ f)(u) + \alpha \cdot (g \circ f)(v) && \text{via } \circ \end{aligned}$$

Nach dem Linearitätskriterium 1.7 folgt die Behauptung. □

Wir nennen zwei Vektorräume zueinander *isomorph*, wenn ein Isomorphismus zwischen ihnen existiert. In diesem Fall haben die zwei Vektorräume abstrakt die gleiche Struktur, beispielsweise gilt:

Lemma 1.9. *Sei $f : V \xrightarrow{\sim} W$ ein Isomorphismus. Dann ist auch $f^{-1} : W \xrightarrow{\sim} V$ ein Isomorphismus, und für Familien von Vektoren $v_i \in V$ sind äquivalent:*

a) Die Familie $(v_i)_{i\in I}$ bildet eine Basis von V,

b) Die Familie $(f(v_i))_{i\in I}$ bildet eine Basis von W.

Insbesondere gilt dann $\dim_K(V) = \dim_K(W)$.

Beweis. Mit f ist auch f^{-1} ein Homomorphismus: Denn für $u, v \in W$, $\alpha \in K$ gilt

$$\begin{aligned} f\big(f^{-1}(u) + \alpha \cdot f^{-1}(v)\big) &= f(f^{-1}(u)) + \alpha \cdot f(f^{-1}(v)) && \text{da } f \text{ linear ist,} \\ &= u + \alpha \cdot v && \text{da } f \circ f^{-1} = id_W \end{aligned}$$

Indem wir f^{-1} anwenden, folgt $f^{-1}(u) + \alpha \cdot f^{-1}(v) = f^{-1}(u + \alpha v)$. Also ist f^{-1} linear nach dem Kriterium in Lemma 1.7. Zu zeigen bleibt die Äquivalenz von $a)$ und $b)$. Die K-Linearität von f gibt $f(\langle v_i \mid i \in I\rangle) = \langle f(v_i) \mid i \in I\rangle$, und falls f surjektiv ist, erhalten wir die Implikation:

$$(v_i)_{i\in I} \text{ Erzeugendensystem von } V \quad \Longrightarrow \quad (f(v_i))_{i\in I} \text{ Erzeugendensystem von } W$$

Falls f ein Isomorphismus ist, so auch f^{-1} und dann gilt per Symmetrie auch die Umkehrung der obigen Implikation. In diesem Fall gilt dasselbe für Basen, da diese genau die minimalen Erzeugendensysteme sind. □

Für die Konstruktion von linearen Abbildungen zwischen Vektorräumen ist das folgende Lemma sehr hilfreich:

Lemma 1.10. *Sei V ein Vektorraum über K und $\mathscr{B} = (v_1, \ldots, v_n)$ eine Familie von Vektoren darin. Dann ist*

$$\Phi_{\mathscr{B}} : \quad K^n \longrightarrow V, \quad (\alpha_1, \ldots, \alpha_n) \mapsto \sum_{i=1}^{n} \alpha_i v_i$$

ein Homomorphismus von K-Vektorräumen.

Beweis. Schreibe $I = \{1, \ldots, n\}$. Seien $\alpha \in K$ und $u = (\alpha_i)_{i\in I}, v = (\beta_i)_{i\in I} \in K^n$ gegeben. Dann gilt

$$u + \alpha v = (\alpha_i + \alpha\beta_i)_{i\in I} \quad \text{in} \quad K^n.$$

Daher gilt

$$\begin{aligned} \Phi_{\mathscr{B}}(u + \alpha v) = \Phi_{\mathscr{B}}((\alpha_i + \alpha\beta_i)_{i\in I}) &= \sum_{i\in I} (\alpha_i + \alpha\beta_i) \cdot v_i \\ &= \sum_{i\in I} \alpha_i v_i + \alpha \cdot \sum_{i\in I} \beta_i v_i \\ &= \Phi_{\mathscr{B}}(u) + \alpha \cdot \Phi_{\mathscr{B}}(v) \end{aligned}$$

Nach dem Lemma 1.7 zeigt dies, dass $\Phi_{\mathscr{B}}$ eine K-lineare Abbildung ist. □

Beispiel 1.11. Im reellen Vektorraum

$$V := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x+y+z=0 \right\}$$

betrachte man die beiden Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Für $\mathscr{B} = (v_1, v_2)$ erhalten wir die lineare Abbildung

$$\Phi_{\mathscr{B}}: \quad \mathbb{R}^2 \longrightarrow V, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ -x-y \end{pmatrix}.$$

Diese Abbildung ist sogar ein Isomorphismus, wie man leicht nachrechnet.

Dieses Beispiel verallgemeinert sich zu folgender Aussage, die zeigt, dass jeder endlich erzeugten Vektorraum isomorph ist zu einem Standardvektorraum:

Satz 1.12. *Sei V ein Vektorraum über K, und sei $\mathscr{B} = (v_1, \ldots, v_n)$ eine Familie von Vektoren darin. Dann sind äquivalent:*

a) $\mathscr{B}$ ist eine Basis von V über K.

b) Die Abbildung

$$\Phi_{\mathscr{B}}: \quad K^n \longrightarrow V, \quad (\alpha_1, \ldots, \alpha_n) \mapsto \sum_{i=1}^{n} \alpha_i v_i \quad \textit{ist ein Isomorphismus.}$$

Beweis. Nach dem Lemma 1.10 ist $\Phi_{\mathscr{B}}$ ein Homomorphismus, es geht also nur um die Frage, wann dieser Homomorphismus bijektiv ist. Dies ist genau dann der Fall, wenn jedes $v \in V$ sich als Linearkombination $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ mit eindeutigen Koeffizienten $\alpha_1, \ldots, \alpha_n \in K$ schreibt, also genau dann, wenn B eine Basis ist. □

Um eine lineare Abbildung zu definieren, genügt es, ihre Werte auf einer Basis vorzugeben, und diese Werte können beliebig gewählt werden:

Korollar 1.13. *Seien V und W Vektorräume über K. Gegeben seien*

- *eine Basis $\mathscr{B} = (v_1, \ldots, v_n)$ von V,*
- *eine Familie $\mathscr{C} = (w_1, \ldots, w_n)$ von Vektoren in W.*

Dann gibt es genau ein $g \in \mathrm{Hom}_K(V, W)$ mit $g(v_i) = w_i$ für alle $i \in \{1, \ldots, n\}$.

Beweis. Die Eindeutigkeit von g ist klar, für $v = \sum_{i=1}^n a_i v_i$ muß per Linearität gelten:

$$g(v) = g\left(\sum_{i=1}^n a_i v_i\right) = \sum_{i=1}^n a_i g(v_i) = \sum_{i=1}^n a_i w_i.$$

Für die Existenz kann man benutzen, dass nach dem Struktursatz $\Phi_{\mathscr{B}} : K^n \to V$ ein Isomorphismus ist: Die Abbildung $g = \Phi_{\mathscr{C}} \circ \Phi_{\mathscr{B}}^{-1} : V \to W$ besitzt dann die Eigenschaft $g(v_i) = w_i$ für alle i. □

2 Abbildungsräume und Dualität

Sei W ein K-Vektorraum. Für jede Menge X ist dann die Menge

$$\mathrm{Abb}(X,W) \;=\; \{\text{Abbildungen } f : X \to W\}$$

ein K-Vektorraum mit punktweiser Addition und Skalarmultiplikation. Wenn $X = V$ ebenfalls ein K-Vektorraum ist, gilt:

Lemma 2.1. *Die Teilmenge* $\mathrm{Hom}_K(V,W) \subseteq \mathrm{Abb}(V,W)$ *ist ein Untervektorraum.*

Beweis. Für alle $f,g \in \mathrm{Hom}_K(V,W)$ ist $f+g \in \mathrm{Hom}_K(V,W)$, denn für alle $u,v \in V$ und $\alpha \in K$ ist

$$\begin{aligned}(f+g)(u+\alpha v) &= f(u+\alpha v)+g(u+\alpha v)\\ &= (f(u)+\alpha f(v))+(g(u)+\alpha g(v))\\ &= (f(u)+g(u))+\alpha(f(v)+g(v))\\ &= (f+g)(u)+\alpha\cdot(f+g)(v)\end{aligned}$$

Analog folgt $\alpha \cdot f \in \mathrm{Hom}_K(V,W)$. Die Behauptung folgt also aus Lemma 1.7. □

Beispiel 2.2. Sei V ein K-Vektorraum. Dann ist $\varphi\colon \mathrm{Hom}_K(K,V) \xrightarrow{\sim} V, f \mapsto f(1)$ ein Isomorphismus von Vektorräumen über K:

- φ ist bijektiv: Jedes $v \in V$ ist das Bild $v = \varphi(f)$ von genau einem $f \in \mathrm{Hom}_K(K,V)$, nämlich der linearen Abbildung f mit $f(a) := a \cdot v$.
- φ ist ein Homomorphismus: Für alle $f,g \in \mathrm{Hom}_K(K,V)$, $\alpha \in K$ ist

$$\varphi(f+\alpha g) \;=\; (f+\alpha g)(1) \;=\; f(1)+\alpha g(1) \;=\; \varphi(f)+\alpha\varphi(g).$$

Interessanter wird es, wenn wir V als Definitionsbereich wählen:

Definition 2.3. Sei V ein Vektorraum über K. Der *Dualraum* von V ist definiert als der Vektorraum

$$V^* \;:=\; \mathrm{Hom}_K(V,K).$$

Seine Elemente sind die linearen Abbildungen $f\colon V \to K$, man nennt solche linearen Abbildungen von einem Vektorraum in den Grundkörper auch *Linearformen*.

Die Möglichkeit, Linearformen zu addieren und mit Skalaren zu multiplizieren, spielt in vielen Anwendungen eine Rolle:

Beispiel 2.4 (Numerische Integration). Sei $V = \{f\colon [0,1] \to \mathbb{R} \text{ stetige Funktion}\}$ der reelle Vektorraum aller stetigen Funktionen auf dem Einheitsintervall, versehen mit der punktweisen Vektorraumstruktur. Für jeden festen Wert $t \in [0,1]$ ist dann die Abbildung

$$\mathrm{ev}_t\colon \quad V \longrightarrow \mathbb{R}, \quad f \mapsto f(t)$$

eine Linearform, also ein Vektor im Dualraum $V^* = \mathrm{Hom}_{\mathbb{R}}(V,\mathbb{R})$. Wir können nun weitere Linearformen durch Linearkombinationen solcher Vektoren im Dualraum bilden: Z.B. entspricht

$$\tfrac{1}{2}\cdot(\mathrm{ev}_0 + \mathrm{ev}_1) \in V^*$$

der Linearform

$$V \longrightarrow \mathbb{R}, \quad f \mapsto \tfrac{1}{2}(f(0)+f(1)),$$

die für die näherungsweise Berechnung des Integrals von Funktionen verwendet werden kann:

$$\int_0^1 f(x)dx \approx \tfrac{1}{2}(f(0)+f(1)).$$

Es gibt viele andere Näherungsformeln, die Integrale durch Linearkombinationen von Funktionswerten an verschiedenen Stellen $t_1,\ldots,t_n \in [0,1]$ approximieren. Eine Näherungsformel

$$\int_0^1 f(x)dx \approx \sum_{i=1}^{n} \alpha_i \cdot f(t_i)$$

mit $\alpha_i \in \mathbb{R}$ entspricht dabei dem Vektor $\alpha_1 \cdot \mathrm{ev}_{t_1} + \cdots + \alpha_n \cdot \mathrm{ev}_{t_n}$ im Dualraum.

Jeder linearen Abbildung $g\colon U \to V$ von Vektorräumen kann man eine lineare Abbildung $g^*\colon V^* \to U^*$ in der umgekehrten Richtung zwischen ihren Dualräumen zuordnen. Dies ist der Spezialfall $W = K$ der folgenden Konstruktion:

Lemma 2.5. *Seien U,V,W Vektorräume über K. Für $g \in \mathrm{Hom}_K(U,V)$ gilt:*

a) Die Abbildungen

$$g^*\colon \quad \mathrm{Hom}_K(V,W) \longrightarrow \mathrm{Hom}_K(U,W), \quad f \mapsto f\circ g$$

$$g_*\colon \quad \mathrm{Hom}_K(W,U) \longrightarrow \mathrm{Hom}_K(W,V), \quad f \mapsto g\circ f$$

sind K-linear bezüglich der punktweisen Vektorraumstruktur.

b) Ist g ein Isomorphismus, dann sind auch g_* und g^* Isomorphismen.

Beweis. *a*) Die Abbildung g^* ist K-linear: Seien $f_1,f_2 \in \mathrm{Hom}_K(V,W)$, $\alpha \in K$. Zu zeigen ist dann

$$g^*(f_1+\alpha f_2) = g^*(f_1) + \alpha g^*(f_2).$$

Dies folgt punktweise durch Auswerten der linearen Abbildungen: Für alle $u \in U$ gilt

$$\begin{aligned}(g^*(f_1+\alpha f_2))(u) &= ((f_1+\alpha f_2)\circ g)(u) &&= (f_1+\alpha f_2)(g(u))\\ &&&= f_1(g(u)) + \alpha f_2(g(u))\\ &&&= (g^*(f_1))(u) + (\alpha g^*(f_2))(u)\\ &&&= (g^*(f_1) + \alpha g^*(f_2))(u)\end{aligned}$$

und somit $g^*(f_1+\alpha f_2) = g^*(f_1) + \alpha g^*(f_2)$. Die K-Linearität von g_* folgt analog.

b) Ist $g : U \to V$ ein Isomorphismus und $h = g^{-1} : V \to U$ sein Inverses, dann gilt für alle Homomorphismen $f \in \mathrm{Hom}_K(V,W)$:

$$h^*(g^*(f)) \;=\; ((f \circ g) \circ h) \;=\; (f \circ (g \circ h)) \;=\; f \circ id \;=\; f.$$

Ebenso sieht man $g^*(h^*(f)) = f$ für alle $f \in \mathrm{Hom}_K(U,W)$. Also ist g^* invertierbar mit h^* als Inversem. Die Aussage für g_* folgt analog. □

Neben Dualräumen sind ein weiteres wichtiges Beispiel für Vektorräume von linearen Abbildungen solche, bei denen der Definitions- und Zielraum gleich sind:

Definition 2.6. Ein *Endomorphismus* eines Vektorraumes V über K ist eine lineare Abbildung $f : \quad V \longrightarrow V$ des Vektorraumes in sich. Bijektive Endomorphismen heißen *Automorphismus* des Vektorraumes. Wir schreiben kurz

$$\begin{aligned} \mathrm{End}_K(V) \;&=\; \mathrm{Hom}_K(V,V), \\ \mathrm{Aut}_K(V) \;&=\; \{f : V \to V \mid f \text{ ist ein Automorphismus}\}. \end{aligned}$$

Bezüglich der Verkettung von Abbildungen ist $\mathrm{Aut}_K(V)$ eine Gruppe. Ebenso ist $\mathrm{End}_K(V)$ nicht nur ein Vektorraum, sondern auch ein Ring mit punktweiser Addition und der Verkettung als Multiplikation:

$$\begin{aligned} (f+g)(v) \;&=\; f(v)+g(v), \\ (f \circ g)(v) \;&=\; f(g(v)). \end{aligned}$$

Genauer ist $\mathrm{End}_K(V)$ ein Beispiel für eine sogenannte K-Algebra:

Definition 2.7. Eine *K-Algebra* ist ein Ring $(R,+,\circ)$, der zugleich ein Vektorraum über K ist, sodass die Multiplikation des Ringes mit der Skalarmultiplikation des Vektorraumes verträglich ist:

$$\forall a \in K \;\; \forall f,g \in R : \quad a(f \circ g) \;=\; (af) \circ g \;=\; f \circ (ag)$$

Im nächsten Abschnitt werden wir mittels von Matrizen eine explizite Beschreibung der K-Algebra $\mathrm{End}_K(V)$ für $V = K^n$ bekommen, mit der sich z.B. die Verkettung von Drehungen oder Spiegelungen in der Ebene sehr einfach ausrechnen lässt.

3 Von linearen Abbildungen zu Matrizen

Wir haben bereits viele Beispiele linearer Abbildungen gesehen. Z.B. haben wir Drehungen um den Winkel φ um den Ursprung beschrieben durch Abbildungen der Form

$$f : \quad \mathbb{R}^2 \;\longrightarrow\; \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;\mapsto\; \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$$

mit den Koeffizienten $a_{11} = a_{22} = \cos\varphi$ und $a_{21} = -a_{12} = \sin\varphi$. Allgemein gilt:

Lemma 3.1. *Die K-linearen Abbildungen $f : K^n \to K^m$ sind genau die Abbildungen der Form*

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

mit Koeffizienten $a_{ij} \in K$ für $i = 1,\ldots,m$ und $j = 1,\ldots,n$.

Beweis. Dass jede Abbildung der gegebenen Form K-linear ist, rechnet man direkt nach. Umgekehrt sei $f : K^n \longrightarrow K^m$ eine beliebige K-lineare Abbildung. Wir müssen zeigen, dass sich diese in der angegebenen Form schreiben lässt. Um die a_{ij} zu erraten, beachte man: Wenn f die angegebene Form besitzt, so haben die Bilder der Standardbasisvektoren $e_1,\ldots,e_n \in K^n$ die Form

$$f(e_j) \;=\; \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

Wir drehen den Spieß um und *definieren* für gegebenes f Koeffizienten $a_{ij} \in K$ durch

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \;:=\; f(e_j) \;\in\; K^m \quad \text{für} \quad j \;=\; 1,\ldots,n.$$

Sei $g : K^n \longrightarrow K^m$ die lineare Abbildung definiert durch

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

Per Konstruktion gilt $f(e_j) = g(e_j)$ für $j = 1,\ldots,n$, und wegen $K^n = \langle e_1,\ldots,e_n\rangle$ folgt dann $f = g$. □

Um vernünftig mit linearen Abbildungen arbeiten zu können, benötigen wir eine weniger umständliche Notation. Dazu führen wir den Begriff einer Matrix ein:

Definition 3.2. Eine *Matrix* der Größe $m \times n$ über K ist ein rechteckiges Schema mit Einträgen $a_{ij} \in K$, welches m Zeilen und n Spalten umfasst:

$$M \;=\; \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Wir bezeichnen

- den Index $i \in \{1,2,\ldots,m\}$ als *Zeilenindex*,
- den Index $j \in \{1,2,\ldots,n\}$ als *Spaltenindex*.

Wir bezeichnen im Folgenden mit $\mathrm{Mat}(m \times n, K) = K^{m\times n}$ die Menge aller Matrizen vom Format $m \times n$ mit Einträgen in K. Für solche Matrizen verwenden wir auch die Kurznotation

$$M \;=\; (a_{ij}) \;\in\; \mathrm{Mat}(m \times n, K).$$

Beipielsweise ist

- $\mathrm{Mat}(m \times 1, K) \;=\; K^m$ die Menge der *Spaltenvektoren*,
- $\mathrm{Mat}(1 \times n, K) \;=\; K^n$ die Menge der *Zeilenvektoren*.

Für einen Zeilenvektor $a = (a_1, \ldots, a_n) \in \mathrm{Mat}(1 \times n, K)$ und einen ebenso langen Spaltenvektor

$$x \;=\; \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;\in\; \mathrm{Mat}(n \times 1, K),$$

schreiben wir kurz

$$a \cdot x \;:=\; \sum_{i=1}^{n} a_i x_i \;\in\; K.$$

Allgemeiner sei

$$M \;=\; \begin{pmatrix} — & a_1 & — \\ & \vdots & \\ — & a_i & — \\ & \vdots & \\ — & a_m & — \end{pmatrix} \;\in\; \mathrm{Mat}(m \times n, K)$$

mit Zeilen $a_i = (a_{i1}, \ldots, a_{in})$. Für Spaltenvektoren $x \in K^n = \mathrm{Mat}(n \times 1, K)$ setzen wir

$$M \cdot x \;:=\; \begin{pmatrix} a_1 \cdot x \\ \vdots \\ a_m \cdot x \end{pmatrix} \;\in\; K^m \;=\; \mathrm{Mat}(m \times 1, K),$$

mit $a_i \cdot x \in K$ wie zuvor definiert. Explizit sieht dieses Produkt einer $m \times n$ Matrix mit einem Spaltenvektor der Länge n so aus:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \;=\; \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}$$

Lemma 3.3. *Jede lineare Abbildung zwischen Standard-Vektorräumen hat die Form*

$$f: \quad K^n \longrightarrow K^m, \quad v \mapsto M \cdot v$$

für eine eindeutig bestimmte Matrix $M \in \mathrm{Mat}(m \times n, K)$*. Wir erhalten somit eine Bijektion*

$$\varphi: \quad \mathrm{Mat}(m \times n, K) \xrightarrow{\sim} \mathrm{Hom}_K(K^n, K^m), \quad M \mapsto M \cdot (-).$$

Beweis. Nach Lemma 3.1 bleibt nur noch die Eindeutigkeit der zu einer gegebenen linearen Abbildung $f: K^n \longrightarrow K^m$ gehörigen Matrix $M = (a_{ij})$ zu zeigen. Die ist aber nichts Neues: Die j-te Spalte jeder solchen Matrix ist gegeben durch

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

also gleich dem Spaltenvektor $f(e_j)$. □

Wir halten aus dem Beweis als Slogan fest: Die Spalten einer Matrix sind die Bilder der Standardbasisvektoren unter der zugehörigen linearen Abbildung! Damit kann man die Matrix zu einer linearen Abbildung direkt ablesen:

Beispiel 3.4. Eine Drehung um einen Winkel $\alpha \in \mathbb{R}$ in der reellen Ebene ist gegeben durch Multiplikation mit der Matrix

$$M = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \in \mathrm{Mat}(2 \times 2, \mathbb{R}).$$

Denn die Spalten dieser Matrix sind genau die Vektoren, die man durch Drehung der Standardbasisvektoren um den Winkel α erhält.

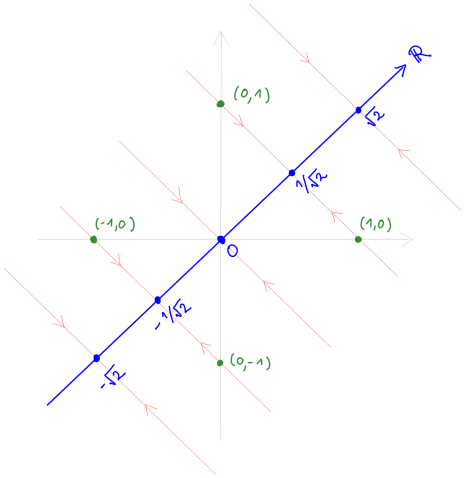
Beispiel 3.5. Als nächstes Beispiel betrachten wir die orthogonale Projektion von der reellen Ebene auf die Diagonale: Diese Projektion ist die lineare Abbildung

$$f: \quad \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \tfrac{1}{\sqrt{2}} \cdot (x_1 + x_2)$$

gegeben durch die Multiplikation mit der Matrix

$$M = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \in \mathrm{Mat}(1 \times 2, \mathbb{R}),$$

denn aus dem folgenden Bild lesen wir unmittelbar $f(e_1) = f(e_2) = \frac{1}{\sqrt{2}}$ ab:



Bisher haben wir nur mit Standard-Vektorräumen $V = K^n$ gearbeitet, also mit Spaltenvektoren. Um in beliebigen endlich erzeugten Vektorräumen ebenso rechnen zu können, müssen wir Basen wählen: Jede Basis $\mathscr{B} = (u_1, \ldots, u_n)$ eines endlich erzeugten Vektorraumes U über K liefert einen Isomorphismus

$$\Phi_{\mathscr{B}} : \quad K^n \xrightarrow{\sim} U, \quad (a_1, \ldots, a_n) \mapsto \sum_{i=1}^{n} a_i u_i,$$

der U mit einem Standard-Vektorraum identifiziert. Um lineare Abbildungen von beliebigen endlich erzeugten Vektorräumen durch Matrizen zu beschreiben, müssen wir sowohl im Definitions- als auch im Zielraum eine Basis wählen:

Definition 3.6. Es sei

- U ein K-Vektorraum mit einer Basis $\mathscr{B} = (u_1, \ldots, u_n)$,
- V ein K-Vektorraum mit einer Basis $\mathscr{C} = (v_1, \ldots, v_m)$.

Für $f \in \mathrm{Hom}_K(U,V)$ ist $\Phi_{\mathscr{C}}^{-1} \circ f \circ \Phi_{\mathscr{B}} \in \mathrm{Hom}_K(K^n, K^m)$. Die zugehörige Matrix bezeichnen wir mit

$$_{\mathscr{C}}f_{\mathscr{B}} = M_{\mathscr{B},\mathscr{C}}(f) \in \mathrm{Mat}(m \times n, K),$$

sie heißt die *Abbildungsmatrix von f in den Basen $\mathscr{B}, \mathscr{C}$*. Im folgenden Diagramm liefert also die Verkettung entlang beider möglichen Pfade von links unten nach rechts oben dieselbe Abbildung (wir sagen auch, das Diagramm sei *kommutativ*):

$$\begin{array}{ccc} U & \xrightarrow{\quad f \quad} & V \\ {\scriptstyle \Phi_{\mathscr{B}}}\uparrow & & \uparrow{\scriptstyle \Phi_{\mathscr{C}}} \\ K^n & \xrightarrow[\quad {}_{\mathscr{C}}f_{\mathscr{B}} \quad]{} & K^m \end{array}$$

Die Koeffizienten der Abbildungsmatrix ${}_{\mathscr{C}}f_{\mathscr{B}} = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ zu den gewählten Basen sind durch

$$f(u_j) = \sum_{i=1}^{m} a_{ij} v_i$$

gegeben. Im Fall einer bereits durch Multiplikation mit einer Matrix M gegebenen Abbildung

$$f\colon \quad U = K^n \longrightarrow V = K^m, \quad u \mapsto M \cdot u$$

von Standardvektorräumen erhalten wir bezüglich der Standardbasen $\mathscr{B}$ und $\mathscr{C}$ die gegebene Matrix ${}_{\mathscr{C}}f_{\mathscr{B}} = M$ zurück, aber bezüglich anderer Basen erhalten wir auch andere Matrizen:

Beispiel 3.7. Sei

$$f\colon \quad U = \mathbb{R}^2 \longrightarrow V = \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2x+y \\ 3x+y \\ 4x+y \end{pmatrix}.$$

Für die Standardbasen $\mathscr{B} = (e_1^U, e_2^U)$ von $U = \mathbb{R}^2$ und $\mathscr{C} = (e_1^V, e_2^V, e_3^V)$ von $V = \mathbb{R}^3$ erhalten wir

$${}_{\mathscr{C}}f_{\mathscr{B}} = \begin{pmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix}, \quad \text{denn} \quad \begin{cases} f(e_1^U) = 2 \cdot e_1^V + 3 \cdot e_2^V + 4 \cdot e_3^V, \\ f(e_2^U) = 1 \cdot e_1^V + 1 \cdot e_2^V + 1 \cdot e_3^V. \end{cases}$$

Wenn wir stattdessen die Basen $\mathscr{B}' = (u_1, u_2)$ von U und $\mathscr{C}' = (v_1, v_2, v_3)$ von V verwenden mit

$$u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

erhalten wir

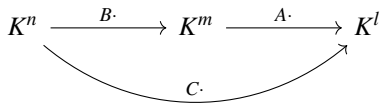
$${}_{\mathscr{C}'}f_{\mathscr{B}'} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 3 & 2 \end{pmatrix}, \quad \text{denn} \quad \begin{cases} f(u_1) = 2 \cdot v_1 + 1 \cdot v_2 + 3 \cdot v_3, \\ f(u_2) = 1 \cdot v_1 + 0 \cdot v_2 + 2 \cdot v_3. \end{cases}$$

4 Das Matrizenprodukt

Wir haben gesehen, dass die Verkettung linearer Abbildungen wieder linear ist. Für Matrizen $A \in \mathrm{Mat}(l \times m, K)$ und $B \in \mathrm{Mat}(m \times n, K)$ mit den zugehörigen linearen Abbildungen

$$g = \varphi(B)\colon \quad K^n \longrightarrow K^m \quad \text{und} \quad f = \varphi(A)\colon \quad K^m \longrightarrow K^l$$

gibt es somit nach Lemma 3.3 genau eine Matrix $C \in \mathrm{Mat}(l \times n, K)$ mit $f \circ g = \varphi(C)$ wie im folgenden kommutativen Diagramm angedeutet:



Um die Matrix C zu berechnen, erinnern wir uns an den Slogan, dass die Spalten einer Matrix die Bilder der Standardbasisvektoren unter der zugehörigen linearen Abbildung sind. Für $k \leq n$ bezeichne

- $w_k = C \cdot e_k \in K^l$ die k-te Spalte von C,
- $v_k = B \cdot e_k \in K^m$ die k-te Spalte von B,

dann folgt

$$\begin{aligned} w_k &= C \cdot e_k && \text{per Definition von } w_k \\ &= (f \circ g)(e_k) && \text{wegen } f \circ g = \varphi(C) \\ &= f(g(e_k)) && \text{per Definition von } \circ \\ &= f(B \cdot e_k) && \text{wegen } g = \varphi(B) \\ &= f(v_k) && \text{per Definition von } v_k \\ &= A \cdot v_k && \text{wegen } f = \varphi(A) \end{aligned}$$

Wir halten fest: Die Spalten der gesuchten Matrix C können wir als Spaltenvektoren bekommen, indem wir einfach das Produkt der Matrix A mit den Spaltenvektoren von B ausrechnen. Sei $A = (a_{ij})$, $B = (b_{jk})$, $C = (c_{ik})$, dann ist also

$$\begin{pmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{lk} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l1} & a_{l2} & \cdots & a_{lm} \end{pmatrix} \cdot \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{mk} \end{pmatrix} \quad \text{oder kurz} \quad c_{ik} = \sum_{j=1}^{m} a_{ij} \cdot b_{jk}.$$

Dies führt auf die folgende Definition:

Definition 4.1. Für Matrizen $A = (a_{ij}) \in \mathrm{Mat}(l \times m, K)$, $B = (b_{jk}) \in \mathrm{Mat}(m \times n, K)$ definieren wir das *Matrizenprodukt*

$$A \cdot B := (c_{ik}) \in \mathrm{Mat}(l \times n, , K) \quad \text{durch} \quad c_{ik} := \sum_{j=1}^{m} a_{ij} \cdot b_{jk}.$$

Schematisch kann man die Produktmatrix $C = A \cdot B$ nach dem folgenden Muster berechnen, wobei aus der i-ten Zeile a_i von A und der k-ten Spalte b_k von B der Eintrag

$$c_{ik} = a_i \cdot b_k$$

gebildet wird:

$$\begin{pmatrix} — & a_1 & — \\ & \vdots & \\ — & a_i & — \\ & \vdots & \\ — & a_l & — \end{pmatrix} \cdot \begin{pmatrix} | & & | & & | \\ b_1 & \cdots & b_k & \cdots & b_n \\ | & & | & & | \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1k} & \cdots & c_{1n} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \cdots & c_{ik} & \cdots & c_{in} \\ \vdots & & \vdots & & \vdots \\ c_{l1} & \cdots & c_{lk} & \cdots & c_{lk} \end{pmatrix}$$

Für $a,b,c,x,y,z \in K$ gilt beispielsweise

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} x & a \\ y & b \\ z & c \end{pmatrix} = \begin{pmatrix} x+2y+3z & a+2b+3c \\ 4x+5y+6z & 4a+5b+6c \end{pmatrix} \in \mathrm{Mat}(2\times 2, K).$$

Aus der Diskussion vor unserer Definition des Matrizenproduktes folgt, dass für Matrizen

$$A \in \mathrm{Mat}(l\times m, K) \iff \varphi(A) \in \mathrm{Hom}_K(K^m, K^l),$$
$$B \in \mathrm{Mat}(m\times n, K) \iff \varphi(B) \in \mathrm{Hom}_K(K^n, K^m),$$

gilt:

$$\varphi(A\cdot B) = \varphi(A)\circ\varphi(B) \in \mathrm{Hom}_K(K^n, K^l).$$

Damit kann man die Verkettung von linearen Abbildungen sehr leicht ausrechnen:

Beispiel 4.2. Für Drehmatrizen

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$$

bekommen wir aus der Definition des Matrizenproduktes

$$A\cdot B = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \quad \text{mit} \quad \begin{cases} c = \cos\alpha\cos\beta - \sin\alpha\sin\beta, \\ s = \sin\alpha\cos\beta + \cos\alpha\sin\beta. \end{cases}$$

Anschaulich ist die Verkettung von $\varphi(B)\circ\varphi(A)$ eine Drehung um den Winkel $\alpha+\beta$, wir erhalten somit die Additionstheoreme für sinus und cosinus:

$$\begin{aligned} \cos\alpha\cos\beta - \sin\alpha\sin\beta &= \cos(\alpha+\beta), \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta &= \sin(\alpha+\beta). \end{aligned}$$

Beispiel 4.3. Die Projektion $f : \mathbb{R}^2 \to \mathbb{R}, (x_1,x_2) \mapsto x_1$ ist die Multiplikation mit der Matrix

$$A = \begin{pmatrix} 1 & 0 \end{pmatrix} \in \mathrm{Mat}(1\times 2, \mathbb{R}).$$

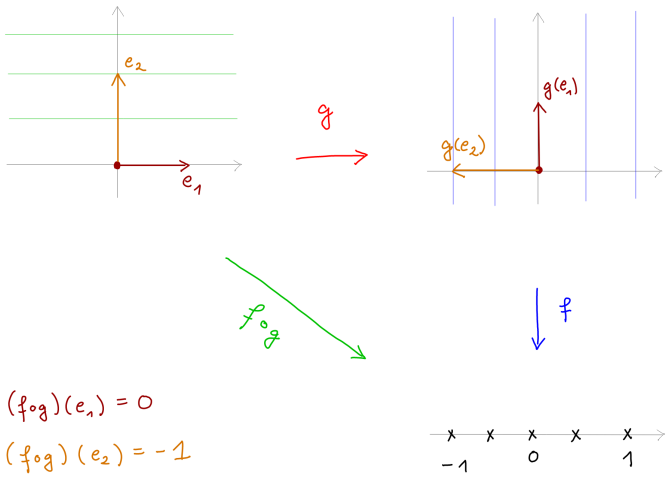
Sei ferner $g : \mathbb{R}^2 \to \mathbb{R}^2$ die Drehung um den Ursprung um den Winkel $\frac{\pi}{2}$, gegeben durch Multiplikation mit

$$B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{Mat}(2 \times 2, \mathbb{R}).$$

Die Verkettung $f \circ g$ ist dann die Multiplikation mit

$$A \cdot B = \begin{pmatrix} 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \end{pmatrix} \in \mathrm{Mat}(1 \times 2, \mathbb{R}),$$

also gegeben durch $(x, y) \mapsto -y$ wie in der folgenden Skizze gezeigt:



Wir können Matrizen vom gleichen Format auch komponentenweise zueinander addieren und mit Skalaren multiplizieren: Für $A = (a_{ij}), B = (b_{ij}) \in \mathrm{Mat}(m \times n, K)$ und $\alpha \in K$ definieren wir

$$\begin{aligned} A + B &:= (a_{ij} + b_{ij}) \in \mathrm{Mat}(m \times n, K), \\ \alpha \cdot A &:= (\alpha \cdot a_{ij}) \in \mathrm{Mat}(m \times n, K). \end{aligned}$$

Damit wird $\mathrm{Mat}(m \times n, K)$ ein Vektorraum über K, und diese Struktur ist kompatibel mit dem Matrizenprodukt; genauer gelten die folgenden Rechenregeln:

Lemma 4.4. *Für $A \in \mathrm{Mat}(k \times l, K)$, $B, B' \in \mathrm{Mat}(l \times m, K)$, $C \in \mathrm{Mat}(m \times n, K)$ gilt:*

a) Assoziativgesetz: Es ist $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

b) Distributivgesetz: $A \cdot (B + B') = A \cdot B + A \cdot B'$ und $(B + B') \cdot C = B \cdot C + B' \cdot C$.

c) Es ist $A \cdot \mathbf{1}_l = A$ und $\mathbf{1}_l \cdot B = B$ für die Einheitsmatrix

$$\mathbf{1}_l := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in \mathrm{Mat}(l \times l, K),$$

die Einsen auf der Diagonalen hat und Nullen überall sonst.

d) Für alle $\alpha \in K$ ist $\alpha \cdot (A \cdot B) = (\alpha \cdot A) \cdot B = A \cdot (\alpha \cdot B)$.

Beweis. Das Assoziativgesetz folgt aus der Assoziativität $(f \circ g) \circ h = f \circ (g \circ h)$ der Verkettung von Abbildungen

$$K^n \xrightarrow{h} K^m \xrightarrow{g} K^l \xrightarrow{f} K^k$$

durch Anwenden der Bijektion $\varphi\colon \mathrm{Mat}(k \times n) \longrightarrow \mathrm{Hom}_K(K^n, K^k)$. Die übrigen Aussagen rechnet man ebenfalls direkt nach. □

Bemerkung 4.5. Für die Menge $\mathrm{Mat}(n \times n, K)$ der *quadratische Matrizen* haben wir drei Verknüpfungen betrachtet:

- Die *Addition*

$$+\colon \quad \mathrm{Mat}(n \times n, K) \times \mathrm{Mat}(n \times n, K) \to K, \quad (A, B) \mapsto A + B.$$

- Die *Skalarmultiplikation*

$$\cdot\colon \quad K \times \mathrm{Mat}(n \times n, K) \to K, \quad (\alpha, B) \mapsto \alpha B.$$

- Die *Matrixmultiplikation*

$$\cdot\colon \quad \mathrm{Mat}(n \times n, K) \times \mathrm{Mat}(n \times n, K) \to K, \quad (A, B) \mapsto A \cdot B.$$

Nach dem obigen Lemma ist die Menge $\mathrm{Mat}(n \times n, K)$ ein K-Vektorraum bezüglich Addition und Skalarmultiplikation, und zugleich ein Ring bezüglich der Addition und Matrixmultiplikation. Nach Eigenschaft $d)$ im obigen Lemma ist $\mathrm{Mat}(n \times n, K)$ somit eine K-Algebra, wir nennen sie die *Matrixalgebra* vom Rang n.

Da quadratische Matrizen einen Ring bilden, können wir insbesondere Potenzen einer quadratischen Matrix $A \in \mathrm{Mat}(n \times n, K)$ bilden. Für $m \in \mathbb{N}$ sind diese rekursiv definiert durch

$$A^0 := \mathbf{1}_n \quad \text{und} \quad A^{k+1} := A \cdot A^k \quad \text{für } k \in \mathbb{N}_0.$$

Die K-Algebra-Struktur erlaubt es uns dann auch, für $\alpha_0,\dots,\alpha_n \in K$ polynomiale Ausdrücke

$$\alpha_d \cdot A^d + \alpha_{d-1} \cdot A^{d-1} + \cdots + \alpha_1 \cdot A + \alpha_0 \cdot \mathbf{1}_n \ \in \ \mathrm{Mat}(n \times n, K)$$

zu bilden. Die Distributivität liefert außerdem Formeln wie

$$(A+B)^2 \ = \ A^2 + AB + BA + B^2 \quad \text{für} \quad A, B \ \in \ \mathrm{Mat}(n \times n, K).$$

Achtung, dabei ist im Allgemeinen $AB \neq BA$, die Matrizenmultiplikation ist nicht kommutativ: Z.B. ist

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

Nur im Fall $AB = BA$ hat man die üblichen binomischen Formeln:

Lemma 4.6. *Seien $A, B \in \mathrm{Mat}(n \times n, K)$ Matrizen mit $AB = BA$. Für alle $m \in \mathbb{N}$ ist dann*

$$(A+B)^m \ = \ \sum_{k=0}^{m} \binom{m}{k} \cdot A^k \cdot B^{m-k} \quad \textit{mit} \quad \binom{m}{k} \ := \ \frac{m!}{k!(n-k)!} \ \in \ \mathbb{N}.$$

Beweis. Der Beweis geht genauso wir in den reellen Zahlen, wenn man beachtet, dass wir wegen $AB = BA$ alle beim Ausmultiplizieren erhaltenen Terme als $A^k B^{m-k}$ mit $k \in \{0, 1, \dots, m\}$ schreiben können. □

5 Mehr über Zeilen und Spalten

Wir haben $\mathrm{Mat}(m \times n, K)$ zu einem Vektorraum gemacht mit komponentenweiser Addition und Skalarmultiplikation: Für $A = (a_{ij}), B = (b_{ij}) \in \mathrm{Mat}(m \times n, K), \alpha \in K$ hatten wir

$$\begin{aligned} A + B \ &:= \ (a_{ij} + b_{ij}) \ \in \ \mathrm{Mat}(m \times n, K) \\ \alpha \cdot A \ &:= \ (\alpha \cdot a_{ij}) \quad \in \ \mathrm{Mat}(m \times n, K) \end{aligned}$$

gesetzt. Das passt zur punktweisen Vektorraumstruktur für lineare Abbildungen:

Lemma 5.1. *Die Abbildung*

$$\varphi\colon \quad \mathrm{Mat}(m \times n, K) \ \longrightarrow \ \mathrm{Hom}_K(K^n, K^m), \quad A \ \mapsto \ A \cdot (-)$$

ist ein Isomorphismus von Vektorräumen über K.

Beweis. Wir wissen bereits, dass φ bijektiv ist. Zu zeigen bleibt nur, dass unter dieser Bijektion die obige komponentenweise Addition und Skalarmultiplikation

von Matrizen genau der punktweisen Addition und Skalarmultiplikation linearer Abbildungen entspricht. Dafür reicht es zu zeigen, dass für alle $A, B \in \mathrm{Mat}(m \times n, K)$ und $\alpha \in K$ die Identität

$$\varphi(A + \alpha \cdot B) \;=\; \varphi(A) + \alpha \cdot \varphi(B) \quad \text{in} \quad \mathrm{Hom}_K(K^n, K^m)$$

gilt, also $(\varphi(A + \alpha \cdot B))(v) = \varphi(A)(v) + \alpha \cdot \varphi(B)(v)$ für alle $v \in K^n$ ist. Das läuft hinaus auf

$$(A + \alpha \cdot B) \cdot v \;=\; A \cdot v + \alpha \cdot (B \cdot v),$$

eine Gleichung von Spaltenvektoren. Prüfen wir für $i = 1, \ldots, n$ den i-ten Eintrag dieser Spaltenvektoren nach:

$$\begin{aligned}
\big((A + \alpha \cdot B) \cdot v\big)_i \;=\; \sum_{j=1}^{n} (A + \alpha \cdot B)_{ij} \cdot v_j \;&=\; \sum_{j=1}^{n} (a_{ij} + \alpha \cdot b_{ij}) \cdot v_j \\
&=\; \sum_{j=1}^{n} a_{ij} \cdot v_j \;+\; \sum_{j=1}^{n} \alpha \cdot b_{ij} \cdot v_j \\
&=\; \big(A \cdot v\big)_i + \alpha \cdot \big(B \cdot v\big)_i \\
&=\; \big(A \cdot v + \alpha \cdot (B \cdot v)\big)_i .
\end{aligned}$$

□

Korollar 5.2. *Für endlich erzeugte K-Vektorräume V, W ist*

$$\dim_K\big(\mathrm{Hom}_K(V, W)\big) \;=\; \dim_K(V) \cdot \dim_K(W).$$

Beweis. Nach Wahl von Basen in V und in W dürfen wir $V = K^m$ und $W = K^n$ mit $m, n \in \mathbb{N}_0$ annehmen. Offenbar ist eine Basis des Vektorraumes $\mathrm{Mat}(m \times n, K)$ gegeben durch die Matrizen

$$E_{ij} \;:=\; \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \quad \begin{matrix} \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \end{matrix}$$

$$\uparrow$$

$$j\text{-te Spalte}$$

für $i = 1, \ldots, m$ und $j = 1, \ldots, n$, man nennt diese auch *Standardmatrizen*. Es gibt genausoviele Standardmatrizen wie Paare (i, j), also genau $m \cdot n$. Somit ist $\dim_K(\mathrm{Hom}_K(K^m, K^n)) = m \cdot n$, und es folgt die Behauptung. □

Der Übergang von einem endlich erzeugten Vektorraum zu seinem Dualraum lässt sich in der Sprache von Matrizen als Übergang von Spalten- zu Zeilenvektoren verstehen:

Beispiel 5.3. Sei $V = K^n$ der Standardvektorraum der Dimension n. Der natürliche Isomorphismus

$$V \;\simeq\; \mathrm{Hom}_K(K,V) \;=\; \mathrm{Mat}(n\times 1, K)$$

identifiziert diesem Standardvektorraum nach unseren Konventionen für Matrizen als Vektorraum von *Spaltenvektoren*. Die Standardmatrizen sind in diesem Fall die Standardbasisvektoren

$$e_i \;:=\; \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \;\leftarrow\; i\text{-te Zeile}$$

für $i = 1,\ldots,n$. Ebenso ist der Dualraum $V^* = \mathrm{Hom}_K(V,K) = \mathrm{Mat}(1\times n, K)$ der Vektorraum von *Zeilenvektoren*, die Standardmatrizen sind hierbei die 'dualen' Standardbasisvektoren

$$\begin{array}{l} e_j^* \;:=\; (0,\ldots,0,1,0,\ldots,0) \\ \qquad\qquad\quad\;\; \uparrow \\ \qquad\qquad\quad\;\; j\text{-te Spalte} \end{array}$$

für $j = 1,\ldots,n$. Insbesondere gilt

$$\dim_K(V^*) \;=\; \dim_K(V) \;=\; n,$$

wie nach dem vorigen Korollar erwartet. Jeder Zeilenvektor $w = (a_1,\ldots,a_n) \in V^*$ entspricht einer Linearform auf dem Vektorraum der Spaltenvektoren mittels dem Matrizenprodukt

$$V^* \times V \;=\; \mathrm{Mat}(1\times n, K) \times \mathrm{Mat}(n\times 1, K) \;\longrightarrow\; \mathrm{Mat}(1\times 1, K) \;=\; K,$$

diese Linearform ist gegeben durch

$$V \;\longrightarrow\; K, \quad v \;=\; \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;\mapsto\; w\cdot v \;=\; (a_1,\ldots,a_n)\cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;=\; a_1x_1 + \cdots a_nx_n.$$

Wenn wir die dualen Standardbasisvektoren e_j^* auf den Standardbasisvektoren e_i auswerten, erhalten wir

$$e_j^* \cdot e_i \;=\; \delta_{ij} \quad \text{für das Kronecker-Delta} \quad \delta_{ij} \;:=\; \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

Für endlich erzeugte Vektorräume gilt nach Wahl einer Basis analog:

Lemma 5.4. *Sei V ein endlich erzeugter K-Vektorraum und $\mathscr{B} = (v_1, \ldots, v_n)$ eine Basis desselben. Für $j = 1, \ldots, n$ definiere eine Linearform $f_j \in V^* = \mathrm{Hom}_K(V, K)$ durch*

$$f_j(v_i) \;=\; \delta_{ij} \quad \textit{für} \quad i \;=\; 1, \ldots, n.$$

Dann ist $\mathscr{B}^* = (f_1, \ldots, f_n)$ eine Basis von V^* (wir nennen sie die zu $\mathscr{B}$ duale Basis*).*

Beweis. Wegen $\dim_K(V^*) = \dim_K(V) = n$ müssen wir nur zeigen, dass $\mathscr{B}$ ein linear unabhängiges System von Vektoren im Dualraum bildet. Seien dazu $a_i \in K$ gegeben mit $a_1 f_1 + \cdots + a_n f_n = 0$. Für alle i ist dann

$$\begin{aligned} a_i \;=\; a_1 \cdot 0 + \cdots + a_i \cdot 1 + \cdots + a_n \cdot 0 \;&=\; a_1 \cdot f_1(e_i) + \cdots + a_n \cdot f_n(e_i) \\ &=\; (a_1 f_1 + \cdots + a_n f_n)(e_i) \;=\; 0 \end{aligned}$$

und es folgt die Behauptung. □

Ebenso wie jeder Vektorraum einen Dualraum besitzt, kann man zu jeder linearen Abbildung eine duale Abbildung definieren: Sei $f\colon U \to V$ eine lineare Abbildung von K-Vektorräumen. Die dazu *duale Abbildung* hatten wir definiert durch

$$f^*\colon \quad V^* \;=\; \mathrm{Hom}_K(V, K) \;\longrightarrow\; U^* \;=\; \mathrm{Hom}_K(U, K), \quad g \;\mapsto\; g \circ f$$

Wie lässt sich die duale Abbildung explizit beschreiben? Wenn wir eine Basis $\mathscr{B}$ von U und eine Basis $\mathscr{C}$ von V wählen, können wir f angeben durch die zugehörige Abbildungsmatrix

$$M_{\mathscr{B},\mathscr{C}}(f) \;\in\; \mathrm{Mat}(m \times n, K)$$

mit $n = \dim_K(U)$, $m = \dim_K(V)$. In den dualen Basen $\mathscr{B}^*$ von U^* und $\mathscr{C}^*$ von V^* wird dann die duale Abbildung beschrieben durch eine Abbildungsmatrix

$$M_{\mathscr{C}^*,\mathscr{B}^*}(f) \;\in\; \mathrm{Mat}(n \times m, K).$$

Wir wollen diese aus der vorigen Abbildungmatrix berechnen. Eine einfache Art, aus einer Matrix vom Format $m \times n$ eine vom Format $n \times m$ zu machen, ist die Vertauschung der Zeilen- und Spaltenindices:

Definition 5.5. Unter der *Transponierten* einer Matrix $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ verstehen wir die Matrix $A^t := (a_{ji}) \in \mathrm{Mat}(n \times m, K)$, die aus A durch Vertauschen der Zeilen und Spalten hervorgeht wie im folgenden Beispiel gezeigt:

$$A \;=\; \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad\Longrightarrow\quad A^t \;=\; \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

Wir können die duale Abbildung zu einer linearen Abbildung wie folgt beschreiben:

Satz 5.6. *Seien U, V zwei endlich erzeugte K-Vektorräume mit Basen $\mathscr{B}$, $\mathscr{C}$, und sei $f\colon U \to V$ eine lineare Abbildung. Seien*

$$A := M_{\mathscr{B},\mathscr{C}}(f) \in \mathrm{Mat}(m\times n,K) \quad \textit{bzw.} \quad B := M_{\mathscr{C}^*,\mathscr{B}^*}(f) \in \mathrm{Mat}(n\times m,K)$$

die Abbildungsmatrizen von f zu den gegebenen Basen bzw. von f^* zu den dualen Basen. Dann gilt

$$B = A^t.$$

Beweis. Sei $\mathscr{B} = (u_1,\dots,u_n)$, $\mathscr{B}^* = (u_1^*,\dots,u_n^*)$, $\mathscr{C} = (v_1,\dots,v_m)$, $\mathscr{C}^* = (v_1^*,\dots,v_m^*)$, und es sei

$$A = (a_{ij}) \quad \text{und} \quad B = (b_{ij}).$$

Per Definition von Abbildungsmatrizen gilt:

$$f(u_i) = \sum_{k=1}^{m} a_{ki}\, v_k \quad \text{und} \quad f^*(v_j^*) = \sum_{l=1}^{n} b_{lj}\, u_l^*.$$

Die zweite Gleichung ist eine Gleichung von Linearformen auf dem Vektorraum U und für deren Werte im Punkt $u_i \in U$ folgt:

$$\big(f^*(v_j^*)\big)(u_i) = \sum_{l=1}^{n} b_{lj}\, u_l^*(u_i) = \sum_{l=1}^{n} b_{lj}\,\delta_{li} = b_{ij}.$$

Per Definition der dualen Abbildung f^* gilt andererseits:

$$\begin{aligned}\big(f^*(v_j^*)\big)(u_i) = v_j^*(f(u_i)) = v_j^*\left(\sum_{k=1}^{m} a_{ki}\, v_k\right) &= \sum_{k=1}^{m} a_{ki}\, v_j^*(v_k)\\ &= \sum_{k=1}^{m} a_{ki}\,\delta_{jk}\\ &= a_{ji}.\end{aligned}$$

Also ist $b_{ij} = a_{ji}$ wie gewünscht. □

Insbesondere kehrt sich bei der Transposition von Matrizen die Reihenfolge der Faktoren in Matrixprodukten um:

Korollar 5.7. *Für $A \in \mathrm{Mat}(l\times m,K)$, $B \in \mathrm{Mat}(m\times n,K)$ gilt $(A\cdot B)^t = B^t \cdot A^t$.*

Beweis. Für $K^n \xrightarrow{g} K^m \xrightarrow{f} K^l$ und $\varphi \in \mathrm{Hom}_K(K^l,K)$ gilt

$$\begin{aligned}(f\circ g)^*(\varphi) = \varphi\circ(f\circ g) &= (\varphi\circ f)\circ g\\ &= f^*(\varphi)\circ g\\ &= g^*(f^*(\varphi))\\ &= (g^*\circ f^*)(\varphi)\end{aligned}$$

und somit $(f\circ g)^* = g^*\circ f^*$. Das hätten wir natürlich ebensogut direkt nachrechnen können, was einen zweiten Beweis derselben Aussage liefert: Sei $A = (A_{ij})$. Wir schreiben kurz $A^t = (A^t_{ij}) = (A_{ji})$ etc. Dann gilt

$$\big(A\cdot B\big)^t_{ij} \;=\; \big(A\cdot B\big)_{ji} \;=\; \sum_k A_{jk}B_{ki} \;=\; \sum_k B_{ki}A_{jk} \;=\; \sum_k B^t_{ik}A^t_{kj} \;=\; \big(B^t\cdot A^t\big)_{ij}$$

für alle i,j. □

Für endlich erzeugte Vektorräume V kann man einen Isomorphismus $V \simeq V^*$ z.B. angeben, indem man eine Basis wählt und diese abbildet auf die dazu duale Basis. Dieser Isomorphismus hängt allerdings von der gewählten Basis ab. Wenn wir zweimal dualisieren, gibt es einen *kanonischen*, d.h. von willkürlichen Wahlen unabhängigen Isomorphismus:

Satz 5.8 (Bidualität). *Sei V ein Vektorraum über K. Dann ist*

$$\iota:\quad V \;\longrightarrow\; V^{**} \;=\; \mathrm{Hom}_K\big(\mathrm{Hom}_K(V,K),K\big),\quad v\mapsto (f\mapsto f(v))$$

eine K-lineare Abbildung, und im Fall $\dim_K(V)<\infty$ ist ι ein Isomorphismus.

Beweis. Wir zeigen zunächst, dass die Abbildung ι den angegebenen Zielbereich hat: Für jedes feste $v\in V$ ist

$$\iota(v):\quad V^* = \mathrm{Hom}_K(V,K) \;\longrightarrow\; K,\quad f\;\mapsto\; f(v)$$

eine K-lineare Abbildung, denn für $f,g\in V^*$ und $\alpha\in K$ ist

$$\iota(v)(f+\alpha g) \;=\; (f+\alpha g)(v) \;=\; f(v)+\alpha g(v) \;=\; \iota(v)(f)+\alpha\iota(v)(g).$$

Weiter zeigen wir, dass $\iota: V\to V^{**}$ eine K-lineare Abbildung ist: Seien $v,w\in V$ und $\alpha\in K$. Für alle $f\in V^*$ ist dann

$$\begin{aligned}(\iota(v+\alpha w))(f) \;=\; f(v+\alpha w) \;=\; f(v)+\alpha\cdot f(w) \;&=\; (\iota(v))(f)+\alpha\cdot(\iota(w))(f)\\ &=\; (\iota(v)+\alpha\cdot\iota(w))(f)\end{aligned}$$

Also ist wie gewünscht

$$\iota(v+\alpha w) \;=\; \iota(v)+\alpha\cdot\iota(w).$$

Zu zeigen bleibt, dass für $\dim_K(V)<\infty$ die Abbildung ι ein Isomorphismus ist: Sei dazu $\mathscr{B} = (v_1,\ldots,v_n)$ eine Basis von V. Sei $\mathscr{B}^* = (f_1,\ldots,f_n)$ die dazu duale Basis von V^*. Für $g_i := \iota(v_i)\in V^{**}$ gilt

$$g_i(f_j) \;=\; (\iota(v_i))(f_j) \;=\; f_j(v_i) \;=\; \delta_{ij}.$$

Also ist $(g_1,\ldots,g_n) \overset{!}{=} \mathscr{B}^{**}$ die zu $\mathscr{B}^*$ duale Basis, insbesondere eine Basis. Damit bildet ι eine Basis von V auf eine Basis von V^{**} ab, ist also ein Isomorphismus. □

Bemerkung 5.9. Im Fall $\dim_K(V) = \infty$ ist die lineare Abbildung $\iota : V \to V^{**}$ zwar noch immer injektiv, aber kein Isomorphismus mehr. Dies sieht man bereits für den Vektorraum

$$V \;=\; \{(a_i)_{i\in\mathbb{N}} \in K^{\mathbb{N}} \mid a_i = 0 \text{ für alle bis auf endlich viele } i\}$$

der endlichen Folgen. Dieser besitzt eine abzählbare Basis, aber in den Übungen werden wir sehen, dass sein Dualraum V^* und somit erst recht sein Doppeldual V^{**} keine abzählbare Basis besitzt.

Kapitel IV
Bild, Kern und Lineare Gleichungssysteme

Zusammenfassung Wir werfen nun einen genaueren Blick auf die Struktur linearer Abbildungen: Wie berechnet man ihr Bild, was sind ihre Fasern? Aus rechnerischer Sicht führt uns dies auf lineare Gleichungssysteme, die mit dem Gauß-Algorithmus gelöst werden können. Der Gauß-Algorithmus wird auch viele weitere Fragen der linearen Algebra beantworten: Wie sieht man z.B., ob eine gegebene quadratische Matrix invertierbar ist, und wie findet man in diesem Fall ihre inverse Matrix?

1 Struktur der Lösungsmengen von LGS

Der Kern und das Bild linearer Abbildungen sind wie für Homomorphismen von Gruppen definiert:

Definition 1.1. Für lineare Abbildungen $f : V \to W$ nennen wir

$$\ker(f) \;=\; \{v \in V \mid f(v) = 0\} \;\subseteq\; V \qquad \textit{den} \text{ Kern } \textit{von } f,$$
$$\operatorname{im}(f) \;=\; \{f(v) \in W \mid v \in V\} \;\subseteq\; W \qquad \textit{das} \text{ Bild } \textit{von } f.$$

Für lineare Abbildungen von Vektorräumen sind dies Untervektorräume:

Lemma 1.2. *Für $f \in \operatorname{Hom}_K(V,W)$ gilt:*

a) $\ker(f) \subseteq V$ *und* $\operatorname{im}(f) \subseteq W$ *sind Untervektorräume.*
b) Es ist $\operatorname{im}(f) = W$ *genau dann, wenn f surjektiv ist.*
c) Es ist $\ker(f) = \{0\}$ *genau dann, wenn f injektiv ist.*

Beweis. Die Aussage $a)$ folgt direkt aus der Linearität

$$f(v + \alpha w) = f(v) + \alpha f(w) \quad \text{für} \quad v, w \in V,\ \alpha \in K.$$

Die Aussage $b)$ gilt per Definition von Surjektivität und hat gar nichts mit linearen Abbildungen zu tun. Zu zeigen bleibt daher nur die Aussage $c)$. Per Definition ist

eine Abbildung f injektiv genau dann, wenn $f(u) = f(v)$ nur für $u = v$ gilt. Im Fall linearer Abbildungen gilt andererseits die Äquivalenz:

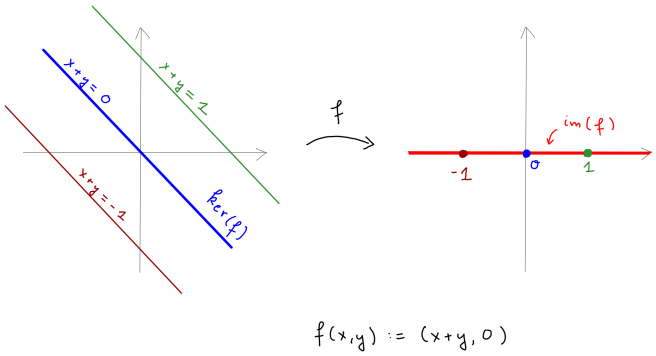
$$f(u) = f(v) \iff f(u-v) = 0 \iff u - v \in \ker(f)$$

Die Injektivität von f ist somit gleichbedeutend mit $\ker(f) = \{0\}$. □

Beispiel 1.3. Für $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $(x,y) \mapsto (x+y, 0)$ gilt

$$\ker(f) \;=\; \{(x,-x) \mid x \in \mathbb{R}\} \quad \text{und} \quad \operatorname{im}(f) \;=\; \{(w,0) \mid w \in \mathbb{R}\}.$$

Die folgende Skizze illustriert die Situation:



Allgemein ist der Kern einer linearen Abbildung zwischen Standardvektorräumen die Lösungsmenge eines linearen Gleichungssystems (LGS):

Beispiel 1.4. Sei $A = (a_{ij}) \in \operatorname{Mat}(m \times n, K)$. Der Kern der linearen Abbildung

$$f: \quad K^n \longrightarrow K^m \quad \text{mit} \quad f(x) := A \cdot x$$

ist die Menge aller Lösungen $x = (x_1, \ldots, x_n)$ des LGS

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &=& 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &=& 0 \\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \quad \vdots & & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &=& 0 \end{array}$$

Ein solches LGS, in dem auf der rechten Seite überall Nullen stehen, bezeichnet man auch als ein *homogenes* LGS. In der obigen Situation schreiben wir auch

$$\ker(A) := \ker(f) \quad \text{und} \quad \operatorname{im}(A) := \operatorname{im}(f)$$

und sprechen vom *Kern* bzw. vom *Bild* der Matrix $A \in \operatorname{Mat}(m \times n, K)$.

Korollar 1.5. *Die Lösungsmenge eines homogenen LGS in n Variablen über K ist ein Untervektorraum von K^n.*

Beweis. Die Lösungsmenge ist der Kern einer linearen Abbildung $f\colon K^n \to K^m$ und nach Lemma 1.2 somit ein Untervektorraum. □

Allgemeiner kann man LGS betrachten, worin auf der rechten Seite von Null verschiedene Konstanten stehen dürfen. Solche *inhomogenen* LGS beschreiben die Fasern linearer Abbildungen:

Beispiel 1.6. Für die lineare Abbildung $f : K^n \to K^m$ im vorigen Beispiel besteht die Faser

$$f^{-1}(b) \quad \text{über einem Vektor} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m$$

genau aus den Lösungen des inhomogenen LGS

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

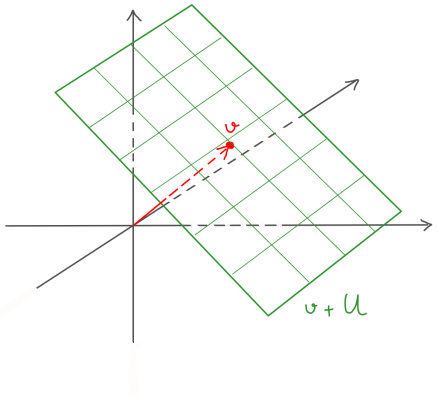
Die Lösungsmenge inhomogener LGS mit von Null verschiedener rechter Seite ist kein Untervektorraum, denn sie enthält den Nullvektor nicht:

Beispiel 1.7. Für $f : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto x+y$ entsteht $f^{-1}(b) = \{(x, b-x) \mid x \in \mathbb{R}\}$ für $b \in \mathbb{R}$ durch Verschiebung der Geraden $\ker(f)$ (siehe Beispiel 1.4).

Definition 1.8. Ein *affiner Unterraum* eines Vektorraumes V ist eine Teilmenge der Form

$$v + U := \{v + u \mid u \in U\} \subseteq V$$

für einen Untervektorraum $U \subseteq V$ und einen beliebigen Vektor $v \in V$.



Lemma 1.9. *Für Untervektorräume $U, U' \subseteq V$ und Vektoren $v, v' \in V$ sind folgende zwei Aussagen äquivalent:*

a) Es ist $v+U = v'+U'$.

b) Es ist $U = U'$ und $v - v' \in U$.

Beweis. $b) \Rightarrow a)$: Betrachte den Vektor $u = v - v' \in U$. Für beliebige Vektoren $w \in V$ gilt

$$\begin{aligned} w \in v+U &\iff w - v \in U && \text{per Def. von } v+U \\ &\iff w - v + u \in U && \text{wegen } u \in U \\ &\iff w - v' \in U && \text{wegen } u = v - v' \\ &\iff w \in v' + U && \text{per Def. von } v'+U \end{aligned}$$

Wenn $b)$ gilt, ist $U = U'$ und dann folgt $v+U = v'+U'$ wie in $a)$ behauptet.

$a) \Rightarrow b)$: Der affine Raum $A = v+U$ bestimmt den Untervektorraum $U \subseteq V$ eindeutig wegen

$$\begin{aligned} U &= \{u_1 - u_2 \in V \mid u_1, u_2 \in U\} \\ &= \{(v+u_1) - (v+u_2) \in V \mid u_1, u_2 \in U\} \\ &= \{w_1 - w_2 \in V \mid w_1, w_2 \in A\} \end{aligned}$$

Es bleibt also nur die Eindeutigkeit des Fußpunktes v zu diskutieren. Für $v, v' \in V$ mit $v+U = v'+U$ gilt $v = v+0 \in v'+U$ und somit $v - v' \in U$, also gilt $b)$. □

Wir halten fest: Aus einem affinen Unterraum $A = v+U \subseteq V$ können wir den Untervektorraum $U \subseteq V$ und bis auf Addition eines beliebigen Vektors aus U auch den Fußpunkt $v \in V$ rekonstruieren. Insbesondere ist die Dimension

$$\dim_K(A) := \dim_K(U)$$

wohldefiniert. Wie wir im Beispiel zu Beginn bereits gesehen haben, sind die Fasern linearer Abbildungen affine Unterräume. Genauer gilt:

Lemma 1.10. *Für $f \in \mathrm{Hom}_K(V,W)$ und $w \in W$ gilt:*

a) Entweder ist $f^{-1}(w) = \varnothing$.

b) Oder es gibt ein $v \in V$ mit $w = f(v)$, und für jedes solche ist $f^{-1}(w) = v + \ker(f)$.

Beweis. Sei $v \in f^{-1}(w)$. Wir müssen zeigen, dass $f^{-1}(w) = v + \ker(f)$ ist. Sei dazu ein beliebiger weiterer Vektor $v' \in f^{-1}(w)$ gegeben, dann gilt

$$f(v'-v) = f(v') - f(v) = w - w = 0$$

wegen der Linearität von f. Somit ist $v' - v \in \ker(f)$, also $v' \in v + \ker(f)$. □

Korollar 1.11 (Lösungsmenge inhomogener LGS). *Sei $A \in \mathrm{Mat}(m \times n, K)$, und für $b \in K^m$ bezeichne*

$$\mathscr{L}(A,b) \;:=\; \big\{x \in K^n \mid Ax = b\big\}$$

die Lösungsmenge des zugehörigen inhomogenen LGS. Dann gilt:

a) Entweder ist $\mathscr{L}(A,b) = \varnothing$.

b) Oder man erhält aus einer beliebigen Lösung $x \in \mathscr{L}(A,b)$ alle weiteren durch Addition von Lösungen des homogenen LGS, d.h.

$$\mathscr{L}(A,b) \;=\; \big\{x+y \mid Ay = 0\big\} \;=\; x + \mathscr{L}(A,0).$$

Beweis. Folgt unmittelbar aus dem vorigen Lemma. □

Wir wollen diese Resultate abschließend noch in kompakter Form mithilfe von Matrizen zusammenfassen. Für Matrizen $A \in \mathrm{Mat}(m \times n, K)$ betrachten wir den Kern und das Bild

$$\ker(A) \;=\; \{v \in K^n \mid Av = 0\} \quad \text{und} \quad \mathrm{im}(A) \;=\; \{Av \in K^m \mid v \in K^n\}.$$

Der *Spaltenrang* der Matrix ist definiert als $\mathrm{rk}(A) := \dim_K \mathrm{im}(A)$.

Lemma 1.12. *Für $A \in \mathrm{Mat}(m \times n, K)$ gilt:*

a) Es ist $\mathrm{im}(A) = \langle v_1, \ldots, v_n \rangle_K$ für die Spalten $v_i = Ae_i$ der Matrix A.

b) Sei $(A \mid b)$ die aus A durch Anhängen einer Spalte b erhaltene Matrix. Dann gilt:

$$b \in \mathrm{im}(A) \quad \Longleftrightarrow \quad \mathrm{rk}(A) = \mathrm{rk}(A \mid b).$$

Beweis. Für $a)$ beachte man, dass jeder Vektor $v \in K^n$ eine Linearkombination der Standardbasisvektoren ist:

$$v = \sum_{i=1}^{n} \alpha_i \cdot e_i \quad \text{mit} \quad \alpha_i \in K.$$

Somit ist

$$Av = \sum_{i=1}^{n} \alpha_i \cdot Ae_i$$

eine Linearkombination der Spalten Ae_i der Matrix. Die Aussage in $b)$ folgt durch Anwenden von $a)$ auf die beiden Matrizen A und $(A \mid b)$. □

Der Spaltenrang einer Matrix lässt sich aus Teil $a)$ des Lemmas sofort ablesen, beispielsweise gilt

$$\mathrm{rk}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} \;=\; 2 \quad \text{und} \quad \mathrm{rk}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix} \;=\; 1.$$

Allgemein erfüllt der Spaltenrang von Matrizen nach dem Lemma stets die obere Abschätzung

$$\mathrm{rk}(A) \;=\; \dim_K \big\langle \text{Spalten von } A \big\rangle_K \;\leq\; \min\{m,n\} \quad \text{für} \quad A \in \mathrm{Mat}(m\times n, K),$$

da die lineare Hülle von n Vektoren im K^m höchstens Dimension $\min\{m,n\}$ haben kann. Wir wollen abschließend unsere Resultate über LGS noch unter Benutzung des Spaltenranges zusammenfassen:

Fazit 1.13. Für $A \in \mathrm{Mat}(m\times n, K)$, $b \in K^m$ sei $\mathscr{L}(A,b) := \{x \in K^n \mid A\cdot x = b\}$, dann gilt:

$$\mathscr{L}(A,b) \neq \varnothing \quad\Longleftrightarrow\quad b \in \mathrm{im}(A) \quad\Longleftrightarrow\quad \mathrm{rk}(A) = \mathrm{rk}(A \mid b).$$

Wenn diese drei äquivalenten Bedingungen erfüllt sind, ist die Lösungsmenge ein affiner Raum

$$\mathscr{L}(A,b) \;=\; x + \ker(A).$$

2 Die Dimensionsformel

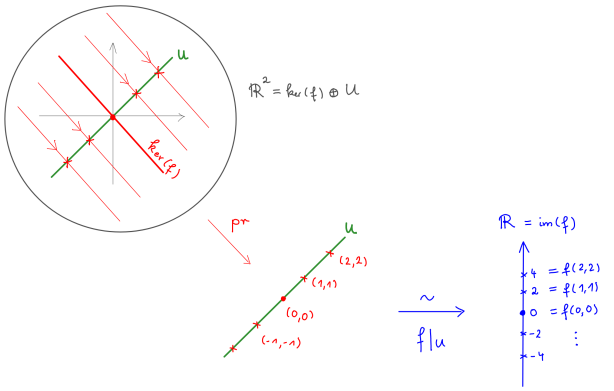
Als nächstes wollen wir uns überlegen, was der Kern und das Bild miteinander zu tun haben. Wir beginnen mit einem Beispiel:

Beispiel 2.1. Für $f : V = \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x,y) \mapsto x+y$ gilt:

a) f ist konstant entlang jeder Parallelen zu der Geraden $\ker(f) = \mathbb{R}\cdot(e_1 - e_2)$.

b) f schränkt sich ein zu einem Isomorphismus auf der Geraden $U = \mathbb{R}\cdot(e_1 + e_2)$.

Man kann f geometrisch verstehen als eine Projektion auf den Untervektorraum U gefolgt von einem Isomorphismus $f|_U : U \xrightarrow{\sim} \mathrm{im}(f) = \mathbb{R}$ auf das Bild:



Wir haben hier den Vektorraum zerlegt als direkte Summe $V = U \oplus \ker(f)$, und für die Projektion haben wir diese Zerlegung benutzt:

Bemerkung 2.2. Sei V ein Vektorraum. Nach dem Basisergänzungssatz können wir zu jedem Untervektorraum $U \subseteq V$ ein *Komplement*, d.h. einen Unterraum $U' \subseteq V$ mit $V = U \oplus U'$ finden. Nach Wahl eines solchen Komplementes hat jedes $v \in V$ eine eindeutige Zerlegung

$$v \;=\; u+u' \quad \text{mit} \quad u \in U \quad \text{und} \quad u' \in U'.$$

Direktes Nachrechnen zeigt, dass $\mathrm{pr}\colon V \to U, v = u+u' \mapsto u$ eine lineare Abbildung ist. Man beachte, dass diese nicht nur von dem gegebenen Untervektorraum $U \subseteq V$, sondern auch vom gewählten Komplement abhängt:

BILD

Das zu Beginn gegebene Beispiel verallgemeinert sich wie folgt:

Satz 2.3. *Sei $f : V \to W$ eine lineare Abbildung und $U \subseteq V$ ein Komplement ihres Kerns, also*

$$V \;=\; U \oplus \ker(f).$$

Sei $\mathrm{pr}\colon V \to U$ *die Projektion auf den ersten Summanden. Dann ist* $f = f|_U \circ \mathrm{pr}$*, und*

$$f|_U : \quad U \xrightarrow{\sim} \mathrm{im}(f) \quad \textit{ist ein Isomorphismus.}$$

Beweis. Nach Voraussetzung hat jeder Vektor $v \in V$ eine Zerlegung $v = u+u'$ mit eindeutigen $u \in U$, $u' \in \ker(f)$. Dabei ist

$$f(v) \;=\; f(u+u') \;=\; f(u)+f(u') \;=\; f(u),$$

also folgt $f = f|_U \circ \mathrm{pr}$. Zudem ist $\ker(f|_U) = U \cap \ker(f) = \{0\}$ wegen der Direktheit der Summe, also ist die lineare Abbildung $f|_U\colon U \hookrightarrow W$ injektiv und somit ein Isomorphismus auf ihr Bild. □

Die wichtigste Folgerung aus diesem Satz ist die Dimensionsformel für lineare Abbildungen, die den Zusammenhang zwischen Bild und Kern klärt:

Korollar 2.4 (Dimensionsformel für lineare Abbildungen). *Sei $f : V \to W$ eine lineare Abbildung von Vektorräumen. Dann ist*

$$\dim_K V \;=\; \dim_K \ker(f) + \dim_K \mathrm{im}(f).$$

Beweis. Sei $U \subseteq V$ ein Komplement von $\ker(f)$. Dann gilt $V = \ker(f) \oplus U$ und somit

$$\dim_K V \;=\; \dim_K \ker(f) + \dim_K U.$$

Nach dem Satz ist dabei $U \simeq \operatorname{im}(f)$, also folgt die Behauptung. Man beachte, dass dieser Beweis auch im Fall $\dim_K V = \infty$ gültig bleibt. □

In Matrizensprache sagt die Dimensionsformel, dass die Dimension des Kerns einer Matrix sich aus ihrem Spaltenrang wie folgt berechnen lässt:

Korollar 2.5. *Für alle $A \in \operatorname{Mat}(m \times n, K)$ gilt $\dim_K \ker(A) = n - \operatorname{rk}(A)$.*

Das ist beispielsweise nützlich, um die Dimension des Kerns von Matrizen mit wenigen Zeilen, aber vielen Spalten zu bestimmen. Für $n \geq 2$ sind z.B. die Spalten von

$$A \;=\; \begin{pmatrix} 1 & 3 & * & \cdots & * \\ 2 & 7 & * & \cdots & * \end{pmatrix} \;\in\; \operatorname{Mat}(2 \times n, \mathbb{R})$$

nicht alle proportional, also ist $\operatorname{rk}(A) = 2$ und wir erhalten $\dim_{\mathbb{R}} \ker(A) = n - 2$.

Korollar 2.6. *Sei $\dim_K(V) < \infty$. Für $f \in \operatorname{End}_K(V)$ sind dann äquivalent:*

a) f ist ein Monomorphismus.

b) f ist ein Epimorphismus.

c) f ist ein Isomorphismus.

Beweis. Es ist

$$\ker(f) = \{0\} \iff \dim \ker(f) = 0 \iff \dim \operatorname{im}(f) = \dim(V) \iff \operatorname{im}(f) = V$$

wobei die Dimensionsformel im zweiten Schritt benutzt wurde. □

Korollar 2.7. *Für quadratische Matrizen $A \in \operatorname{Mat}(n \times n, K)$ gilt:*

$$\operatorname{rk}(A) = n \quad\iff\quad \ker(A) = \{0\} \quad\iff\quad A \text{ ist invertierbar}$$

Dabei heißt eine Matrix $A \in \operatorname{Mat}(n \times n, K)$ *invertierbar*, wenn sie als Element des Ringes $\operatorname{Mat}(n \times n, K)$ multiplikativ invertierbar ist. Die invertierbaren Matrizen bilden offenbar eine Gruppe, die Einheitengruppe des Ringes $\operatorname{Mat}(n \times n, K)$. Wir bezeichnen diese Gruppe mit

$$\operatorname{GL}_n(K) \;:=\; \{A \in \operatorname{Mat}(n \times n, K) \mid A \text{ ist invertierbar}\}$$

und nennen sie die *allgemeine lineare Gruppe* (engl. *G*eneral *L*inear Group). Per Definition ist $A \in \operatorname{Mat}(n \times n, K)$ genau dann, wenn

$$A \cdot B \;=\; \mathbf{1} \;=\; B \cdot A \qquad (\star)$$

für ein $B \in \operatorname{Mat}(n \times n, K)$ und die Einheitsmatrix $\mathbf{1} \in \operatorname{Mat}(n \times n, K)$ ist. Nach dem Kapitel über Gruppen ist dann die *inverse Matrix* $B = A^{-1}$ eindeutig bestimmt. Es

genügt, nur eine der beiden Gleichungen in $(\star)$ zu prüfen: Die erste zeigt $\operatorname{rk}(A) = n$, die zweite $\ker(A) = \{0\}$; in beiden Fällen ist $A \in \mathrm{GL}_n(K)$ nach dem Korollar, und in einer Gruppe sind rechtsinverse auch linksinvers und umgekehrt.

Beispiel 2.8. Sei $c \in K$ und

$$A := \begin{pmatrix} 1 & 1 \\ 1 & c \end{pmatrix} \in \mathrm{Mat}(2 \times 2, K).$$

Die Gleichung $A \cdot B = \mathbf{1}$ mit $B = (b_{ij}) \in \mathrm{Mat}(2 \times 2, K)$ läuft hinaus auf das LGS

$$\begin{aligned} b_{11} + b_{21} &= 1, & b_{12} + b_{22} &= 0, \\ b_{11} + cb_{21} &= 0, & b_{12} + cb_{22} &= 1. \end{aligned}$$

Für $c = 1$ hat dieses LGS keine Lösung, und für $c \neq 1$ ist seine eindeutige Lösung gegeben durch

$$B = \frac{1}{c-1} \begin{pmatrix} c & -1 \\ -1 & 1 \end{pmatrix}.$$

Man rechnet leicht nach, dass dasselbe auch für das (andere) LGS $B \cdot A = \mathbf{1}$ gilt.

Bemerkung 2.9. Die Äquivalenz $A \cdot B = \mathbf{1} \iff B \cdot A = \mathbf{1}$ gilt nur für *quadratische* Matrizen. Beispielsweise ist

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{1}, \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq \mathbf{1}.$$

3 Basiswechsel für lineare Abbildungen

Wenn wir im Beweis der Dimensionsformel aus dem vorigen Abschnitt explizit ein Komplement des Kerns konstruieren, erhalten wir:

Satz 3.1 (Struktursatz für lineare Abbildungen). *Sei $f : V \longrightarrow W$ eine lineare Abbildung zwischen Vektorräumen endlicher Dimension. Dann gibt es Basen $\mathscr{A}$ und $\mathscr{B}$, in denen f durch eine Blockmatrix*

$$M_{\mathscr{A},\mathscr{B}}(f) = \begin{pmatrix} \mathbf{1}_{r\times r} & 0_{r\times b} \\ 0_{a\times r} & 0_{a\times b} \end{pmatrix} \quad \textit{mit} \quad \begin{cases} r = \dim \operatorname{im}(f) \\ b = \dim(V) - r \\ a = \dim(W) - r \end{cases}$$

gegeben ist. Dabei bezeichnet

$$\begin{aligned} \mathbf{1}_{r\times r} &\in \mathrm{Mat}(r \times r, K) \textit{ die Einheitsmatrix}, \\ 0_{s\times t} &\in \mathrm{Mat}(s \times t, K) \textit{ die Nullmatrix für } s,t \in \mathbb{N}. \end{aligned}$$

Beweis. Wir wählen mit dem Basisergänzungssatz eine Basis $\mathscr{A} = (v_1, \ldots, v_n)$ von V mit der Eigenschaft, dass die letzten $n-r$ Vektoren davon eine Basis des Kerns $\ker(f)$ bilden. Dann ist

$$V \;=\; U \oplus \ker(f) \quad \text{für den Unterraum} \quad U \;:=\; \langle v_1, \ldots, v_r \rangle_K.$$

Die Vektoren $v_1, \ldots, v_r$ bilden dann eine Basis von U. Zudem ist nach der Diskussion von Komplementen im vorigen Kapitel die Einschränkung $f|_U : U \xrightarrow{\sim} \mathrm{im}(f)$ ein Isomorphismus. Da Isomorphismen von Vektorräumen Basen auf Basen abbilden, ist $(w_1, \ldots, w_r) := (f(v_1), \ldots, f(v_r))$ eine Basis von $\mathrm{im}(f)$. Wir können diese nach dem Basisergänzungssatz ergänzen zu einer Basis

$$\mathscr{B} \;=\; (w_1, \ldots, w_m) \quad \text{von } W.$$

Per Konstruktion gilt

$$f(v_i) \;=\; \begin{cases} w_i & \text{für } i \le r, \\ 0 & \text{für } i > r. \end{cases}$$

Somit folgt die Behauptung. □

Um den obigen Satz in Matrizensprache zu übersetzen, müssen wir verstehen, wie sich Abbildungsmatrizen unter Basiswechseln verhalten. Sei dazu zunächst V ein endlich erzeugter K-Vektorraum. Nach Wahl einer Basis $\mathscr{B} = (v_1, \ldots, v_n)$ können wir ihn mit dem Standardvektorraum identifizieren mittels

$$\Phi_{\mathscr{B}} : \quad K^n \;\xrightarrow{\sim}\; V, \quad (x_1, \ldots, x_n) \;\mapsto\; \sum_{i=1}^{n} x_i v_i.$$

Definition 3.2. Der *Basiswechsel* zwischen zwei Basen $\mathscr{B}, \mathscr{B}'$ des Vektorraumes V ist der Automorphismus

$$\Phi_{\mathscr{B}',\mathscr{B}} \;:=\; \Phi_{\mathscr{B}}^{-1} \circ \Phi_{\mathscr{B}'} \;\in\; \mathrm{Aut}_K(K^n)$$

in dem folgenden kommutativen Diagramm:

$$\begin{array}{ccc} K^n & \xrightarrow{\;\Phi_{\mathscr{B}',\mathscr{B}}\;} & K^n \\ {\scriptstyle \Phi_{\mathscr{B}'}}\downarrow & & \downarrow{\scriptstyle \Phi_{\mathscr{B}}} \\ V & = & V \end{array}$$

Unter der kanonischen Identifikation $\mathrm{Aut}_K(K^n) = \mathrm{GL}_n(K)$ entspricht dieser einer invertierbaren Matrix, und wir schreiben auch

$$\Phi_{\mathscr{B}',\mathscr{B}} \;\in\; \mathrm{GL}_n(K)$$

für diese sog. *Basiswechselmatrix*. In der Notation des Kapitels III, Definition 3.6 ist dies die Abbildungsmatrix $\Phi_{\mathscr{B}',\mathscr{B}} = M_{\mathscr{B}',\mathscr{B}}(id_V)$ der Identitätsabbildung.

Beispiel 3.3. Oft möchte man von der Standardbasis des Vektorraumes $V = K^n$ zu einer anderen Basis wechseln. Die zugehörige Basiswechselmatrix lässt sich direkt ablesen: Im Fall der Standardbasis $\mathscr{B} = (e_1, \ldots, e_n)$ ist $\Phi_{\mathscr{B}} \colon K^n \longrightarrow V = K^n$ die Identitätsabbildung, also ist

$$\Phi_{\mathscr{B}',\mathscr{B}} \;=\; \Phi_{\mathscr{B}'} \;=\; \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} \;\in\; \mathrm{GL}_n(K)$$

die Matrix, deren Spalten die Vektoren der neuen Basis $\mathscr{B}' = (v_1, \ldots, v_n)$ sind.

Wir wissen nun, wie man zwischen Basen *eines* Vektorraumes wechselt. Für die Beschreibung linearer Abbildungen haben wir es mit *zwei* Vektorräumen zu tun und müssen daher eine Basis im Definitions- und eine im Zielbereich wählen: Sei

- V ein Vektorraum mit einer Basis $\mathscr{A} = (v_1, \ldots, v_n)$,
- W ein Vektorraum mit einer Basis $\mathscr{B} = (w_1, \ldots, w_m)$.

Für eine lineare Abbildung $f : V \to W$ betrachten wir wie im Kapitel III, Def. 3.6 die Abbildungsmatrix

$$M_{\mathscr{A},\mathscr{B}}(f) \;=\; \Phi_{\mathscr{B}}^{-1} \circ f \circ \Phi_{\mathscr{A}} \;\in\; \mathrm{Mat}(m \times n, K) \;=\; \mathrm{Hom}_K(K^n, K^m).$$

Bei Basiswechseln gilt die folgende *Transformationsformel*:

Proposition 3.4. *Sei $f : V \to W$ ein Homomorphismus zwischen endlich erzeugten Vektorräumen über K. Weiter seien*

- *$\mathscr{A}$, $\mathscr{A}'$ zwei Basen von V mit Basiswechselmatrix $S = \Phi_{\mathscr{A}',\mathscr{A}}$,*
- *$\mathscr{B}$, $\mathscr{B}'$ zwei Basen von W mit Basiswechselmatrix $T = \Phi_{\mathscr{B}',\mathscr{B}}$.*

Dann gilt für die entsprechenden Abbildungsmatrizen

$$M_{\mathscr{A}',\mathscr{B}'}(f) \;=\; T^{-1} \cdot M_{\mathscr{A},\mathscr{B}}(f) \cdot S.$$

Beweis. Folgt aus dem kommutativen Diagramm

$$\begin{array}{ccccccc} V & = & V & \xrightarrow{\;f\;} & W & = & W \\ {\scriptstyle \Phi_{\mathscr{A}'}}\uparrow & & {\scriptstyle \Phi_{\mathscr{A}}}\uparrow & & {\scriptstyle \Phi_{\mathscr{B}}}\uparrow & & {\scriptstyle \Phi_{\mathscr{B}'}}\uparrow \\ K^n & \xrightarrow{\;S\;} & K^n & \xrightarrow{\;M_{\mathscr{A},\mathscr{B}}(f)\;} & K^m & \xrightarrow{\;T^{-1}\;} & K^m \end{array}$$

und der Definition der Basiswechselmatrizen. □

Korollar 3.5. *Für jede Matrix $M \in \mathrm{Mat}(m \times n, K)$ gibt es $S \in \mathrm{GL}_n(K)$, $T \in \mathrm{GL}_m(K)$, sodass gilt:*

$$T^{-1}MS \;=\; \begin{pmatrix} \mathbf{1}_{r\times r} & 0_{r\times b} \\ 0_{a\times r} & 0_{a\times b} \end{pmatrix} \quad \textit{mit} \quad r \;=\; \mathrm{rk}(M), \quad a+r \;=\; m, \quad b+r \;=\; n.$$

Beweis. Wir wenden die Proposition an auf $f : K^n \to K^m, f(v) = Mv$ und wählen für $\mathscr{A}$ und $\mathscr{B}$ die Standardbasen. Die Behauptung folgt dann unmittelbar aus dem Struktursatz für lineare Abbildungen (Satz 3.1). □

Im Struktursatz für lineare Abbildungen $f\colon V \to W$ haben wir in V und in W Basen $\mathscr{B}$ und $\mathscr{C}$ gewählt. Im Fall von *Endomorphismen* interessiert uns meist eine feinere Frage: Hier ist $W = V$ und wir wollen dann meist $\mathscr{C} = \mathscr{B}$ wählen. Wir schreiben daher kurz

$$M_{\mathscr{B}}(f) := M_{\mathscr{B},\mathscr{B}}(f) \quad \text{für} \quad f \in \operatorname{End}_K(V)$$

und betrachten in der Proposition 3.4 den Spezialfall $T = S$. Eine so einfache Form wie im obigen Korollar erreicht man nur für Projektionsabbildungen:

Lemma 3.6. *Sei V ein Vektorraum mit $\dim_K(V) < \infty$. Für $f \in \operatorname{End}_K(V)$ sind die folgenden Eigenschaften äquivalent:*

a) Es ist $f \circ f = f$.

b) Man kann $V = V_1 \oplus V_2$ als direkte Summe von Untervektorräumen $V_1, V_2 \subseteq V$ auf solche Weise zerlegen, dass $f = pr$ die Projektion auf den ersten Summanden wird, d.h. $f(v_1 + v_2) = v_1$ für alle $v_1 \in V_1$ und alle $v_2 \in V_2$ gilt:

$$\begin{array}{ccc} V & \xrightarrow{\ f\ } & V \\ \| & & \uparrow \\ V_1 \oplus V_2 & \xrightarrow{\ pr\ } & V_1 \end{array}$$

c) Es gibt eine Basis $\mathscr{B}$ von V mit

$$M_{\mathscr{B}}(f) = \begin{pmatrix} \mathbf{1}_{r\times r} & 0_{r\times s} \\ 0_{s\times r} & 0_{s\times s} \end{pmatrix} \quad \textit{für} \quad \begin{cases} r = \dim_K \operatorname{im}(f), \\ s = \dim_K \ker(f). \end{cases}$$

Beweis. Zunächst gelte $a)$, also $f \circ f = f$. Um $b)$ zu zeigen, setzen wir $V_1 := \operatorname{im}(f)$ und $V_2 := \ker(f)$. Es ist $V = V_1 + V_2$, denn jeder Vektor $v \in V$ lässt sich schreiben als Summe $v = f(v) + (v - f(w))$ mit

- $f(v) \in V_1 = \operatorname{im}(f)$,
- $v - f(v) \in V_2 = \ker(f)$ wegen $f(v - f(v)) = f(v) - f(f(v)) = 0$,

wobei die letzte Gleichung aus $f \circ f = f$ folgt. Zudem gilt $V_1 \cap V_2 = \{0\}$, denn:

- $f(v_1) = v_1$ für alle $v_1 \in V_1 = \operatorname{im}(f)$ wegen $f \circ f = f$,
- $f(v_2) = 0$ für alle $v_2 \in V_2 = \ker(f)$ per Definition des Kerns.

Es folgt $V = V_1 \oplus V_2$ und $f(v_1 + v_2) = v_1$ für alle $v_1 \in V_1, v_2 \in V_2$. Also gilt $b)$.

Aus $b)$ folgt $c)$, indem wir Basen $(v_1, \ldots, v_r)$ von V_1 und $(v_{r+1}, \ldots, v_{r+s})$ von V_2 wählen und $\mathscr{B} = (v_1, \ldots, v_r, v_{r+1}, \ldots, v_{r+s})$ setzen. Aus $c)$ folgt sofort $a)$, weil die angegebene Matrix gleich ihrem Quadrat ist. □

Die Annahme $\dim_K(V) < \infty$ diente nur dazu, in Teil $c)$ Matrizen schreiben zu können: Die Äquivalenz $a) \Leftrightarrow b)$ gilt ganz allgemein. Endomorphismen f mit $f \circ f$ nennt man auch *Projektoren*, *Projektionen* oder *Projektionsabbildungen*.

4 Ein zweiter Blick auf den Gauß-Algorithmus

Wir wollen nun einige Anwendungen des Gauß-Algorithmus in der linearen Algebra betrachten. Seine Aufgabe ist zunächst die Lösung von LGS $A \cdot x = b$ mit $b \in K^m$ und $A \in \mathrm{Mat}(m \times n, K)$. Jede Zeile von A entspricht einer Gleichung in dem LGS, also schreiben wir

$$A = \begin{pmatrix} \text{---} & a_1 & \text{---} \\ & \vdots & \\ \text{---} & a_m & \text{---} \end{pmatrix}$$

als Matrix von Zeilenvektoren. Zeilenumformungen von A kann man beschreiben durch Linksmultiplikation mit den Standardmatrizen $E_{ij} \in \mathrm{Mat}(m \times m, K)$, die den Eintrag 1 an der Stelle (i, j) haben und sonst nur Nullen enthalten:

Bemerkung 4.1. Die Linksmultiplikation mit E_{ij} ersetzt eine Matrix A durch die Matrix, deren i-te Zeile die j-ten Zeile der ursprünglichen Matrix ist und deren übrige Zeilen nur Nullen enthalten:

$$E_{ij} \cdot \begin{pmatrix} & \vdots & \\ \text{---} & a_{i-1} & \text{---} \\ \text{---} & a_i & \text{---} \\ \text{---} & a_{i+1} & \text{---} \\ & \vdots & \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \text{---} & 0 & \text{---} \\ \text{---} & a_j & \text{---} \\ \text{---} & 0 & \text{---} \\ & \vdots & \end{pmatrix} \longleftarrow i$$

Da jedes $S \in \mathrm{Mat}(l \times m, K)$ eine Linearkombination von Standardmatrizen ist, sind somit für beliebige Matrizen $A \in \mathrm{Mat}(m \times n, K)$ die Zeilen der Produktmatrix $S \cdot A$ Linearkombinationen der Zeilen von A. Also: Lineare Operationen auf den Zeilen einer Matrix sind gegeben durch Linksmultiplikation mit einer Matrix! Betrachten wir dies für die im Gauß-Algorithmus benutzten *elementaren Umformungen*:

- Multipliziere die i-te Zeile von A mit einem Skalar $\alpha \in K^\times = K \setminus \{0\}$: Das ist die Linksmultiplikation mit der Matrix

$$S_i(\alpha) = \mathbf{1} + (\alpha - 1)E_{ii} \in \mathrm{GL}_m(K).$$

- Addiere zur i-ten Zeile von A die j-te: Das ist die Linksmultiplikation mit der Matrix

$$S_{ij} = \mathbf{1} + E_{ij} \in \mathrm{GL}_m(K).$$

Hieraus lassen sich weitere elementare Umformungen zusammensetzen:

- Addiere zur i-ten Zeile von A das α-fache der j-ten für ein $j \neq i$: Dies ist die Linksmultiplikation mit

$$S_{ij}(\alpha) \;=\; \mathbf{1} + \alpha E_{ij} \;=\; S_j(\alpha^{-1}) \cdot S_{ij} \cdot S_j(\alpha) \in \mathrm{GL}_m(K).$$

- Vertausche zwei Zeilen von A miteinander: Dies ist die Linksmultiplikation mit

$$T_{ij} \;=\; \mathbf{1} + E_{ij} + E_{ji} - E_{ii} - E_{jj} \;=\; S_{ji} \cdot S_{ij}(-1) \cdot S_{ji} \cdot S_j(-1) \;\in\; \mathrm{GL}_m(K).$$

Matrizen der Form $S_i(\alpha)$, $S_{ij}(\alpha)$ und T_{ij} heißen auch *Elementarmatrizen*. Man kann die obigen Zeilenumformungen zum Berechnen des Spaltenranges von A und der Lösungsmenge $\mathscr{L}(A,b) = \{x \in K^n \mid Ax = b\}$ von LGS verwenden:

Satz 4.2. *Sei* $A \in \mathrm{Mat}(m \times n, K)$

a) Man kann A durch Zeilenumformungen auf sog. reduzierte Zeilenstufenform *bringen: Es gibt ein Produkt $S \in \mathrm{GL}_m(K)$ von Elementarmatrizen mit*

$$SA \;=\; \begin{pmatrix} 1 & * \cdots * & 0 & * \cdots * & 0 & * \cdots * & \cdots & 0 & * \cdots * \\ & & 1 & * \cdots * & 0 & * \cdots * & \cdots & 0 & * \cdots * \\ & & & & 1 & * \cdots * & \cdots & 0 & * \cdots * \\ & & & & & & \cdots & \cdots & \\ & & & & & & & 1 & * \cdots * \\ & & & & & & & & \\ & & & & & & & & \end{pmatrix}$$

wobei alle nicht bezeichneten Stellen Null sind und $*$ beliebige Einträge sind.

b) Für alle $S \in \mathrm{GL}_m(K)$, $b \in K^m$ ist $\mathscr{L}(A,b) = \mathscr{L}(SA, Sb)$ und $\mathrm{rk}(A) = \mathrm{rk}(SA)$.

Beweis. Siehe Seite 6 in der Einleitung. Wir haben für die Rechnungen dort nur die Körperaxiome benutzt, das Ergebnis gilt daher über jedem Körper K. Die Aussage über den Rang folgt daraus, dass $S\colon \mathrm{im}(A) \to \mathrm{im}(SA)$ ein Isomorphismus ist. □

Anwendung 1 — Der klassische Gauß-Algorithmus

1. Betrachte für das LGS $Ax = b$ die erweiterte Koeffizientenmatrix $M = (A|b)$.
2. Forme diese mit elementaren Zeilenoperationen um zu einer Matrix $\tilde{M} = (\tilde{A}|\tilde{b})$ mit $\tilde{A}$ in reduzierter Zeilenstufenform, wie im Beweis auf Seite 6.
3. Sei r die Anzahl der von Null verschiedenen Zeilen von $\tilde{A}$, dann ist $\mathrm{rk}(A) = r$ und es gibt zwei Fälle:
   - Wenn $\tilde{b}$ einen von Null verschiedenen Eintrag in der i-ten Zeile für ein $i > r$ hat, dann hat das LGS $Ax = b$ keine Lösung.
   - Sonst liest man alle Lösungen wie auf Seite 7 ab: Man wähle dazu die freien Variablen beliebig und bestimme die übrigen durch Einsetzen in $\tilde{A} \cdot x = \tilde{b}$.

Beispiel 4.3. Für $c \in \mathbb{R}$ betrachten wir das LGS

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 + 2x_2 + 3x_3 + 4x_4 &= 2 \\ 3x_1 + 4x_2 + 5x_3 + 6x_4 &= c \end{aligned}$$

Der Gauß-Algorithmus liefert:

$$\left(\begin{array}{cccc|c} 1&1&1&1&1\\ 1&2&3&4&2\\ 3&4&5&6&c \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1&1&1&1&1\\ 0&1&2&3&1\\ 0&1&2&3&c-3 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1&1&1&1&1\\ 0&1&2&3&1\\ 0&0&0&0&c-4 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1&0&-1&-2&0\\ 0&1&2&3&1\\ 0&0&0&0&c-4 \end{array}\right)$$

Das LGS ist also lösbar genau für $c = 4$, und dann sind seine Lösungen genau die Vektoren

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \lambda + 2\mu \\ 1 - 2\lambda - 3\mu \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

mit $\lambda, \mu \in \mathbb{R}$. Dies liefert eine Parametrisierung der Lösungsmenge des LGS als affiner Unterraum, angegeben durch einen Fußpunkt und eine Basis des zugehörigen Untervektorraumes (der Lösungsmenge des zugehörigen homogenen LGS).

Anwendung 2 — Invertieren von Matrizen

Nach der Dimensionsformel ist eine Matrix $A \in \mathrm{Mat}(n \times n, K)$ invertierbar genau für $\ker(A) = \{0\}$. Wenn wir mit dem Gauß-Algorithmus prüfen, ob dies der Fall ist, können wir die inverse Matrix im Fall ihrer Existenz gleich mit ausrechnen:

Lemma 4.4. *Sei $A \in \mathrm{Mat}(n \times n, K)$ gegeben, und sei $M = (A \mid \mathbf{1}) \in \mathrm{Mat}(n \times 2n, K)$ hieraus durch Anhängen der Einheitsmatrix erhalten. Man forme M mit elementaren Zeilenoperationen um zu $\tilde{M} = (\tilde{A} \mid B)$ mit $\tilde{A}$ in reduzierter Zeilenstufenform. Dann gilt:*

- *Falls $\tilde{A} \neq \mathbf{1}$ ist, dann ist A nicht invertierbar.*
- *Falls $\tilde{A} = \mathbf{1}$ ist, dann ist A invertierbar und die inverse Matrix ist $A^{-1} = B$.*

Insbesondere ist jede invertierbare Matrix A ein Produkt von Elementarmatrizen.

Beweis. Sei $S \in \mathrm{GL}_n(K)$ mit $\tilde{A} = SA$ in reduzierter Zeilenstufenform. Dann gilt

$$A \in \mathrm{GL}_n(K) \iff \tilde{A} = SA \in \mathrm{GL}_n(K) \iff \tilde{A} = \mathbf{1},$$

wobei die zweite Äquivalenz benutzt, dass die einzige invertierbare quadratische Matrix in reduzierter Zeilenstufenform die Einheitsmatrix ist. Im Fall $\tilde{A} = \mathbf{1}$ erhalten wir $(\mathbf{1} \mid B) = (\tilde{A} \mid B) = \tilde{M} = S \cdot M = S \cdot (A \mid \mathbf{1}) = (SA \mid S)$. Es folgt $SA = \mathbf{1}$ und $B = S$, also $A^{-1} = B$ und dann ist $A = B^{-1}$ ein Produkt von Elementarmatrizen. □

Beispiel 4.5. Für welche $\lambda \in \mathbb{R}$ ist die Matrix

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 3 & 4 \\ 1 & 4 & \lambda \end{pmatrix} \in \mathrm{Mat}(3\times 3, \mathbb{R})$$

invertierbar, und wie sieht dann die inverse Matrix aus? Die ersten Schritte des Gauß-Algorithmus liefern

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 1 & 3 & 4 & 0 & 1 & 0 \\ 1 & 4 & \lambda & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 4 & -1 & 1 & 0 \\ 0 & 2 & \lambda & -1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 4 & -1 & 1 & 0 \\ 0 & 0 & \lambda-8 & 1 & -2 & 1 \end{array}\right).$$

Somit ist A invertierbar genau für $\lambda \neq 8$. Wenn wir die inverse Matrix berechnen wollen, müssen wir weitere Umformungen anwenden, bis im linken Matrixblock die 3×3 Einheitsmatrix steht. Für $\lambda = 9$ liefert der Gauß-Algorithmus z.B.

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 4 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -5 & 9 & -4 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 11 & -18 & 8 \\ 0 & 1 & 0 & -5 & 9 & -4 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right).$$

Nach dem vorigen Lemma folgt

$$\begin{pmatrix} 1 & 2 & 0 \\ 1 & 3 & 4 \\ 1 & 4 & 9 \end{pmatrix}^{-1} = \begin{pmatrix} 11 & -18 & 8 \\ -5 & 9 & -4 \\ 1 & -2 & 1 \end{pmatrix}.$$

Anwendung 3 — Basen für den Aufspann von Vektoren

Um eine Basis für den Aufspann von gegebenen Spaltenvektoren $v_1, \ldots, v_n \in K^m$ zu finden, betrachten wir den Aufspann als das Bild $\mathrm{im}(A)$ der aus den Spaltenvektoren gebildeten Matrix

$$A = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} \in \mathrm{Mat}(m\times n, K).$$

Da wir mit Spaltenvektoren arbeiten, sollten wir nun statt Zeilenumformungen besser Spaltenumformungen betrachten. Zeilenumformungen hatten wir durch die Linksmultiplikation $A \mapsto S\cdot A$ mit invertierbaren Matrizen S beschrieben, analog sieht man, dass Spaltenumformungen durch die Rechtsmultiplikation $A \mapsto A\cdot S$ mit invertierbaren Matrizen S gegeben sind. Wenn wir für S Elementarmatrizen wählen, erhalten wir *elementare Spaltenumformungen*:

- Multiplizieren einer Spalte von A mit einem Skalar $\alpha \neq 0$.
- Addieren eines Vielfachen einer Spalte von A zu einer anderen Spalte von A.
- Vertauschen zweier Spalten von A.

Satz 4.6. *Sei $A \in \mathrm{Mat}(m \times n, K)$.*

a) Man kann A durch Spaltenumformungen auf eine reduzierte Spaltenstufenform umformen: Es gibt ein Produkt $S \in \mathrm{GL}_n(K)$ von Elementarmatrizen mit

$$A \cdot S = \begin{pmatrix} & & & & \\ 1 & & & & \\ * & & & & \\ * & & & & \\ 0 & 1 & & & \\ * & * & & & \\ * & * & & & \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & \cdots & 0 & 1 \\ * & * & & * & * \\ * & * & \cdots & * & * \end{pmatrix}$$

b) Die von Null verschiedenen Spalten von $A \cdot S$ bilden dann eine Basis von $\mathrm{im}(A)$.

Beweis. Satz 4.2 angewandt auf die transponierte Matrix $A^t \in \mathrm{Mat}(n \times m, K)$ liefert ein Produkt $T \in \mathrm{GL}_n(K)$ von Elementarmatrizen mit der Eigenschaft, dass $T \cdot A^t$ in reduzierter Zeilenstufenform ist. Mit T ist auch ihre Transponierte $S = T^t$ ein Produkt von Elementarmatrizen, und

$$A \cdot S = (T \cdot A^t)^t$$

ist in reduzierter Spaltenstufenform. Wegen $S \in \mathrm{GL}_m(K)$ gilt ferner

$$\mathrm{im}(A) = \{A \cdot v \mid v \in K^n\} = \{A \cdot S \cdot w \mid w \in K^n\} = \mathrm{im}(A \cdot S),$$

und wegen der Spaltenstufenform bilden die von Null verschiedenen Spalten der Matrix $A \cdot S$ eine Basis des Bildes $\mathrm{im}(A \cdot S)$. □

Ähnlich wie beim Gauß-Algorithmus für den Kern einer Matrix hätte es für die Aussage in b) bereits genügt, dass $A \cdot S$ in Spaltenstufenform ist. Die reduzierte Spaltenstufenform macht es aber zugleich besonders bequem, eine Darstellung des Bildes $\mathrm{im}(A) = \mathrm{im}(A \cdot S)$ als Lösungsmenge eines LGS abzulesen:

Beispiel 4.7. Sei $V = \langle v_1, v_2, v_3 \rangle_{\mathbb{R}} \subset \mathbb{R}^4$ der Aufspann von

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 8 \\ 14 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 3 \\ 11 \\ 19 \end{pmatrix} \quad \text{und} \quad v_3 = \begin{pmatrix} 1 \\ 4 \\ 14 \\ 24 \end{pmatrix}.$$

Die Spaltenversion des Gauß-Algorithmus liefert:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 8 & 11 & 14 \\ 14 & 19 & 24 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 2 \\ 8 & 3 & 6 \\ 14 & 5 & 10 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 0 \end{pmatrix} = A \cdot S.$$

Somit hat V eine Basis bestehend aus den beiden Vektoren

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 4 \end{pmatrix} \quad \text{und} \quad u_2 = \begin{pmatrix} 0 \\ 1 \\ 3 \\ 5 \end{pmatrix}.$$

Hieraus liest man zugleich ab, dass der Untervektorraum $V \subseteq \mathbb{R}^4$ die Lösungsmenge des folgenden LGS ist:

$$\begin{cases} 2x+3y = z \\ 4x+5y = w \end{cases}$$

Abschließend wollen wir noch das Verhältnis zwischen dem Spaltenrang und dem Zeilenrang einer Matrix klären:

Lemma 4.8. *Für jede Matrix $A \in \mathrm{Mat}(m \times n, K)$ ist ihr Spaltenrang gleich ihrem Zeilenrang, d.h. es gilt*

$$\mathrm{rk}(A) := \dim_K \langle \textit{Spalten von } A \rangle_K = \dim_K \langle \textit{Zeilen von } A \rangle_K.$$

Beweis. Nach Satz 4.2(b) ist der Spaltenrang $\mathrm{rk}(A)$ unter Zeilenumformungen der Matrix A invariant, d.h. es gilt

$$\mathrm{rk}(SA) = \mathrm{rk}(A) \quad \text{für alle } S \in \mathrm{GL}_m(K).$$

Der Zeilenrang $\mathrm{rk}(A^t)$ erfüllt ebenfalls

$$\mathrm{rk}((SA)^t) = \mathrm{rk}(A^t S^t) = \mathrm{rk}(A^t) \quad \text{für alle } S \in \mathrm{GL}_m(K),$$

denn es ist $\mathrm{im}(A^t S^t) = \mathrm{im}(A^t)$, weil mit S auch S^t eine invertierbare Matrix ist. Um die Gleichheit von Spalten- und Zeilenrang zu zeigen, dürfen wir also A ersetzen durch $\tilde{A} = SA$. Nach Satz 4.2 können wir annehmen, dass $\tilde{A}$ in Zeilenstufenform ist, und dann ist die zu beweisende Gleichung $\mathrm{rk}(\tilde{A}) = \mathrm{rk}(\tilde{A}^t)$ offensichtlich. □

Bemerkung 4.9. Anschaulich misst $\mathrm{rk}(A^t)$ die Anzahl unabhängiger Gleichungen in dem LGS $Ax = 0$. Dies ist ein LGS in n Variablen; nach der Dimensionsformel besitzt seine Lösungsmenge die Dimension $n - \mathrm{rk}(A)$. Das obige Lemma liefert als Slogan für die Dimensionformel:

Dimension der Lösungsmenge = #(Variablen) − #(unabhängige Gleichungen)

5 Exkurs: Der Satz von Skolem-Noether

Zeilen- und Spaltenumformungen sind auch für theoretische Argumente nützlich; wir wollen hier als Beispiel alle Automorphismen von Matrixalgebren über einem

Körper K bestimmen. Zur Erinnerung: Eine *K-Algebra* ist ein Ring R, der zugleich ein Vektorraum über K mit derselben Addition ist, sodass die Multiplikation des Ringes mit der Skalarmultiplikation des Vektorraumes verträglich ist:

$$\forall a \in K \ \forall f,g \in R: \quad a(f \cdot g) \ = \ (af) \cdot g \ = \ f \cdot (ag)$$

Ein *Automorphismus* einer K-Algebra R ist eine bijektive Abbildung $\varphi : R \longrightarrow R$, die ein Homomorphismus sowohl von Ringen als auch von Vektorräumen ist. Wir interessieren uns hier für die Matrixalgebra $R := \mathrm{Mat}(n \times n, K)$. Jede invertierbare Matrix $A \in \mathrm{GL}_n(K)$ liefert einen Automorphismus

$$\gamma_A : \quad R \xrightarrow{\sim} R, \quad B \mapsto A \cdot B \cdot A^{-1}$$

Automorphismen von dieser Form nennt man auch *innere Automorphismen*.

Beispiel 5.1. Die Abbildung

$$\varphi : \ \mathrm{Mat}(2 \times 2, K) \longrightarrow \mathrm{Mat}(2 \times 2, K), \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} d & c \\ b & a \end{pmatrix}$$

ist ein Automorphismus der K-Algebra $\mathrm{Mat}(2 \times 2, K)$. Dieser hat die Form $\varphi = \gamma_A$ für die Matrix

$$A \ = \ A^{-1} \ = \ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ \in \ \mathrm{GL}_2(K).$$

Denn Linksmultiplikation mit A vertauscht die Zeilen, Rechtsmultiplikation mit A die Spalten einer Matrix, und beides zusammen ergibt φ. Allgemein gilt:

Satz 5.2 (Skolem-Noether). *Zu jedem Automorphismus*

$$\varphi : \quad \mathrm{Mat}(n \times n, K) \xrightarrow{\sim} \mathrm{Mat}(n \times n, K)$$

von K-Algebren gibt es eine invertierbare Matrix $A \in \mathrm{GL}_n(K)$ mit $\varphi = \gamma_A$.

Beweis. Wenn es so ein A gibt, gilt für die $v_k := Ae_k$ jedenfalls

$$\varphi(E_{ij}) \cdot v_k \ = \ AE_{ij}A^{-1} \cdot Ae_k \ = \ A \cdot E_{ij} \cdot e_k \ = \ \delta_{jk} \cdot A \cdot e_i \ = \ \delta_{jk} \cdot v_i$$

und φ ist dadurch eindeutig bestimmt. Wir werden umgekehrt Vektoren $v_1, \ldots, v_n$ mit der Eigenschaft $\varphi(E_{ij}) \cdot v_k = \delta_{jk} \cdot v_i$ konstruieren und zeigen, dass die aus diesen Vektoren als Spalten gebildete quadratische Matrix A das Gewünschte leistet. Um die Vektoren zu konstruieren, setze $F_{ij} := \varphi(E_{ij}) \in \mathrm{Mat}(n \times n, K)$. Dann gilt

$$F_{ij} \cdot F_{kl} \ = \ \varphi(E_{ij} \cdot E_{kl}) \ = \ \varphi(\delta_{jk} E_{il}) \ = \ \delta_{jk} F_{il}.$$

weil φ ein Ringhomomorphismus ist. Wegen $F_{11} \neq 0$ können wir ein $u \in K^n$ wählen mit der Eigenschaft $F_{11} \cdot u \neq 0$. Wir definieren nun $v_k := F_{k1} \cdot u$. Für diese Vektoren gilt wie gewünscht:

$$\varphi(E_{ij}) \cdot v_k \ = \ F_{ij} \cdot v_k \ = \ F_{ij} \cdot F_{k1} \cdot u \ = \ \delta_{jk} \cdot F_{i1} \cdot u \ = \ \delta_{jk} \cdot v_i.$$

Als nächstes zeigen wir, dass die Vektoren $v_1, \ldots, v_n$ eine Basis von K^n bilden. Da es sich um n Vektoren handelt, genügt es, die lineare Unabhängigkeit zu prüfen. Sei dazu

$$0 \;=\; \sum_{k=1}^{n} \alpha_k v_k \quad \text{für Skalare} \quad \alpha_1, \ldots, \alpha_n \in K.$$

Multiplikation mit der Matrix F_{1i} liefert

$$0 \;=\; \sum_{k=1}^{n} \alpha_k \cdot F_{1i} \cdot v_k \;=\; \sum_{k=1}^{n} \alpha_k \cdot F_{1i} \cdot F_{k1} \cdot u \;=\; \sum_{k=1}^{n} \alpha_k \cdot \delta_{ik} \cdot F_{11} \cdot u \;=\; \alpha_i \cdot F_{11} u$$

und somit $\alpha_i = 0$ wegen $F_{11} u \neq 0$. Also sind die Vektoren linear unabhängig und bilden somit eine Basis. Die aus diesen Vektoren als Spalten gebildete Matrix ist somit invertierbar, d.h.

$$A \;:=\; \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} \;\in\; \mathrm{GL}_n(K).$$

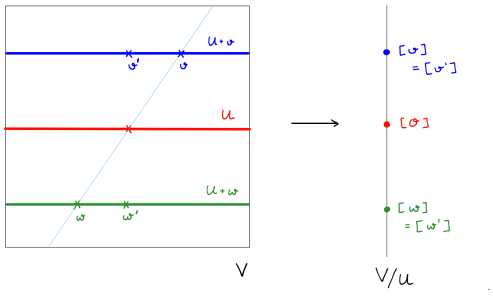
Dann gilt $\gamma_A(E_{ij}) \cdot v_k = A E_{ij} A^{-1} \cdot A e_k = \delta_{jk} \cdot v_i = \varphi(E_{ij}) \cdot v_k$ und somit $\varphi = \gamma_A$. □

6 Quotientenvektorräume

Im Beweis der Dimensionsformel haben wir ein Komplement des Kerns *gewählt*; wir wollen uns nun überlegen, wie man solche willkürlichen Wahlen vermeiden kann. Sei V ein K-Vektorraum und $U \subseteq V$ ein Untervektorraum. Dann ist $U \subseteq V$ eine additive Untergruppe. Die Quotientengruppe V/U besteht per Definition aus den Äquivalenzklassen $[v]$ von Vektoren $v \in V$ modulo der Äquivalenzrelation

$$v \sim v' \quad \overset{\text{def}}{\Longleftrightarrow} \quad v - v' \;\in\; U$$

Die Punkte von V/U sind also affine Unterräume $[v] = v + U$ von V:



Lemma 6.1. *Der Quotient V/U trägt eine eindeutige Vektorraumstruktur, sodass die Quotientenabbildung*

$$p: \quad V \twoheadrightarrow V/U, \quad v \mapsto [v]$$

ein Homomorphismus von Vektorräumen ist. Es ist $U = \ker(p)$.

Beweis. Die Abbildung p ist per Definition surjektiv. Wenn die Gruppe V/U die Struktur eines Vektorraumes trägt, sodass p linear ist, dann muß die Addition und Skalarmultiplikation repräsentantenweise erfolgen:

$$[v] + \alpha \cdot [w] \;=\; p(v) + \alpha \cdot p(w) \;=\; p(v + \alpha \cdot w) \;=\; [v + \alpha \cdot w].$$

Dies zeigt die Eindeutigkeit der gesuchten Vektorraumstruktur. Für die Existenz bleibt nur zu zeigen, dass die repräsentantenweise Addition $[v] + [w] := [v + w]$ und Skalarmultiplikation $\alpha \cdot [v] := [\alpha \cdot v]$ wohldefiniert sind, d.h. nicht von der Wahl der Repräsentanten abhängen. Für $+$ hatten wir uns das bereits im Kapitel über Gruppen überlegt, das Argument für $\cdot$ ist analog. Wir machen beides auf einmal: Seien $v, v', w, w' \in V$ mit $[v] = [v']$ und mit $[w] = [w']$. Dann ist $v \sim v'$ und $w \sim w'$, also

$$v - v' \in U \quad \text{und} \quad w - w' \in U.$$

Da $U \subseteq V$ ein Untervektorraum ist, folgt für $\alpha \in K$ auch

$$(v + \alpha w) - (v' + \alpha w') \;=\; (v - v') + \alpha(w - w') \;\in\; U$$

Es ist also $v + \alpha w \sim v' + \alpha w'$ und somit $[v + \alpha w] = [v' + \alpha w']$. □

Korollar 6.2 (Dimensionsformel für Quotienten). *Für Untervektorräume $U \subseteq V$ gilt*

$$\dim(V) \;=\; \dim(U) + \dim(V/U).$$

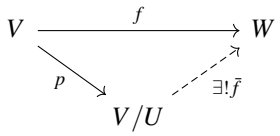
Beweis. Wende die Dimensionsformel an auf die lineare Abbildung $p : V \twoheadrightarrow V/U$ mit $\ker(p) = U$ und $\operatorname{im}(p) = V/U$. □

Homomorphismen von Quotientenvektorräumen in andere Vektorräume kann man mit folgender sogenannter *universeller Eigenschaft* verstehen:

Satz 6.3 (Homomorphiesatz). *Sei V ein Vektorraum über K, und sei $U \subseteq V$ ein Untervektorraum. Dann gibt es für jede K-lineare Abbildung $f\colon V \to W$ mit der Eigenschaft $U \subseteq \ker(f)$ genau eine K-lineare Abbildung*

$$\bar{f}: \quad V/U \longrightarrow W$$

mit $f = \bar{f} \circ p$ für die Quotientenabbildung $p\colon V \twoheadrightarrow V/U$. Wir fassen diese Aussage mit dem Diagramm



zusammen. Für die eindeutig bestimmte Abbildung $\bar{f}$ gilt:

a) $\operatorname{im}(\bar{f}) = \operatorname{im}(f)$.

b) *$\bar{f}$ ist injektiv genau für* $U = \ker(f)$.

Beweis. Die Eindeutigkeit ist klar, denn für jede Abbildung $\bar{f}$ mit $f = \bar{f} \circ p$ muß gelten:

$$\bar{f}([v]) \;=\; \bar{f}(p(v)) \;= (\bar{f} \circ p)(v) \;= f(v).$$

Für die Existenz wollen wir umgekehrt

$$\bar{f}\colon \quad V/U \;\longrightarrow\; W, \quad [v] \;\mapsto\; f(v)$$

setzen. Dann gilt per Konstruktion $f = \bar{f} \circ p$, allerdings ist die Wohldefiniertheit der Abbildung, also die Unabhängigkeit der soeben versuchten Definition von $\bar{f}$ von den gewählten Repräsentanten, nachzuprüfen: Für $[v] = [v']$ ist $v - v' \in U \subseteq \ker(f)$, also $f(v) = f(v')$ wegen

$$f(v) - f(v') \;=\; f(v - v') \;=\; 0.$$

Damit ist $\bar{f}$ durch $\bar{f}([v]) := f(v)$ wohldefiniert. Per Konstruktion ist $\operatorname{im}(f) = \operatorname{im}(\bar{f})$, zudem gilt

$$\begin{aligned} \ker(\bar{f}) \;&=\; \{[v] \in V/U \mid \bar{f}([v]) = 0\} \\ &=\; \{[v] \in V/U \mid f(v) = 0\} \\ &=\; \{[v] \in V/U \mid v \in \ker(f)\} \end{aligned}$$

und somit

$$\begin{aligned} \ker(\bar{f}) = 0 \;&\iff\; \forall v \in \ker(f)\colon\ [v] = [0] \text{ in } V/U \\ &\iff\; \forall v \in \ker(f)\colon\ v \;\in\; U \\ &\iff\; \ker(f) = U \end{aligned}$$

□

Korollar 6.4. *Jeder Homomorphismus $f : V \to W$ von Vektorräumen induziert einen Isomorphismus*

$$\bar{f}\colon \quad V/\ker(f) \;\xrightarrow{\sim}\; \operatorname{im}(f).$$

Beweis. Für $U \subseteq \ker(f)$ gibt der Homomorphiesatz eine Faktorisierung $f = i \circ \bar{f} \circ p$ wie im folgenden Diagram, wobei p die Quotientenabbildung und i die Inklusion bezeichnet:

$$\begin{array}{ccc} V & \xrightarrow{\;f\;} & W \\ {\scriptstyle p}\downarrow & & \uparrow{\scriptstyle i} \\ V/U & \xrightarrow{\;\bar{f}\;} & \operatorname{im}(f) \end{array}$$

Für $U = \ker(f)$ ist $\bar{f}$ injektiv, also ein Isomorphismus auf das Bild. □

Kapitel V
Die Determinante

Zusammenfassung Der Flächeninhalt von Parallelogrammen und das Volumen von Parallelotopen führen auf Determinanten. In diesem Kapitel werden wir allgemein die Determinante von n Vektoren in K^n durch die Leibniz-Formel definieren als eine alternierende Summe über alle Permutationen. Nach einigen Beispielen werden wir die Multiplikativität von Determinanten für Matrixprodukte zeigen und sehen, wie man Determinanten leicht durch Entwickeln nach Zeilen oder Spalten berechnet.

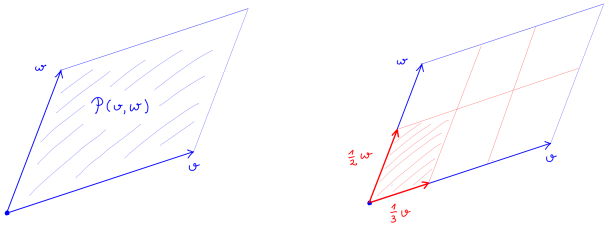
1 Motivation: Flächeninhalte

Gegeben sei das von zwei Vektoren $u, v \in \mathbb{R}^2$ in der reellen Ebene aufgespannte Parallelogramm

$$\mathscr{P}(v,w) \;:=\; \big\{\, \alpha v + \beta w \in \mathbb{R}^2 \mid \alpha, \beta \in [0,1] \,\big\}.$$

Wir wollen diesem einen Flächeninhalt zuordnen, der im Folgenden mit $\det(v,w)$ bezeichnet sei. Dazu legen wir zunächst den Maßstab fest durch die *Normierung*, dass für das Einheitsquadrat gilt: $\det(e_1,e_2) = 1$. Außerdem fordern wir für $a, b \in \mathbb{R}$ die *Skalierungseigenschaft*

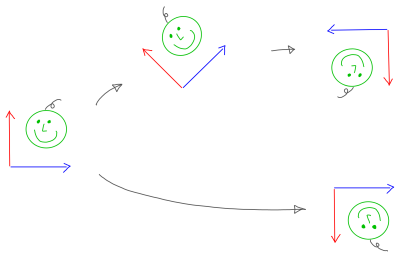
$$\det(av, bw) \;=\; ab \cdot \det(v,w)\,:$$



Damit die Skalierungseigenschaft auch für $a < 0$ oder $b < 0$ gilt, sollte $\det(v,w)$ als *orientierter Flächeninhalt* angesehen werden, für dessen Vorzeichen gilt:

- $\det(v,w) = -\det(w,v)$.
- Drehungen in der Ebene erhalten den orientierte Flächeninhalt.
- Achsenspiegelungen ändern das Vorzeichen des orientierten Flächeninhalts.

Wir sagen auch, Drehungen seien *orientierungserhaltend* und Achsenspiegelungen seien *orientierungsumkehrend*:



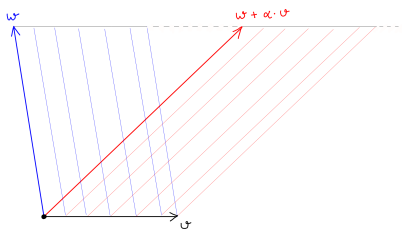
Wenn $\alpha \in \mathbb{R}$ den Winkel zwischen $v,w \in \mathbb{R}^2$ im Gegenuhrzeigersinn ist, sollte für das Vorzeichen gelten:

$$\det(v,w) \begin{cases} > 0 & \text{für } \sin(\alpha) > 0, \\ < 0 & \text{für } \sin(\alpha) < 0, \\ = 0 & \text{für } \sin(\alpha) = 0. \end{cases}$$

Mit dieser Konvention soll der Flächeninhalt *additiv* werden, d.h. für alle $u,v,w \in \mathbb{R}^2$ fordern wir:

$$\begin{aligned} \det(v+u,w) &= \det(v,w) + \det(u,w), \\ \det(v,w+u) &= \det(v,w) + \det(v,u). \end{aligned}$$

Wenn man in der zweiten Gleichung $u = -w$ wählt, versteht man, warum wir hier mit *orientierten* Flächen arbeiten. Wenn man andererseits $u = \alpha v$ mit $\alpha \in \mathbb{R}$ wählt und $\det(v,v) = 0$ nutzt, erhält man die Scherungsinvarianz des Flächeninhalts:



Durch diese Eigenschaften ist der orientierte Flächeninhalt eindeutig bestimmt:

Lemma 1.1. *Es gibt genau eine Funktion $f\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, sodass gilt:*

a) Bilinearität: Für alle $\alpha \in \mathbb{R}$ und alle $v,w,x \in \mathbb{R}^2$ ist

$$\begin{aligned} f(v,w+\alpha x) &= f(v,w)+\alpha f(v,x), \\ f(v+\alpha x,w) &= f(v,w)+\alpha f(x,w). \end{aligned}$$

b) Antisymmetrie: Für alle $v,w \in \mathbb{R}^2$ ist $f(v,w) = -f(w,v)$.

c) Normierung: Die Standardbasis hat $f(e_1,e_2) = 1$.

Beweis. Wir zeigen zunächst die Eindeutigkeit: Wenn f eine Funktion mit den drei Eigenschaften $a),b),c)$ ist, dann berechnet man für $v = (v_1,v_2), w = (w_1,w_2) \in \mathbb{R}^2$ sofort

$$\begin{aligned} f(v,w) &= f(v_1e_1+v_2e_2,w) \overset{a)}{=} \sum_{i=1,2} v_i \cdot f(e_i,w) = \sum_{i=1,2} v_i \cdot f(e_i, w_1e_1+w_2e_2) \\ &\overset{a)}{=} \sum_{i=1,2}\sum_{j=1,2} v_i \cdot w_j \cdot f(e_i,e_j). \end{aligned}$$

Wegen $b)$ und $c)$ ist dabei

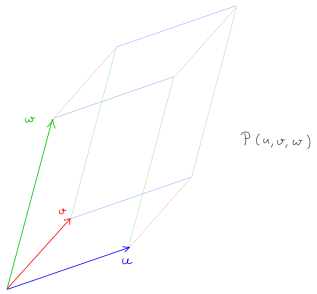
$$f(e_i,e_j) = \begin{cases} +1 & \text{für } i<j, \\ -1 & \text{für } i>j, \\ 0 & \text{für } i=j, \end{cases}$$

also $f(v,w) = v_1w_2 - v_2w_1$. Umgekehrt rechnet man direkt nach, dass der Ausdruck auf der rechten Seite der letzten Gleichung die Eigenschaften $a),b),c)$ besitzt. □

Die obige Diskussion lässt sich ohne Mühe verallgemeinern auf Volumina von Polytopen

$$\mathscr{P}(u,v,w) := \{\alpha u+\beta v+\gamma w \in \mathbb{R}^3 \mid \alpha,\beta,\gamma \in [0,1]\},$$

die von drei Vektoren $u,v,w \in \mathbb{R}^3$ im Raum aufgespannt werden:



Aber an dieser Stelle sind zunächst einige Worte über Permutationen angebracht.

2 Exkurs zu Permutationen

Für orientierte Volumina und ihre höherdimensionalen Verallgemeinerungen spielt die Reihenfolge der Variablen eine Rolle. Beim Umsortieren der Variablen hilft uns die symmetrische Gruppe

$$\mathfrak{S}_n \;=\; \Big\{\text{bijektive Abbildungen } \sigma : \{1,\dots,n\} \to \{1,\dots,n\}\Big\}.$$

Die Elemente dieser Gruppe heißen *Permutationen*. Wir schreiben sie manchmal als Wertetabellen

$$\sigma \;=\; \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} \in \mathfrak{S}_n \quad \text{mit} \quad i_\nu \;=\; \sigma(\nu).$$

Auch wenn diese Notation aussieht wie eine $2 \times n$ Matrix, hat die Verkettung $\circ$ von Permutationen nichts mit dem Matrizenprodukt zu tun. Die Verkettung $\circ$ macht $\mathfrak{S}_n$ zu einer Gruppe, diese ist für $n \geq 3$ nicht abelsch:

Beispiel 2.1. Für

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},\; \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in \mathfrak{S}_3$$

berechnet man

$$\sigma \circ \tau \;=\; \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \;=\; \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$

$$\tau \circ \sigma \;=\; \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \;=\; \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Die symmetrische Gruppe hat

$$|\mathfrak{S}_n| \;=\; n!$$

Elemente. Für $n = 60$ sind das mehr, als es Atome im beobachtbaren Universum gibt! Trotzdem können wir mit symmetrischen Gruppen sehr gut rechnen:

Definition 2.2. Eine *Transposition* ist eine Permutation, die zwei Indices vertauscht und die übrigen Indices nicht verändert. Für $i \neq j$ bezeichnen wir mit $\tau_{ij} \in \mathfrak{S}_n$ die Transposition mit

$$\tau_{ij}(k) \;:=\; \begin{cases} j & \text{für } k = i, \\ i & \text{für } k = j, \\ k & \text{sonst.} \end{cases}$$

Beispiel 2.3. Es gilt

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \;=\; \tau_{12} \circ \tau_{23} \;=\; \tau_{23} \circ \tau_{13} \;=\; \tau_{23} \circ \tau_{12} \circ \tau_{23} \circ \tau_{12}.$$

Satz 2.4. *Jede Permutation $\sigma \in \mathfrak{S}_n$ ist ein Produkt von Transpositionen.*

Beweis. Für $\sigma = id$ kann man das leere Produkt nehmen, das per Konvention das neutrale Element ist. Sei jetzt also $\sigma \neq id$. In diesem Fall können wir den kleinsten Index betrachten, der von σ bewegt wird. Wir setzen

$$i := \min\{k \mid \sigma(k) \neq k\} \quad \text{und} \quad j := \sigma(i).$$

Wegen $\sigma(k) = k$ für $k = 1, \ldots, i-1$ ist hierbei $j > i$. Sei $\tau = \tau_{ij}$ die Transposition, die i und j vertauscht. Wir setzen $\hat{\sigma} = \tau \circ \sigma$. Dann gilt

$$\hat{\sigma}(k) \;=\; \tau(\sigma(k)) \;=\; \begin{cases} \tau(k) = k & \text{für } k < i, \\ \tau(j) = i & \text{für } k = i, \\ ***** & \text{für } k > i. \end{cases}$$

Im Fall $\hat{\sigma} = id$ ist $\sigma = \tau$ und wir sind fertig. Andernfalls ist

$$\hat{i} := \min\{k \mid \hat{\sigma}(k) \neq k\} \;>\; i.$$

Per Induktion ist $\hat{\sigma}$ ein Produkt von Transpositionen. Dann gilt dies auch für σ. □

Um die Anzahl der Faktoren in einer Zerlegung als Produkt von Transpositionen zu kontrollieren, betrachten wir $\sigma \in \mathfrak{S}_n$ als Funktion

$$\sigma: \quad \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}.$$

Diese ist streng monoton wachsend nur für $\sigma = id$. Die Vertauschung benachbarter Indices zerstört die Monotonie nur an einer einzigen Stelle. Wir definieren daher:

Definition 2.5. Ein *Fehlstand* von $\sigma \in \mathfrak{S}_n$ ist ein Paar (i, j) mit der Eigenschaft $i < j$ und $\sigma(i) > \sigma(j)$. Wir definieren das *Signum* von σ durch

$$sgn(\sigma) \;=\; \begin{cases} +1 & \text{falls die Anzahl solcher Fehlstände gerade ist,} \\ -1 & \text{falls die Anzahl solcher Fehlstände ungerade ist.} \end{cases}$$

Man nennt σ eine

- *gerade Permutation* im Fall $sgn(\sigma) = +1$,
- *ungerade Permutation* im Fall $sgn(\sigma) = -1$.

Beispiel 2.6. Für $n = 3$ gilt:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \text{hat } sgn(\sigma) = -1: \quad \begin{cases} \sigma(1) > \sigma(2), \\ \sigma(1) < \sigma(3), \\ \sigma(2) < \sigma(3). \end{cases}$$

In diesem Beispiel ist σ eine Transposition. Allgemein gilt:

Lemma 2.7. *Jede Transposition $\tau \in \mathfrak{S}_n$ hat $sgn(\tau) = -1$.*

Beweis. Seien $i < j$ die von τ vertauschten Indices:

$$\tau = \begin{pmatrix} \dots\ i-1\ \ i\ \ i+1\ \dots\ \ j-1\ \ j\ \ j+1\ \dots \\ \dots\ i-1\ \ j\ \ i+1\ \dots\ \ j-1\ \ i\ \ j+1\ \dots \end{pmatrix}.$$

Die Fehlstände von τ sind genau

- das Paar (i, j),
- die m Paare (i,k) mit $i < k < j$,
- die m Paare (k, j) mit $i < k < j$,

wobei $m = j - i - 1$ ist. Insgesamt gibt es somit $2m+1$ Fehlstände. □

Lemma 2.8. *Sei $\sigma \in \mathfrak{S}_n$. Dann gilt*

$$sgn(\sigma) = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Beweis. Die Permutation σ induziert eine Bijektion zwischen

- zweielementigen Teilmengen $\{i, j\} \subset \{1, \dots, n\}$,
- zweielementigen Teilmengen $\{\sigma(i), \sigma(j)\} \subset \{1, \dots, n\}$.

Es gilt also

$$\prod_{i<j} (j - i) = \pm \prod_{i<j} (\sigma(j) - \sigma(i))$$

Links sind alle Faktoren positiv, rechts trägt jeder Fehlstand den Faktor -1 bei. □

Wir haben hier das Produkt über Indexpaare mit $i < j$ laufen lassen. Man hätte auch andere Indexpaare wählen können, solange diese alle zweielementigen Mengen von Indices genau einmal durchlaufen:

Zusatz. *Sei $I \subset \{1, \dots, n\} \times \{1, \dots, n\}$ eine Menge von Paaren von Indices mit der Eigenschaft, dass durch*

$$(i, j) \mapsto \{i, j\}$$

eine Bijektion zwischen

- *der Menge von Paaren I und*
- *der Menge der zweielementigen Teilmengen von $\{1, \dots, n\}$*

gegeben ist. Dann gilt

$$sgn(\sigma) = \prod_{(i,j)\in I} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Beweis. Analog zum Lemma. □

Satz 2.9. *Das Signum* $sgn\colon \mathfrak{S}_n \to \{\pm 1\}$ *ist ein Gruppenhomomorphismus, d.h. es ist* $sgn(\sigma\circ\tau) = sgn(\sigma)\cdot sgn(\tau)$ *für alle* $\sigma,\tau \in \mathfrak{S}_n$.

Beweis. Wir benutzen das Lemma, erweitern den Bruch und ordnen um:

$$\begin{aligned} sgn(\sigma\circ\tau) &= \prod_{i<j} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{j-i} \\ &= \prod_{i<j} \left[\frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \cdot \frac{\tau(j)-\tau(i)}{j-i} \right] \\ &= \left[\prod_{i<j} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \right] \cdot \left[\prod_{i<j} \frac{\tau(j)-\tau(i)}{j-i} \right] \end{aligned}$$

Das zweite Produkt auf der rechten Seite ist $sgn(\tau)$ nach dem Lemma. Für das erste Produkt nutzen wir den Zusatz zum Lemma mit $I = \{(\tau(i),\tau(j)) \mid i<j\}$ und erhalten

$$\prod_{i<j} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \;=\; sgn(\sigma).$$

Insgesamt folgt damit die Behauptung. □

Permutationen σ mit $sgn(\sigma) = +1$ nennt man auch *gerade* Permutationen. Sie bilden eine Untergruppe

$$\mathfrak{A}_n \;:=\; \{\sigma \in \mathfrak{S}_n \mid sgn(\sigma) = +1\} \;\subseteq\; \mathfrak{S}_n,$$

die sogenannte *alternierende Gruppe*. Aus dem obigen Satz folgt:

Korollar 2.10. *Wenn eine Permutation* $\sigma \in \mathfrak{S}_n$ *sich als ein Produkt von genau* r *Transpositionen darstellen lässt, ist*

$$sgn(\sigma) \;=\; (-1)^r.$$

Beweis. Für Transpositionen $\sigma = \tau_{ij}$ ist $sgn(\sigma) = -1$. □

Unsere bisherige Notation für Permutationen ist nicht sehr praktisch. Kürzer ist die sogenannte Zykelnotation:

Definition 2.11. Ein *Zykel der Länge* k oder *k-Zykel* ist eine Permutation $\sigma \in \mathfrak{S}_n$, die k paarweise verschiedene Indices $i_1,\ldots,i_k \in \{1,\ldots,n\}$ sukzessive aufeinander abbildet gemäß

$$i_1 \;\mapsto\; i_2 \;\mapsto\; \cdots \;\mapsto\; i_k \;\mapsto\; i_1 \quad (\text{also } \sigma(i_\alpha) = i_{\alpha+1 \bmod k})$$

und alle übrigen Indices festhält. Wir schreiben für so einen Zykel kurz

$$\sigma \;=\; (i_1, i_2, \ldots, i_k) \;\in\; \mathfrak{S}_n.$$

Beispiel 2.12. Zykel der Länge 2 sind genau die Transpositionen. Die Gruppe $\mathfrak{S}_3$ enthält genau zwei 3-Zykel:

$$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix} = (123) = (231) = (312) \quad \text{und} \quad \begin{pmatrix}1&2&3\\3&1&2\end{pmatrix} = (132) = (321) = (213).$$

Jede Permutation lässt sich auf mehrere Weisen als Produkt von Zykeln schreiben; um eine eindeutige Zerlegung zu erhalten, führen wir folgende Bedingung ein:

Definition 2.13. Zwei Zykel $\sigma = (i_1, \ldots, i_k), \tau = (j_1, \ldots, j_l) \in \mathfrak{S}_n$ heißen *disjunkt*, wenn gilt:

$$\{i_1, \ldots, i_k\} \cap \{j_1, \ldots, j_l\} = \varnothing.$$

Für disjunkte Zykel ist offenbar $\sigma \circ \tau = \tau \circ \sigma$, und es gilt:

Satz 2.14. *Jedes $\sigma \in \mathfrak{S}_n$ besitzt eine bis auf die Reihenfolge der Faktoren eindeutige Zerlegung als ein Produkt von paarweise disjunkten Zykeln der Länge ≥ 2, und diese Darstellung ist bis auf die Reihenfolge der Faktoren eindeutig.*

Beweis. Sei $\sigma \in \mathfrak{S}_n$ gegeben. Wir betrachten die Potenzen $\sigma^\nu \in \mathfrak{S}_n$ für $\nu \in \mathbb{Z}$ und definieren auf der Menge $X = \{1, 2, \ldots, n\}$ eine Relation durch

$$i \sim j \quad :\Longleftrightarrow \quad \exists \nu \in \mathbb{Z} \colon \sigma^\nu(i) = j.$$

Man sieht leicht, dass dies eine Äquivalenzrelation ist. Die Äquivalenzklassen haben die Form

$$\{\sigma^\nu(i) \mid 0 \leq \nu < k(i)\} \quad \text{mit} \quad i \in X \quad \text{und} \quad k(i) = \min\{\nu \geq 1 \mid \sigma^\nu(i) = i\},$$

und σ permutiert die Elemente einer solchen Äquivalenzklasse mittels einem Zykel der Länge $k(i)$. Da X die Vereinigung von paarweise disjunkten Äquivalenzklassen ist, erhalten wir eine Zerlegung von σ als Produkt paarweise disjunkter Zykel:

$$\sigma = (i_{11}, i_{12}, \ldots, i_{1k_1}) \circ \cdots \circ (i_{r1}, i_{r2}, \ldots, i_{rk_r}) \quad \text{mit} \quad k_\nu = k(i_{\nu 1}).$$

Dabei können wir Zykel der Länge 1 weglassen, da sie trivial sind. Wir erhalten so die Existenz der gesuchten Zerlegung. Die Eindeutigkeit folgt analog: In jeder Zerlegung von σ als Produkt paarweise disjunkter Zykel entsprechen die Zykel den Äquivalenzklassen bezüglich $\sim$ und sind daher eindeutig durch σ bestimmt. □

3 Determinantenfunktionen

Wir wollen nun die wesentlichen bei unserer Betrachtung der orientierten Fläche von Parallelogrammen benutzten Eigenschaften verallgemeinern auf Vektorräume höherer Dimension über beliebigen Körpern:

Definition 3.1. Sei V ein Vektorraum über einem Körper K. Eine Abbildung

$$f\colon \quad V^n = V \times \cdots \times V \longrightarrow K$$

heißt *multilinear*, wenn sie linear in jeder Variablen ist, wenn also die Abbildungen

$$V \longrightarrow K, \quad x \mapsto f(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n)$$

für jedes i bei jeweils fest gewählten $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \in V$ linear sind.

Definition 3.2. Eine multilineare Abbildung $f : V^n \longrightarrow K$ heißt *alternierend*, wenn gilt:

$$f(v_1, \ldots, v_n) = 0, \quad \text{wann immer } v_i = v_j \text{ für zwei Indices } i \neq j \text{ ist.}$$

Lemma 3.3. *Jede alternierende multilineare Abbildung f ist antisymmetrisch:*

$$f(\ldots, v_i, \ldots, v_j, \ldots) = -f(\ldots, v_j, \ldots, v_i, \ldots) \quad \textit{für alle } i \neq j.$$

Beweis. Wir setzen $g(v,w) = f(\ldots, v, \ldots, w, \ldots)$, wobei $\ldots$ für beliebige, aber fest gewählte Einträge steht. Dann berechnet man

$$\begin{aligned} 0 &= g(v+w, v+w) && \text{(da } f \text{ alternierend)} \\ &= g(v,v) + g(v,w) + g(w,v) + g(w,w) && \text{(da } f \text{ multilinear)} \\ &= g(v,w) + g(w,v) && \text{(da } f \text{ alternierend)} \end{aligned}$$

und somit $g(w,v) = -g(v,w)$. □

Die Umkehrung des obigen Lemmas gilt genau dann, wenn $2 \in K^\times = K \setminus \{0\}$ ist. Beispielsweise ist die Bilinearform $f\colon K^n \times K^n \to K, (x,y) \mapsto x^t \cdot y$ über dem Körper $K = \mathbb{F}_2$ wegen $-1 = 1 \in \mathbb{F}_2$ sowohl symmetrisch als auch antisymmetrisch, aber sie ist nicht alternierend im Sinn der Definition 3.2: Die Eindeutigkeitsaussage im folgenden Satz 3.5 gilt tatsächlich nur für alternierende Abbildungen.

Definition 3.4. Sei V ein Vektorraum mit $\dim_K V < \infty$. Eine *Determinantenfunktion* auf V ist eine alternierende multilineare Abbildung

$$f\colon \quad V^n \longrightarrow K \quad \text{mit} \quad n = \dim_K V.$$

Wir nennen f *nichttrivial*, wenn es ein $(v_1, \ldots, v_n) \in V^n$ gibt mit $f(v_1, \ldots, v_n) \neq 0$.

Nichttriviale Determinantenfunktionen kann man für $K = \mathbb{R}$ anschaulich sehen als ein Maß für das orientierte Volumen von Parallelotopen

$$\mathscr{P}(v_1, \ldots, v_n) = \{\alpha_1 v_1 + \cdots + \alpha_n v_n \mid 0 \leq \alpha_i \leq 1\} \subset \mathbb{R}^n$$

in Analogie zu unser Diskussion des Flächeninhaltes. Allgemein gilt:

Satz 3.5. *Für jeden endlich erzeugten Vektorraum V über K gibt es eine nichttriviale Determinantenfunktion*

$$\Delta : \quad V^n \longrightarrow K \qquad \textit{mit} \quad n = \dim_K V$$

und diese ist bis auf Multiplikation mit einer Konstanten eindeutig: Jede weitere solche hat die Form

$$f(v_1,\ldots,v_n) = \alpha \cdot \Delta(v_1,\ldots,v_n) \quad \textit{für genau ein} \quad \alpha = \alpha(f) \in K.$$

Beweis. Wir beginnen mit dem Beweis der Eindeutigkeit: Sei $f : V^n \longrightarrow K$ eine Determinantenfunktion. Wir wählen eine Basis $(e_1,\ldots,e_n)$ von V. Um $f(v_1,\ldots,v_n)$ für beliebige Vektoren $v_1,\ldots,v_n \in V$ auszurechnen, schreiben wir diese Vektoren als Linearkombination der Basisvektoren:

$$v_j = \sum_{j=1}^{n} a_{ij} \cdot e_i \quad \text{mit} \quad a_{ij} \in K.$$

Aus der Multilinearität von f folgt

$$\begin{aligned} f(v_1,\ldots,v_n) &= \sum_{i_1=1}^{n} a_{i_1,1} \cdot f(e_{i_1}, v_2, \ldots, v_n) \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{i_1,1} a_{i_2,2} \cdot f(e_{i_1}, e_{i_2}, v_3, \ldots, v_n) \\ &\vdots \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} a_{i_1,1} a_{i_2,2} \cdots a_{i_n,n} \cdot f(e_{i_1}, \ldots, e_{i_n}) \end{aligned}$$

Somit ist f durch die Werte $f(e_{i_1},\ldots,e_{i_n})$ festgelegt. Um diese Werte zu bestimmen, benutzen wir, dass f alternierend ist: Es kann also höchstens dann $f(e_{i_1},\ldots,e_{i_n}) \neq 0$ gelten, wenn $e_{i_1},\ldots,e_{i_n}$ paarweise verschiedene Vektoren sind. Dies ist genau dann der Fall, wenn das Tupel $(i_1,\ldots,i_n)$ aus $(1,\ldots,n)$ durch Umordnen entsteht, wenn es also ein $\sigma \in \mathfrak{S}_n$ gibt mit

$$i_\nu = \sigma(\nu) \quad \text{für} \quad \nu = 1,\ldots,n.$$

Also ist eine alternierende Multilinearform f eindeutig bestimmt durch die Werte

$$f(\sigma) := f(e_{\sigma(1)},\ldots,e_{\sigma(n)}) \quad \text{für} \quad \sigma \in \mathfrak{S}_n.$$

Wenn man hier σ durch $\sigma \circ \tau$ mit einer Transposition τ ersetzt, werden zwei der Variablen vertauscht und an den übrigen ändert sich nichts. Indem wir $\sigma \in \mathfrak{S}_n$ als Produkt von r Transpositionen schreiben und $sgn(\sigma) = (-1)^r$ beachten, erhalten

wir daher $f(\sigma) = sgn(\sigma) \cdot f(id)$. Für beliebige $v_j = a_{1j}e_1 + \cdots + a_{nj}e_n$ liefert die obige Formel somit

$$f(v_1, \ldots, v_n) \;=\; \det(A) \cdot f(e_1, \ldots, e_n)$$

mit der *Determinante*

$$\det(A) \;:=\; \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma) \cdot a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

der Koeffizientenmatrix $A = (a_{ij}) \in \mathrm{Mat}(n \times n, K)$. Statt f können wir in dieser Rechnung auch eine beliebige andere Determinantenfunktion Δ einsetzen. Somit gilt

$$\begin{aligned} f(v_1, \ldots, v_n) \;&=\; \det(A) \cdot f(e_1, \ldots, e_n), \\ \Delta(v_1, \ldots, v_n) \;&=\; \det(A) \cdot \Delta(e_1, \ldots, e_n). \end{aligned}$$

Für Δ nichttrivial ist $\Delta(e_1, \ldots, e_n) \neq 0$. Dann folgt $f(v_1, \ldots, v_n) = \alpha \cdot \Delta(v_1, \ldots, v_n)$ mit der Konstanten

$$\alpha \;=\; \frac{f(e_1, \ldots, e_n)}{\Delta(e_1, \ldots, e_n)} \;\in\; K.$$

Die Existenz einer nichttrivialen Determinantenfunktion ist jetzt einfach zu zeigen, da wir einen Kandidaten kennen: Wir fixieren einen Isomorphismus $\varphi : V \xrightarrow{\sim} K^n$ und definieren

$$\Delta: \quad V^n \longrightarrow K \quad \text{durch} \quad \Delta(v_1, \ldots, v_n) := \det \begin{pmatrix} | & | & & | \\ \varphi(v_1) & \varphi(v_2) & \cdots & \varphi(v_n) \\ | & | & & | \end{pmatrix}.$$

Die nächste Proposition zeigt, dass Δ die gewünschten Eigenschaften hat. □

Proposition 3.6. *Die Determinante*

$$\det: \quad \mathrm{Mat}(n \times n, K) \longrightarrow K, \quad A \;=\; (a_{ij}) \;\mapsto\; \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma) \cdot a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

besitzt die folgenden Eigenschaften:

a) $\det(A)$ *ist linear in jeder Spalte der Matrix A.*

b) $\det(A) = 0$*, falls zwei Spalten von A gleich sind.*

c) Für Dreiecksmatrizen ist die Determinante das Produkt der Diagonaleinträge:

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix} \;=\; \prod_{i=1}^{n} a_{ii}.$$

Beweis. Für obere Dreiecksmatrizen sind in der Definition von det nur Summanden zu Permutationen σ mit $\sigma(j) \leq j$ für alle j relevant. Die einzige solche Permutation ist $\sigma = id$ und somit folgt die Formel für Dreiecksmatrizen in $c)$.

Die Eigenschaft $a)$ ist klar, denn in der Determinante einer beliebigen Matrix enthält jeder Summand $sgn(\sigma) \cdot a_{\sigma(1),1} \cdots a_{\sigma(n),n}$ genau einen Faktor aus jeder Spalte der Matrix. Um $b)$ zu zeigen, seien $j \neq k$ gegeben, sodass die j-te und die k-te Spalte von A gleich sind. Sei

$$\tau \;=\; \tau_{jk} \;\in\; \mathfrak{S}_n$$

die Transposition, welche die Indices j und k vertauscht. Jede ungerade Permutation hat die Form $\sigma \circ \tau$ für genau eine gerade Permutation $\sigma \in \mathfrak{A}_n$. Wir können daher die Summanden von det nach dem Signum sortieren:

$$\det(A) \;=\; \sum_{\sigma \in \mathfrak{A}_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} - \sum_{\sigma \in \mathfrak{A}_n} a_{\sigma(\tau(1)),1} \cdots a_{\sigma(\tau(n)),n}.$$

Da die j-te und die k-te Spalte von A gleich sind, gilt für alle $\sigma \in \mathfrak{S}_n$:

$$\begin{aligned} a_{\sigma(j),j} \;&=\; a_{\sigma(\tau(k)),k}, \\ a_{\sigma(k),k} \;&=\; a_{\sigma(\tau(j)),j}. \end{aligned}$$

Da τ nur die Indices j,k bewegt, gilt für $i \notin \{j,k\}$ außerdem $a_{\sigma(i),i} = a_{\sigma(\tau(i)),i}$. Somit folgt

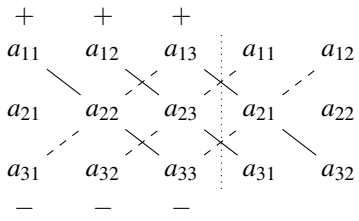
$$a_{\sigma(1),1} \cdots a_{\sigma(n),n} \;=\; a_{\sigma(\tau(1)),1} \cdots a_{\sigma(\tau(n)),n}$$

und daher erhalten wir insgesamt $\det(A) = 0$ wie gewünscht. □

Bemerkung 3.7. Die in Proposition 3.6 angegebene Formel für die Determinante einer Matrix bezeichnet man auch als die *Leibniz-Formel*. Für 3×3 Matrizen lautet sie beispielsweise

$$\det(A) \;=\; a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Diese Formel kann man sich mit der sogenannten *Regel von Sarrus* merken:



Für Determinanten größerer Matrizen gibt es keine so einfache Merkregel! Auch ist die Leibnizformel für explizite Rechnungen nicht zu empfehlen, da sie einen hohen Rechenaufwand erfordert. Besser ist es, die Matrix mit Spaltenumformungen zu vereinfachen, bis man ihre Determinante ablesen kann:

Korollar 3.8. *Seien $A, B \in \mathrm{Mat}(n \times n, K)$. Wenn B aus A hervorgeht durch...*

a) Vertauschen zweier Spalten, dann ist $\det(B) = -\det(A)$.

b) Multiplikation einer Spalte mit einem $\alpha \in K$, dann ist $\det(B) = \alpha \cdot \det(A)$.

c) Addition eines Vielfachen einer Spalte zu einer anderen, so ist $\det(B) = \det(A)$.

Beweis. Folgt sofort daraus, dass $\det(A)$ nach Proposition 3.6 eine alternierende multilineare Funktion der Spalten der Matrix A ist. □

Beispiel 3.9. Es ist

$$\det\begin{pmatrix} 6 & 7 & 4 \\ 7 & 12 & 10 \\ 3 & 6 & 6 \end{pmatrix} = 2 \cdot \det\begin{pmatrix} 6 & 7 & 2 \\ 7 & 12 & 5 \\ 3 & 6 & 3 \end{pmatrix} = 2 \cdot \det\begin{pmatrix} 4 & 3 & 2 \\ 2 & 2 & 5 \\ 0 & 0 & 3 \end{pmatrix} = 2 \cdot \det\begin{pmatrix} 1 & 3 & 2 \\ 0 & 2 & 5 \\ 0 & 0 & 3 \end{pmatrix} = 12.$$

Wir haben hier Spaltenumformungen benutzt. Die Rechenregeln im Korollar 3.8 gelten aber völlig analog auch für Zeilenumformungen, denn die Determinante einer Matrix ändert sich beim Transponieren der Matrix nicht:

Lemma 3.10. *Für $A \in \mathrm{Mat}(n \times n, K)$ gilt* $\det(A) = \det(A^t)$.

Beweis. Sei $A = (a_{ij})$ und sei $A^t = (b_{ij})$ mit $b_{ij} = a_{ji}$ die transponierte Matrix, dann gilt

$$\begin{aligned} \det(A) &= \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma) \cdot a_{\sigma(1),1} \cdots a_{\sigma(n),n} \\ &= \sum_{\sigma \in \mathfrak{S}_n} sgn(\sigma) \cdot a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)} \\ &= \sum_{\tau \in \mathfrak{S}_n} sgn(\tau) \cdot b_{\tau(1),1} \cdots b_{\tau(n),n} = \det(A^t), \end{aligned}$$

wobei wir $\tau = \sigma^{-1}$ gesetzt haben, mit $sgn(\tau) = sgn(\sigma)$. □

Eine einfache Anwendung von Determinanten ist das folgende Kriterium für die lineare Unabhängigkeit von Vektoren und die Invertierbarkeit von Matrizen:

Lemma 3.11. *Sei $f : V^n \to K$ eine nichttriviale Determinantenfunktion. Für $v_i \in V$ sind dann äquivalent:*

a) Es ist $f(v_1, \ldots, v_n) \neq 0$.

b) Die Vektoren $v_1, \ldots, v_n$ bilden eine Basis von V.

Insbesondere gilt für Matrizen $A \in \mathrm{Mat}(n \times n, K)$:

$$\det(A) \neq 0 \iff \mathrm{rk}(A) = n \iff \ker(A) = \{0\}$$

Beweis. Die Matrizenversion folgt mit $V = K^n$ und $f = \det$ aus der Vektorversion, wir zeigen letztere. Alle Werte von f sind skalare Vielfache des Wertes auf einer beliebigen Basis. Wenn $v_1, \dots, v_n$ eine Basis bilden, gilt also:

$$f(v_1, \dots, v_n) = 0 \quad \Longrightarrow \quad f \text{ ist trivial.}$$

Für f nichttrivial muß daher $f(v_1, \dots, v_n) \neq 0$ sein. Wenn andererseits $v_1, \dots, v_n$ keine Basis von V bilden, dann sind sie linear abhängig. Also ist ein Vektor v_i eine Linearkombination der übrigen. Sei etwa

$$v_i = \sum_{j \neq i} \alpha_j v_j,$$

dann folgt aus der Multilinearität

$$f(v_1, \dots, v_n) = \sum_{j \neq i} \alpha_j f(\dots, v_i, \dots, v_{j-1}, v_i, v_{j+1}, \dots).$$

In jedem Summand auf der rechten Seite wird für zwei Variablen derselbe Vektor v_i eingesetzt. Da f alternierend ist, sind somit alle Summanden Null. □

4 Beispiele von Determinanten

Wir haben gesehen, dass Determinanten von Dreiecksmatrizen leicht zu berechnen sind. Ähnlich können wir für Blockdreiecksmatrizen vorgehen:

Lemma 4.1. *Für obere Blockdreiecksmatrizen mit Blöcken $A_{ij} \in \mathrm{Mat}(n_i \times n_j, K)$ gilt:*

$$\det \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} = \det(A_{11}) \cdot \det(A_{22}).$$

Beweis. Durch Zeilenumformungen in den ersten n_1 Zeilen und unabhängig davon in den letzten n_2 Zeilen können wir annehmen, dass

$$A_{11} = \begin{pmatrix} d_1 & \cdots & * \\ & \ddots & \vdots \\ & & d_{n_1} \end{pmatrix} \quad \text{und} \quad A_{22} = \begin{pmatrix} d_{n_1+1} & \cdots & * \\ & \ddots & \vdots \\ & & d_n \end{pmatrix}$$

in Dreiecksform sind. Unsere Blockdreiecksmatrix wird damit eine tatsächliche oberen Dreiecksmatrix, und aus ihren Diagonaleinträgen liest man direkt ab, dass $\det(A) = d_1 \cdots d_n = (d_1 \cdots d_{n_1}) \cdot (d_{n_1+1} \cdots d_n) = \det(A_{11}) \det(A_{22})$ ist. □

Besonders häufig wird das obige Lemma verwendet im Fall $n_1 = 1$. Hier besagt es

$$\det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix} = a_{11} \cdot \det \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Oft bricht man z.B. den Gauß-Algorithmus nach Umformung der ersten Spalte ab und nutzt dann die obige Formel. Ein interessantes Beispiel für Determinanten, die sich so berechnen lassen, tritt in *Interpolationsproblemen* auf:

Satz 4.2. *Für $n \in \mathbb{N}$ seien $a_1, \ldots, a_n \in K$ paarweise verschieden und $y_1, \ldots, y_n \in K$ beliebig. Dann gibt es genau ein Polynom*

$$f(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$$

vom Grad $\leq n-1$ mit Koeffizienten $c_i \in K$, sodass $f(a_i) = y_i$ für alle i gilt.

Beweis. Die Bedingung $f(a_i) = y_i$ für alle i ist äquivalent zu

$$\begin{aligned} c_0 + a_1 c_1 + a_1^2 c_2 + \cdots + a_1^{n-1} c_{n-1} &= y_1 \\ &\vdots \\ c_0 + a_n c_1 + a_n^2 c_2 + \cdots + a_n^{n-1} c_{n-1} &= y_n \end{aligned}$$

Das ist ein inhomogenes LGS in den Variablen $c_1, \ldots, c_n$. Es hat eine eindeutige Lösung, falls die Koeffizientenmatrix invertierbar ist. Die Invertierbarkeit prüfen wir, indem wir die Determinante

$$V(a_1, \ldots, a_n) := \det \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix}.$$

ausrechnen. Diese sogenannte *Vandermonde-Determinante* ist nach dem folgenden Lemma von Null verschieden, da die a_i paarweise verschieden sind. □

Lemma 4.3 (Vandermonde-Determinante). *Es ist $V(a_1, \ldots, a_n) = \prod_{i<j} (a_j - a_i)$.*

Beweis. Für $n = 1$ ist nichts zu zeigen. Sei $n > 1$. Wir subtrahieren in der oben betrachteten Matrix

- von der n-ten Spalte das a_1-fache der $(n-1)$-ten,
- von der $(n-1)$-ten Spalte das a_1-fache der $(n-2)$-ten,
- …

Damit wird $V = V(a_1,\dots,a_n)$ umgeformt zu

$$V \;=\; \det\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & a_2-a_1 & \cdots & (a_2-a_1)a_2^{n-2} \\ 1 & a_3-a_1 & \cdots & (a_3-a_1)a_3^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & a_n-a_1 & \cdots & (a_n-a_1)a_n^{n-2} \end{pmatrix} \;=\; \det\begin{pmatrix} a_2-a_1 & \cdots & (a_2-a_1)a_2^{n-2} \\ a_3-a_1 & \cdots & (a_3-a_1)a_3^{n-2} \\ \vdots & \ddots & \vdots \\ a_n-a_1 & \cdots & (a_n-a_1)a_n^{n-2} \end{pmatrix}$$

In der letzten Determinante sind

- alle Einträge der ersten Zeile Vielfache von a_2-a_1,
- alle Einträge der zweiten Zeile Vielfache von a_3-a_1,
- $\cdots$

Indem wir diese skalaren Faktoren aus den Zeilen herausziehen, bekommen wir

$$V(a_1,\dots,a_n) \;=\; V(a_2,\dots,a_n)\cdot\prod_{i=2}^{n}(a_i-a_1)$$

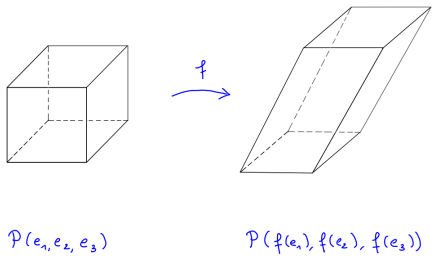
und die Behauptung folgt per Induktion über n. □

Beispiel 4.4. Es gibt eine unendliche Teilmenge $S \subset \mathbb{R}^n$ mit der Eigenschaft, dass jede n-elementige Teilmenge dieser Menge eine Basis bildet: Man wähle

$$S \;=\; \{(1,a,a^2,\dots,a^{n-1}) \mid a \in A\} \quad \text{für eine unendliche Teilmenge } A \subseteq \mathbb{R}.$$

5 Multiplikativität der Determinante

Für $A \in \mathrm{Mat}(n\times n,\mathbb{R})$ hatten wir uns $\det(A)$ vorgestellt als orientiertes Volumen des von den Spalten von A aufgespannten Parallelotops. Dieses Parallelotop ist das Bild des Einheitswürfels unter der linearen Abbildung $f = f_A : \mathbb{R}^n \to \mathbb{R}^n, v \mapsto A\cdot v$:



Also ist $\det(A)$ anschaulich der Dehnungsfaktor, mit dem die lineare Abbildung f_A Volumina reskaliert. Für die Zusammensetzung

$$f_B \circ f_A : \quad \mathbb{R}^n \xrightarrow{f_A} \mathbb{R}^n \xrightarrow{f_B} \mathbb{R}^n$$

von Endomorphismen sollten sich ihre Dehnungsfaktoren multiplizieren. In der Tat gilt über jedem Körper K die folgende Formel:

Satz 5.1. *Für alle $A, B \in \mathrm{Mat}(n \times n, K)$ gilt* $\det(AB) = \det(A)\det(B)$.

Beweis. Mit der Leibniz-Formel sollte man das besser nicht nachrechnen. Ganz ohne Rechnen geht es mit abstrakten Determinantenfunktionen: Aus der Definition des Matrizenproduktes sieht man, dass Spaltenoperationen auf B ebensolche auf AB induzieren. Für festes A ist somit $B \mapsto \det(AB)$ eine alternierende multilineare Abbildung in den Spalten von B. Die Eindeutigkeit von Determinantenfunktionen bis auf Skalare liefert eine Konstante $c = c_A \in K$ mit

$$\det(AB) = c\det(B) \quad \text{für alle} \quad B \in \mathrm{Mat}(n \times n, K).$$

Speziell für $B = \mathbf{1}$ liest man $c = \det(A)$ ab. □

Korollar 5.2. *Die Determinante ist ein Gruppenhomomorphismus*

$$\det : \quad \mathrm{GL}_n(K) \longrightarrow K^\times$$

Insbesondere gilt $\det(A^{-1}) = \det(A)^{-1}$ *für alle* $A \in \mathrm{GL}_n(K)$.

Beweis. Folgt sofort aus der im Satz bewiesenen Multiplikativität von det. □

Der Kern dieses Gruppenhomomorphismus bildet eine Untergruppe von $\mathrm{GL}_n(K)$, die sogenannte *spezielle lineare Gruppe*

$$\mathrm{SL}_n(K) \;=\; \{A \in \mathrm{GL}_n(K) \mid \det(A) = 1\} \;\subset\; \mathrm{GL}_n(K).$$

6 Laplace-Entwicklung

Statt mit Zeilen- und Spaltentransformationen kann man det auch direkt rekursiv berechnen. Dabei nutzt man zunächst nur die Linearität in einer Spalte:

Beispiel 6.1. Sei $A = (a_{ij}) \in \mathrm{Mat}(3 \times 3, K)$. Wegen der Linearität von det in der ersten Spalte ist

$$\det(A) \;=\; a_{11} \cdot \det \begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} \;+\; a_{21} \cdot \det \begin{pmatrix} 0 & a_{12} & a_{13} \\ 1 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} \;+\; a_{31} \cdot \det \begin{pmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 1 & a_{32} & a_{33} \end{pmatrix}$$

Die erste der Matrizen auf der rechten Seite hat Blockdreiecksform. Die anderen beiden erhalten nach Zeilenvertauschung Blockdreiecksform. Somit folgt

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

Das ist ein Spezialfall des folgenden Resultats:

Satz 6.2 (Laplace-Entwicklung). *Für $A=(a_{ij})\in \mathrm{Mat}(n\times n,K)$ bezeichne A_{ij} die aus der Matrix A durch Streichen der i-ten Zeile und der j-ten Spalte erhaltene Matrix. Dann gilt:*

a) Für jedes $j\in\{1,\dots,n\}$ kann man nach der j-ten Spalte entwickeln:

$$\det(A) \;=\; \sum_{i=1}^{n}(-1)^{i+j}\,a_{ij}\,\det(A_{ij})$$

b) Für jedes $i\in\{1,\dots,n\}$ kann man nach der i-ten Zeile entwickeln:

$$\det(A) \;=\; \sum_{j=1}^{n}(-1)^{i+j}\,a_{ij}\,\det(A_{ij})$$

Beweis. Wegen $\det(A)=\det(A^t)$ reicht es, Teil $a)$ zu beweisen. Indem wir die j-te Spalte von A schreiben als

$$a_j \;=\; \sum_{i=1}^{n} a_{ij}e_i,$$

erhalten wir aus der Linearität von det in der j-ten Spalte

$$\det(A) \;=\; \sum_{i=1}^{n} a_{ij}\,\det\begin{pmatrix} & | & | & | & \\ \cdots & a_{j-1} & e_i & a_{j+1} & \cdots \\ & | & | & | & \end{pmatrix}$$

Zu zeigen bleibt daher nur

$$\det\begin{pmatrix} & | & | & | & \\ \cdots & a_{j-1} & e_i & a_{j+1} & \cdots \\ & | & | & | & \end{pmatrix} \;=\; (-1)^{i+j}\,\det(A_{ij}).$$

Durch Zeilen- und Spaltenvertauschungen reduziert man diese letzte Behauptung auf den Fall $i=j=1$. In diesem Fall hat man eine obere Blockdreiecksmatrix und die Aussage ist klar. □

Definition 6.3. Die zur Matrix A *komplementäre* Matrix ist die Matrix

$$A^* \;=\; (a^*_{ij}) \;\in\; \mathrm{Mat}(n\times n,K) \quad \text{mit} \quad a^*_{ij} \;:=\; (-1)^{i+j}\cdot\det(A_{ji}).$$

Man beachte die Vertauschung von i und j auf der rechten Seite! Wir erhalten:

Korollar 6.4 (Cramer'sche Formel). *Es ist* $A \cdot A^* = A^* \cdot A = \det(A) \cdot \mathbf{1}$.

Beweis. Die Einträge des Matrizenproduktes $A^* \cdot A = (c_{ik})$ erhält man per Definition als

$$c_{ik} \;=\; \sum_{j=1}^{n} a_{ij}^* a_{jk} \;=\; \sum_{j=1}^{n} (-1)^{i+j} \det(A_{ji}) \cdot a_{jk} \;=\; \det \begin{pmatrix} & | & | & | & \\ \cdots & a_{i-1} & a_k & a_{i+1} & \cdots \\ & | & | & | & \end{pmatrix}$$

Die Determinante auf der rechten Seite ist gleich $\det(A)$ im Fall $i = k$, und Null im Fall $i \neq k$. Also ist $A^* \cdot A = \det(A) \cdot \mathbf{1}$, analog $A \cdot A^* = \det(A) \cdot \mathbf{1}$. □

Wenn also $A \in \mathrm{Mat}(n \times n, K)$ invertierbar ist, erhält man die inverse Matrix aus der Formel

$$A^{-1} \;=\; \frac{1}{\det(A)} \cdot A^*.$$

Zum Rechnen ist der Gauß-Algorithmus sehr viel effizienter. Aber die Cramer'sche Formel ist für theoretische Argumente gut:

Beispiel 6.5. Für $A = (a_{ij}) \in \mathrm{GL}_n(\mathbb{R})$ zeigt ein Blick auf die Cramer'sche Formel, dass die Einträge der Inversen A^{-1} stetige Funktionen der Matrixeinträge a_{ij} sind.

Beispiel 6.6. Die Teilmenge

$$\mathrm{SL}_n(\mathbb{Z}) \;:=\; \{M \in \mathrm{Mat}(n \times n, \mathbb{Z}) \mid \det(M) = 1\} \;\subseteq\; \mathrm{SL}_n(\mathbb{R})$$

ist eine Untergruppe, denn die Cramer'sche Formel zeigt, dass sie abgeschlossen unter der Inversion von Matrizen ist.

Beispiel 6.7. Im Baby-Fall $n = 2$ ist die Cramer'sche Formel sogar zum Rechnen gut:

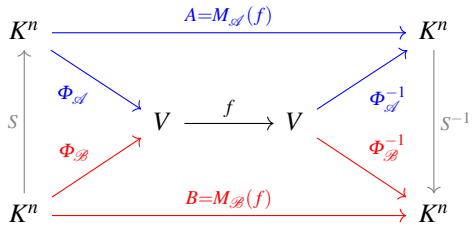
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \;=\; \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad \text{für} \quad ad - bc \;\neq\; 0.$$

Kapitel VI
Eigenwerte und Diagonalisierbarkeit

Zusammenfassung Die Kunst in der linearen Algebra besteht darin, komplizierte Probleme durch geschickte Koordinatenwahl trivial zu machen. In diesem Kapitel überlegen wir uns, unter welchen Bedingungen es zu einem Endomorphismus eine Basis gibt, worin er durch eine Diagonalmatrix gegeben ist. Der Schlüssel dazu sind Eigenwerte, die als Nullstellen des charakteristischen Polynoms berechnet werden können und in unzähligen Anwendungen eine Rolle spielen.

1 Eigenwerte und Eigenvektoren

Sei $f\colon V \to V$ ein Endomorphismus eines endlich-dimensionalen Vektorraumes über einem Körper K. Wie kann man eine Basis finden, in der die Abbildungsmatrix von f möglichst einfache Gestalt hat? Die Abbildungsmatrizen $M_{\mathscr{A}}(f), M_{\mathscr{B}}(f)$ für verschiedene Basen $\mathscr{A}, \mathscr{B}$ hängen durch einen Basiswechel zusammen:



Definition 1.1. Zwei Matrizen $A, B \in \mathrm{Mat}(n \times n, K)$ heißen zueinander *ähnlich*, wenn gilt:

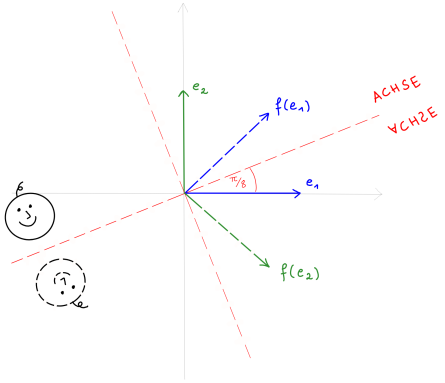
$$B = S^{-1}AS \quad \text{für ein} \quad S \in \mathrm{GL}_n(K).$$

Offenbar ist die Ähnlichkeit eine Äquivalenzrelation auf $\mathrm{Mat}(n \times n, K)$ ist. Wir wollen in jeder Äquivalenzklasse quadratischer Matrizen einen möglichst einfachen Repräsentanten finden, wie im folgenden Beispiel:

Beispiel 1.2. Sei $K = \mathbb{R}$ und $V = \mathbb{R}^2$ die reelle Ebene. Der Endomorphismus

$$f\colon \quad V \longrightarrow V, \quad v \mapsto A \cdot v \qquad \text{mit} \qquad A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

ist eine Spiegelung an der um $\pi/8$ geneigten Achse:



Denn eine direkte Rechnung zeigt

$$M_{\mathscr{B}}(f) = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{für} \quad \mathscr{B} = \left(\begin{pmatrix} 1 \\ t \end{pmatrix}, \begin{pmatrix} -t \\ 1 \end{pmatrix} \right), \quad t = \tan(\tfrac{\pi}{8}) = \sqrt{2} - 1.$$

Definition 1.3. Sei V ein endlich-dimensionaler Vektorraum über K. Wir nennen einen Endomorphismus $f \in \mathrm{End}_K(V)$ *diagonalisierbar*, wenn es eine Basis $\mathscr{B}$ von V gibt, in der seine Abbildungsmatrix eine Diagonalmatrix ist:

$$M_{\mathscr{B}}(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \text{mit} \quad \lambda_1, \ldots, \lambda_n \in K.$$

Eine Matrix $A \in \mathrm{Mat}(n \times n, K)$ heißt *diagonalisierbar*, wenn sie ähnlich zu einer Diagonalmatrix ist, wenn also der Endomorphismus $v \mapsto A \cdot v$ diagonalisierbar ist.

Ein Endomorphismus $f : V \longrightarrow V$ wird bezüglich einer Basis $\mathscr{B} = (v_1, \ldots, v_n)$ dargestellt durch eine Diagonalmatrix mit den Diagonaleinträgen $\lambda_1, \ldots, \lambda_n$ genau dann, wenn $f(v_i) = \lambda_i \cdot v_i$ für alle i ist. Diese Eigenschaft hat einen Namen:

Definition 1.4. Ein Skalar $\lambda \in K$ heißt ein *Eigenwert* von $f \in \mathrm{End}_K(V)$, wenn ein Vektor $v \in V \setminus \{0\}$ existiert mit

$$f(v) = \lambda \cdot v.$$

Wir nennen dann v einen *Eigenvektor* zum Eigenwert λ. Analoge Begriffe benutzen wir für quadratische Matrizen mittels der Identifikation $\mathrm{Mat}(n \times n, K) = \mathrm{End}_K(K^n)$.

Bemerkung 1.5. Es gilt:

a) Der Nullvektor $v = 0$ ist per Definition *kein* Eigenvektor.

b) Der Skalar $\lambda = 0$ ist ein Eigenwert von f genau für $\ker(f) \neq \{0\}$.

c) Wenn v ein Eigenvektor von f zum Eigenwert $\lambda \in K$ ist, dann ist auch $\alpha \cdot v$ einer für jedes $\alpha \in K^\times$ wegen

$$f(\alpha \cdot v) \;=\; \alpha \cdot f(v) \;=\; \alpha \cdot \lambda \cdot v \;=\; \lambda \cdot \alpha \cdot v.$$

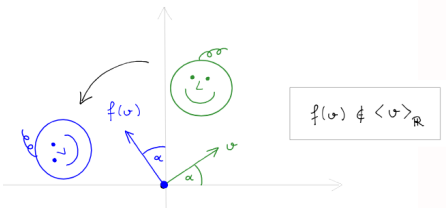
d) Sei $\dim_K(V) < \infty$. Für $f \in \mathrm{End}_K(V)$ und Basen $\mathscr{B}$ von V sind äquivalent:

- Die Basis $\mathscr{B}$ besteht aus Eigenvektoren von f.
- Die Abbildungsmatrix $M_{\mathscr{B}}(f)$ ist eine Diagonalmatrix.

Beispiel 1.6. Eigenwerte gibt es nicht immer: Sei etwa $K = \mathbb{R}$ und sei $V = \mathbb{R}^2$ die reelle Ebene, dann hat die Drehung

$$f: \quad V \longrightarrow V, \quad v \mapsto A \cdot v \quad \text{mit} \quad A \;=\; \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

keine reellen Eigenwerte. Anschaulich sollte das klar sein:



Um es formal nachzurechnen, beachte man: Für $v \in \mathbb{R}^2$ ist $f(v) = \lambda v$ äquivalent zu dem LGS

$$B \cdot v \;=\; 0 \quad \text{mit} \quad B \;=\; \begin{pmatrix} \lambda & 1 \\ -1 & \lambda \end{pmatrix}$$

Für $\lambda \in \mathbb{R}$ ist $\det(B) = \lambda^2 + 1 > 0$ und das LGS hat dann nur $v = 0$ als Lösung, es gibt also keine reellen Eigenwerte. Man beachte: Über den komplexen Zahlen ist das anders, hier ist

$$A \cdot v \;=\; \pm i \cdot v \quad \text{für} \quad v \;=\; \begin{pmatrix} 1 \\ \mp i \end{pmatrix}.$$

Beispiel 1.7. Es kann auch sehr viele Eigenwerte geben: Sei etwa $V = C^\infty(\mathbb{R})$ der reelle Vektorraum der unendlich oft differenzierbaren Funktionen. Die Ableitung von Funktionen definiert einen Endomorphismus

$$V \longrightarrow V, \quad f(x) \mapsto \tfrac{d}{dx} f(x)$$

Dieser hat jede reelle Zahl $\lambda \in \mathbb{R}$ als Eigenwert, mit Eigenvektor $f(x) = \exp(\lambda x)$.

Für endlich-dimensionale Vektorräume kann es aber zu jedem Endomorphismus nur endlich viele Eigenwerte geben:

Satz 1.8. *Sei V ein beliebiger Vektorraum über K und $f \in \mathrm{End}_K(V)$. Dann ist jedes System von Eigenvektoren*

$$v_1, \ldots, v_n \ \in \ V \setminus \{0\}$$

zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_n$ linear unabhängig.

Beweis. Für $n = 1$ ist nichts zu zeigen. Sei also $n > 1$. Per Induktion seien $v_2, \ldots, v_n$ linear unabhängig. Seien $\alpha_1, \ldots, \alpha_n \in K$ mit $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$. Anwenden der linearen Abbildung f liefert $\alpha_1 \lambda_1 v_1 + \cdots + \alpha_n \lambda_n v_n = 0$. Wir ziehen von der letzten Gleichung das λ_1-fache der vorletzten ab:

$$\alpha_2(\lambda_2 - \lambda_1)v_2 + \cdots + \alpha_n(\lambda_n - \lambda_1)v_n \ = \ 0.$$

Da nach Induktionsannahme $v_2, \ldots, v_n$ linear unabhängig sind, folgt

$$\alpha_2(\lambda_2 - \lambda_1) \ = \ \cdots \ = \ \alpha_n(\lambda_n - \lambda_1) \ = \ 0.$$

Wegen $\lambda_i \neq \lambda_1$ für alle i folgt $\alpha_2 = \cdots = \alpha_n = 0$ und damit auch $\alpha_1 = 0$. □

Korollar 1.9. *Sei* $\dim_K(V) = n < \infty$. *Für $f \in \mathrm{End}_K(V)$ gilt dann:*

a) Es gibt höchstens n verschiedene Eigenwerte von f.

b) Wenn es n verschiedene Eigenwerte gibt, ist f diagonalisierbar.

Beweis. Nach dem vorigen Satz sind Eigenvektoren zu paarweise verschiedenen Eigenwerten linear unabhängig. Jedes linear unabhängige System in V besteht aus höchstens $\dim V$ Vektoren, mit Gleichheit nur für Basen. □

Man beachte, dass die Umkehrung von b) nicht korrekt ist: Die Einheitsmatrix ist diagonalisierbar, aber sie besitzt als einzigen Eigenwert $\lambda = 1$. Um Eigenwerte mit Vielfachheiten zu behandeln, machen wir folgende

Definition 1.10. Der *Eigenraum* von $f \in \mathrm{End}_K(V)$ zum Eigenwert $\lambda \in K$ ist der Untervektorraum

$$E(f, \lambda) \ := \ \ker(f - \lambda \cdot id_V) \ = \ \{v \in V \mid f(v) = \lambda v\}.$$

Die von Null verschiedenen Elemente des Eigenraumes $E(f, \lambda)$ sind also genau die Eigenvektoren von f zum Eigenwert λ. Es gilt:

Lemma 1.11. *Für paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n$ von $f \in \mathrm{End}_K(V)$ ist die Summe der zugehörigen Eigenräume direkt:*

$$V' \ = \ E(f, \lambda_1) \oplus \cdots \oplus E(f, \lambda_n) \ \subseteq \ V$$

Beweis. Für Vektoren $v_i \in E(f, \lambda_i)$ kann nach Satz 1.8 nur dann $v_1 + \cdots + v_n = 0$ gelten, wenn $v_1 = \cdots = v_n = 0$ ist. Die Summe der $E(f, \lambda_i) \subset V$ ist also direkt. □

Korollar 1.12. *Für $f \in \mathrm{End}_K(V)$ sind äquivalent:*

a) Es ist f diagonalisierbar.

b) Es ist V die (direkte) Summe seiner Eigenräume.

Beweis. Wenn f diagonalisierbar ist, hat V eine Basis aus Eigenvektoren von f und wird somit aufgespannt von Eigenräumen. Gilt umgekehrt letzteres, so wähle für jedes $E(f,\lambda_i)$ eine Basis. Nach Lemma 1.11 ist die Vereinigung dieser Basen linear unabhängig, also eine Basis, und in dieser Basis hat f Diagonalgestalt. □

2 Das charakteristische Polynom

Wie findet man die Eigenvektoren eines Endomorphismus $f \in \mathrm{End}_K(V)$? Hierzu nehmen wir zunächst $V = K^n$ an. Dann ist f bezüglich der Standardbasis gegeben durch eine Matrix

$$A \;=\; (a_{ij}) \;\in\; \mathrm{Mat}(n\times n, K).$$

Wenn man einen Eigenwert $\lambda \in K$ kennt, erhält man die zugehörigen Eigenvektoren als die von Null verschiedenen Lösungen von $Av = \lambda v$. Es gilt:

$$Av \;=\; \lambda v \iff \lambda v - Av \;=\; 0 \iff (\lambda \mathbf{1} - A)v \;=\; 0 \iff v \;\in\; \ker(\lambda \mathbf{1} - A).$$

Für jeden festen Eigenwert λ ist das ein LGS mit der Koeffizientenmatrix

$$\lambda \mathbf{1} - A \;=\; -\begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{pmatrix}$$

Um die Eigenwerte λ zu finden, kann man in einfachen Fällen wie folgt vorgehen:

Satz 2.1. *Es ist $\lambda \in K$ ein Eigenwert von A genau für* $\det(\lambda \mathbf{1} - A) = 0$.

Beweis. Es gilt:

$$\begin{aligned} \lambda \text{ ist Eigenwert von } A &\iff \exists v \neq 0 :\; Av \;=\; \lambda v \\ &\iff \exists v \neq 0 :\; (\lambda \mathbf{1} - A)v \;=\; 0 \\ &\iff \ker(\lambda \mathbf{1} - A) \neq \{0\} \\ &\iff \mathrm{rk}(\lambda \mathbf{1} - A) < n \\ &\iff \det(\lambda \mathbf{1} - A) \;=\; 0 \end{aligned}$$

□

Beispiel 2.2. Für die bereits bekannte Matrix

$$A \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \;\in\; \mathrm{Mat}(2\times 2,\mathbb{R})$$

berechnet man

$$\det(\lambda\mathbf{1}-A) \;=\; \det\begin{pmatrix} \lambda-\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \lambda+\frac{1}{\sqrt{2}} \end{pmatrix} \;=\; \lambda^2-1$$

Diese Matrix besitzt somit wie erwartet genau die beiden Eigenwerte $\lambda=\pm 1$.

Hier ist $\det(\lambda\mathbf{1}-A)=\lambda^2-1$ ein Polynom in λ. Genauer sollten wir eigentlich sagen, eine Polynomfunktion: Wir hatten *Polynome* definiert als formale Summen der Gestalt

$$f(t) \;=\; a_0+a_1t+a_2t^2+\cdots+a_nt^n \;\in\; K[t]$$

mit $a_i\in K$. Dabei war t nur eine formale Variable! Erst durch Einsetzen konkreter Werte für t erhalten wir die *Polynomfunktion*

$$K \;\longrightarrow\; K, \quad \lambda \;\mapsto\; f(\lambda) \;:=\; a_0+a_1\lambda+\cdots+a_n\lambda^n$$

wobei rechts nun eine echte Summe in K steht. Ein warnendes Beispiel:

Beispiel 2.3. Sei p eine Primzahl und $K=\mathbb{F}_p$. Dann ist $f(t)=t^p-t\in K[t]$ nicht das Nullpolynom, aber

$$f(\lambda) \;=\; \lambda^p-\lambda \;=\; 0 \quad \text{für alle} \quad \lambda \;\in\; \mathbb{F}_p.$$

So etwas passiert nur über endlichen Körpern, aber wir wollen hier klar zwischen Polynomen und Polynomfunktionen unterscheiden. Glücklicherweise lässt sich die Leibnizformel über beliebigen kommutativen Ringen lesen:

Definition 2.4. Sei $n\in\mathbb{N}$. Für eine Matrix $A=(a_{ij})\in\mathrm{Mat}(n\times n,R)$ mit Einträgen in einem beliebigen kommutativen Ring R definieren wir ihre *Determinante* durch die Formel

$$\det(A) \;:=\; \sum_{\sigma\in\mathfrak{S}_n} sgn(\sigma)a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}.$$

Derselbe Beweis wie über Körpern zeigt:

- $\det(A)$ ist multilinear und alternierend in den Spalten.
- $\det(A)=\det(A^t)$ und $\det(AB)=\det(A)\det(B)$.
- Cramer'sche Formel $A\cdot A^*=A^*\cdot A=\det(A)\cdot\mathbf{1}$.

Wir können nun das charakteristische Polynom als Determinante einer Matrix mit Einträgen in dem Polynomring $R=K[t]$ definieren:

Definition 2.5. Das *charakteristische Polynom* von $A = (a_{ij}) \in \mathrm{Mat}(n \times n, K)$ ist die Determinante

$$\chi_A(t) \;:=\; \det\bigl(t \cdot \mathbf{1} - A\bigr) \;=\; \det \begin{pmatrix} t - a_{11} & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & t - a_{nn} \end{pmatrix} \;\in\; K[t].$$

Aus unseren Rechenregeln für Determinanten folgt sofort:

Lemma 2.6. *Das charakteristische Polynom einer Matrix stimmt überein mit dem ihrer transponierten Matrix. Für obere Dreiecksmatrizen $A = (a_{ij})$ mit $a_{ij} = 0$ für alle $i > j$ gilt*

$$\chi_A(t) \;=\; (t - a_{11})(t - a_{22}) \cdots (t - a_{nn}).$$

Beweis. Die erste Aussage folgt aus $(t \cdot \mathbf{1} - A)^t = t \cdot \mathbf{1} - A^t$ und der Invarianz der Determinante unter Transposition. Die zweite folgt daraus, dass mit A auch $t \cdot \mathbf{1} - A$ eine Dreiecksmatrix ist mit Diagonaleinträgen $t - a_{ii}$. □

Einige wenige Koeffizienten des charakteristischen Polynoms kann man auch für beliebige Matrizen schnell ablesen. Wir definieren dazu die *Spur* (engl. *trace*) einer Matrix $A = (a_{ij}) \in \mathrm{Mat}(n \times n, K)$ durch

$$\mathrm{tr}(A) \;:=\; \sum_{i=1}^{n} a_{ii},$$

also die Summe der Diagonaleinträge. Dann gilt:

Lemma 2.7. *Für jede Matrix $A = (a_{ij}) \in \mathrm{Mat}(n \times n, K)$ ist*

$$\chi_A(t) \;=\; t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0 \quad \textit{mit} \quad \begin{cases} c_{n-1} &= -\mathrm{tr}(A), \\ c_0 &= (-1)^n \cdot \det(A). \end{cases}$$

Beweis. Nach der Leibniz-Formel ist das charakteristische Polynom $\chi_A(t) \in K[t]$ die Summe der Polynome

$$p_\sigma(t) \;=\; sgn(\sigma) \cdot \prod_{i=1}^{n} b_{\sigma(i),i}(t) \quad \text{mit} \quad b_{\sigma(i),i}(t) \;=\; \begin{cases} t - a_{i,i} & \text{für } \sigma(i) = i, \\ -a_{\sigma(i),i} & \text{sonst.} \end{cases}$$

Es gilt $\deg(p_\sigma) = \#\{i \mid \sigma(i) = i\} \leq n - 2$ für alle $\sigma \neq id$. Die führenden beiden Terme des charakteristischen Polynoms stimmen also überein mit denen von

$$\begin{aligned} p_{id}(t) \;&=\; b_{11}(t) b_{22}(t) \cdots b_{nn}(t) \\ &=\; (t - a_{11})(t - a_{22}) \cdots (t - a_{nn}) \\ &=\; t^n - (a_{11} + a_{22} + \cdots + a_{nn}) \cdot t^{n-1} + \text{Terme vom Grad} \leq n - 2. \end{aligned}$$

Es bleibt nur der konstante Term von $\chi_A(t)$ zu berechnen. Den konstanten Term eines Polynoms erhält man durch Auswerten des Polynoms in $t = 0$. In unserem Fall ist

$$\begin{aligned}\chi_A(0) &= \sum_{\sigma\in\mathfrak{S}_n} sgn(\sigma)\cdot b_{\sigma(1),1}(0)\cdots b_{\sigma(n),n}(0)\\ &= \sum_{\sigma\in\mathfrak{S}_n} sgn(\sigma)\cdot(-1)^n\cdot a_{\sigma(1),1}\cdots a_{\sigma(n),n} \;=\; (-1)^n\cdot\det(A)\end{aligned}$$

wegen $b_{ij}(0) = -a_{ij}$ und der Leibniz-Formel für $\det(A)$. □

Beispiel 2.8. Für das charakteristische Polynom von 2×2 Matrizen gibt das Lemma die Formel

$$\chi_A(t) \;=\; t^2-(a+d)t+(ad-bc) \quad \text{für} \quad A \;=\; \begin{pmatrix} a & b\\ c & d\end{pmatrix} \;\in\; \mathrm{Mat}(2\times 2,K).$$

Allgemein können wir charakteristische Polynome leicht mit den Rechenregeln für Determinanten aus dem vorigen Kapitel finden. Diese Regeln zeigen auch, dass ähnliche Matrizen dasselbe charakteristische Polynom besitzen:

Lemma 2.9. *Für $B = SAS^{-1}$ mit $B \in \mathrm{GL}_n(K)$ gilt $\chi_A(t) = \chi_B(t)$. Insbesondere ist also*

$$\det(A) \;=\; \det(B) \quad \textit{und} \quad \mathrm{tr}(A) \;=\; \mathrm{tr}(B).$$

Beweis. Aus den Rechenregeln für Determinanten über kommutativen Ringen folgt für $B = SAS^{-1}$ sofort

$$\begin{aligned}\chi_B(t) \;=\; \det(t\cdot\mathbf{1}-SAS^{-1}) \;&=\; \det(S(t\cdot\mathbf{1}-A)S^{-1})\\ &=\; \det(S)\det(t\cdot\mathbf{1}-A)\det(S)^{-1}\\ &=\; \det(t\cdot\mathbf{1}-A) \;=\; \chi_A(t).\end{aligned}$$

Die Behauptung für $\det(B)$ und $\mathrm{tr}(B)$ folgt hieraus. □

Dass ähnliche Matrizen die gleiche Determinante haben, folgt natürlich auch aus der Multiplikativität der Determinante. Die Spur ist zwar *nicht* multiplikativ, im Allgemeinen ist $\mathrm{tr}(AC) \neq \mathrm{tr}(A)\cdot\mathrm{tr}(C)$. Es gilt aber $\mathrm{tr}(AC) = \mathrm{tr}(CA)$, was ebenfalls direkt zeigt, dass ähnliche Matrizen die gleiche Spur haben. Das obige Lemma erlaubt die folgende basisunabhängige Formulierung:

Definition 2.10. Sei V ein endlich-dimensionaler K-Vektorraum. Für $f \in \mathrm{End}_K(V)$ setzen wir

$$\chi_f(t) \;:=\; \chi_A(t), \quad \det(f) \;:=\; \det(A) \quad \text{und} \quad \mathrm{tr}(f) \;:=\; \mathrm{tr}(A),$$

wobei $A = M_{\mathscr{B}}(f)$ die Abbildungsmatrix zu einer beliebigen Basis $\mathscr{B}$ sei. Dabei spielt die Wahl der Basis keine Rolle: Die Abbildungsmatrizen für je zwei Basen haben nach dem vorigen Lemma das gleiche charakteristische Polynom.

3 Nullstellen von Polynomen

Die Eigenwerte einer Matrix sind die Nullstellen ihres charakteristischen Polynoms, aber wie findet man diese? In günstigen Fällen kann man eine Nullstelle erraten und einen Linearfaktor ausklammern:

Lemma 3.1. *Sei $f(t) \in K[t]$ mit $\deg(f) > 0$, und sei $\lambda \in K$ mit $f(\lambda) = 0$. Dann gilt*

$$f(t) \;=\; (t-\lambda)\cdot g(t)$$

für ein eindeutiges $g(t) \in K[t]$ mit $\deg(g) = \deg(f) - 1$.

Beweis. Polynomdivision liefert $f(t) = (t-\lambda)g(t) + r(t)$ für Polynome $g, r \in K[t]$ mit $\deg(r) < \deg(t-\lambda) = 1$, d.h. mit $r \in K$. Einsetzen von $t = \lambda$ gibt $r = 0$. □

Induktiv erhalten wir das folgende Resultat, insbesondere kann jedes $f \in K[t]$ höchstens $\deg(f)$ verschiedene Nullstellen besitzen:

Satz 3.2. *Jedes vom Nullpolynom verschiedene $f(t) \in K[t]$ hat höchstens endlich viele Nullstellen. Seien die paarweise verschiedenen Nullstellen mit $\lambda_1, \ldots, \lambda_k \in K$ bezeichnet, dann gilt*

$$f(t) \;=\; (t-\lambda_1)^{e_1} \cdots (t-\lambda_k)^{e_k} \cdot g(t)$$

mit eindeutigen $e_i \in \mathbb{N}$ und einem eindeutigen $g(t) \in K[t]$ ohne Nullstellen $\lambda \in K$.

Beweis. Induktives Anwenden des Lemmas zeigt $f(t) = (t-\lambda_1)^{e_1} \cdots (t-\lambda_k)^{e_k} \cdot g(t)$ mit $\lambda_1, \ldots, \lambda_k \in K$ paarweise verschieden und $e_k \in \mathbb{N}$. Wählt man $\deg(g)$ minimal, so gilt $g(\lambda) \neq 0$ für alle $\lambda \in K$. Dann ist $f(\lambda) = 0$ genau für $\lambda \in \{\lambda_1, \ldots, \lambda_k\}$. Man überlegt sich leicht, dass

$$e_i \;=\; \max\{e \in \mathbb{N} \mid \exists h(t) \in K[t] : f(t) = (t-\lambda_i)^e \cdot h(t)\}$$

ist. Somit sind die e_i und damit auch das Polynom $g(t) \in K[t]$ eindeutig. □

Definition 3.3. In der obigen Situation schreiben wir auch $e_i = \operatorname{ord}_{t=\lambda_i}(f(t))$ und nennen diese Zahl die *Nullstellenordnung* von f im Punkt $t = \lambda_i$.

Beispiel 3.4. Für $f(t) = t^4 - 5t^3 + 9t^2 - 7t + 2 \in \mathbb{Q}[t]$ errät man $t = 1$ und $t = 2$ als Nullstellen und berechnet mit Polynomdivision

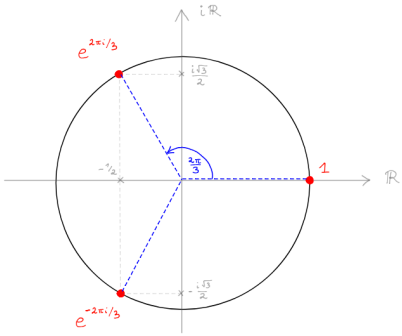
$$f(t) \;=\; (t-1)(t-2)(t^2-2t+1) \;=\; (t-1)^3(t-2).$$

Die Nullstellenordnungen sind also hier $\operatorname{ord}_{t=1}(f(t)) = 3$ und $\operatorname{ord}_{t=2}(f(t)) = 1$.

Beispiel 3.5. Für $f(t) = t^3 - 1$ ist $t = 1$ die einzige reelle Nullstelle. Es gibt jedoch zwei weitere komplexe Nullstellen, denn

$$f(t) \;=\; (t-1)(t^2+t+1) \;=\; (t-1)(t-\tfrac{1+i\sqrt{3}}{2})(t-\tfrac{1-i\sqrt{3}}{2}).$$

Man beachte, dass für jede komplexe Nullstelle λ eines reellen Polynoms auch $\bar{\lambda}$ eine Nullstelle sein muß. Im obigen Beispiel sind die komplexen Nullstellen genau die dritten Einheitswurzeln in der komplexen Ebene:



Definition 3.6. Ein Körper K heißt *algebraisch abgeschlossen*, falls jedes $f \in K[t]$ mit $\deg(f) > 0$ eine Nullstelle in K besitzt, d.h. ein $\lambda \in K$ mit $f(\lambda) = 0$.

Der Körper $K = \mathbb{R}$ ist nicht algebraisch abgeschlossen, denn $f(t) = t^2 + 1$ hat keine reelle Nullstelle. Wir hatten die komplexen Zahlen aus den reellen Zahlen konstruiert, indem wir formal eine Nullstelle dieses einen Polynoms dazugenommen haben. Der Fundamentalsatz der Algebra besagt, dass der so erhaltene Körper $\mathbb{C}$ sogar algebraisch abgeschlossen ist. Es gibt viele weitere interessante Beispiele:

Beispiel 3.7. Die Menge der *algebraischen Zahlen* ist

$$\overline{\mathbb{Q}} := \left\{\lambda \in \mathbb{C} \mid \exists f \in \mathbb{Q}[x] \setminus \{0\} : f(\lambda) = 0\right\} \subset \mathbb{C}.$$

In der Algebra zeigt man, dass die Menge $\overline{\mathbb{Q}}$ einen Körper bezüglich Addition und Multiplikation komplexer Zahlen bildet und dass dieser algebraisch abgeschlossen ist. Seine Elemente sind genau die Nullstellen rationaler Polynome, z.B. ist

$$\sqrt{2},\ \sqrt[3]{6},\ i,\ \exp\left(\tfrac{2\pi i}{7}\right),\ \tfrac{1+i\sqrt{3}}{2},\ \ldots \in \overline{\mathbb{Q}}$$

Da $\overline{\mathbb{Q}}$ abzählbar ist, gibt es sehr viele nicht-algebraische Zahlen, diese bezeichnet man als *transzendent*. Beispielsweise ist $\pi, e \notin \overline{\mathbb{Q}}$. Der Nachweis der Transzendenz ist meist schwierig. Z.B. ist bis heute unbekannt, ob $e + \pi$ transzendent ist!

Für algebraisch abgeschlossene Körper erhält unser Satz über die Abspaltung von Linearfaktoren eine besonders einfache Form:

Korollar 3.8. *Sei K algebraisch abgeschlossen. Dann hat jedes vom Nullpolynom verschiedene Polynom $f(t) \in K[t]$ eine bis auf Umordnung der Faktoren eindeutige Zerlegung als Produkt*

$$f(t) \;=\; c \cdot (t-\lambda_1)^{e_1}(t-\lambda_2)^{e_2}\cdots(t-\lambda_k)^{e_k}$$

mit paarweise verschiedenen $\lambda_i \in K$, einer Konstante $c \in K^\times$ und $e_i \in \mathbb{N}$.

Beweis. In Satz 3.2 ist $g(t) \in K[t]$ ein Polynom ohne Nullstellen in K. Wenn K algebraisch abgeschlossen ist, kann dies nur im Fall $\deg(g) = 0$ passieren, also ist hier $g = c \in K$ eine Konstante. □

Wir sagen in der Situation des obigen Korollars auch, das Polynom $f(t) \in K[t]$ zerfalle über dem Körper K vollständig in Linearfaktoren.

4 Diagonalisierbarkeit

Wir haben am Beispiel von Drehungen gesehen, dass nicht jede reelle Matrix einen reellen Eigenwert besitzt. Über den komplexen Zahlen hatten wir allerdings auch für Drehmatrizen Eigenwerte gefunden. Allgemein gilt:

Lemma 4.1. *Über algebraisch abgeschlossenen Körpen K hat jede quadratische Matrix $A \in \mathrm{Mat}(n \times n, K)$ mindestens einen Eigenvektor.*

Beweis. Es ist $\chi_A(t) \in K[t]$ ein Polynom vom Grad $n \geq 1$ und hat somit in dem algebraisch abgeschlossenen Körper eine Nullstelle $\lambda \in K$. □

Man beachte, dass wir damit erstmal nur *einen* Eigenvektor haben — auch über algebraisch abgeschlossenen Körpern ist nicht jede Matrix diagonalisierbar. Um dies zu verstehen, führen wir die folgenden Begriffe ein:

Definition 4.2. Sei V ein endlich-dimensionaler K-Vektorraum. Für $f \in \mathrm{End}_K(V)$ und $\lambda \in K$ nennen wir

- $d(f,\lambda) := \dim_K(E(f,\lambda))$ die *geometrische Vielfachheit* des Eigenwertes λ,
- $e(f,\lambda) := \mathrm{ord}_{t=\lambda}(\chi_f(t))$ die *algebraische Vielfachheit* des Eigenwertes λ.

Für Matrizen $A \in \mathrm{Mat}(n \times n, K)$ definieren wir die algebraische und geometrische Vielfachheit eines Eigenwertes als die des Endomorphismus $f\colon K^n \to K^n, v \mapsto A \cdot v$ und schreiben $d(A,\lambda) = d(f,\lambda)$ und $e(A,\lambda) = e(f,\lambda)$.

Lemma 4.3. *Für jeden Eigenwert λ von $f \in \mathrm{End}_K(V)$ gilt $1 \leq d(f,\lambda) \leq e(f,\lambda)$.*

Beweis. Wir wählen eine Basis $v_1, \ldots, v_d$ für $E(f,\lambda)$ und ergänzen sie zu einer Basis $\mathscr{B} = (v_1, \ldots, v_d, v_{d+1}, \ldots, v_n)$ des gesamten Vektorraumes. In dieser wird f dargestellt durch eine Blockmatrix

$$M_{\mathscr{B}}(f) \;=\; \left(\begin{array}{c|c} \lambda \cdot \mathbf{1} & C \\ \hline 0 & D \end{array}\right)$$

wobei der linke obere Block das Format $d \times d$ hat. Also hat $\chi_f(t) = (t-\lambda)^d \cdot \chi_D(t)$ im Punkt $t = \lambda$ eine Nullstelle der Ordnung $e(f,\lambda) \geq d = d(f,\lambda)$. □

Beispiel 4.4. Für $\lambda_1, \lambda_2 \in K$ und

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

ist $\chi_A(t) = (t-\lambda_1)(t-\lambda_2)$. Es gibt zwei Fälle:

- Für $\lambda_2 \neq \lambda_1$ ist $d(A,\lambda_i) = e(A,\lambda_i) = 1$.
- Für $\lambda_2 = \lambda_1$ ist $d(A,\lambda_1) = 1 < e(A,\lambda_1) = 2$.

Man prüft leicht nach, dass A im ersten Fall diagonalisierbar ist, im zweiten Fall nicht. Allgemein gilt:

Satz 4.5. *Sei V ein endlich-dimensionaler K-Vektorraum. Für $f \in \mathrm{End}_K(V)$ sind folgende Eigenschaften äquivalent:*

a) Der Endomorphismus f ist diagonalisierbar.

b) Das Polynom $\chi_f(t)$ zerfällt über K vollständig in Linearfaktoren und für alle λ ist

$$d(f,\lambda) \;=\; e(f,\lambda).$$

c) Für die Eigenräume zu den paarweise verschiedenen Eigenwerten $\lambda_1,\ldots,\lambda_k$ von f gilt

$$V \;=\; E(f,\lambda_1) \oplus \cdots \oplus E(f,\lambda_k).$$

Beweis. Zu $(a) \Longrightarrow (b)$: Ein diagonalisierbarer Endomorphismus $f \in \mathrm{End}_K(V)$ wird in einer geeigneten Basis $\mathscr{B}$ von V dargestellt durch eine Diagonalmatrix

$$M_{\mathscr{B}}(f) \;=\; \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} \;\in\; \mathrm{Mat}(n\times n, K).$$

Dann zerfällt $\chi_f(t) = (t-\mu_1)\cdots(t-\mu_n)$ in Linearfaktoren. Dabei müssen die μ_i nicht verschieden sein, aber für jeden Eigenwert $\lambda \in K$ ist

$$e(f,\lambda) \;=\; \#\{i \mid \mu_i = \lambda\} \;=\; d(f,\lambda).$$

Zu $(b) \Longrightarrow (c)$: Wenn das charakteristische Polynom in Linearfaktoren zerfällt, schreiben wir

$$\chi_f(t) \;=\; (t-\lambda_1)^{e_1} \cdots (t-\lambda_k)^{e_k}$$

mit paarweise verschiedenen $\lambda_i \in K$ und $e_i := e(f,\lambda_i)$. Jede Nullstelle $\lambda_i \in K$ ist ein Eigenwert von f. Da die Summe paarweise verschiedener Eigenräume direkt ist, folgt $V' := E(f,\lambda_1) \oplus \cdots \oplus E(f,\lambda_k) \subseteq V$. Wenn die geometrischen gleich den algebraischen Vielfachheiten sind, folgt aus Dimensionsgründen $V' = V$ wegen

$$\dim_K(V') \;=\; \sum_i d(f,\lambda_i) \quad \text{und} \quad \dim_K(V) \;=\; \sum_i e(f,\lambda_i).$$

Zu $(c) \Longrightarrow (a)$: Wenn $V = E(f,\lambda_1) \oplus \cdots \oplus E(f,\lambda_k)$ gilt, wählen wir in jedem der Eigenräume auf der rechten Seite eine Basis. Die Vereinigung dieser Basen ist dann eine Basis $\mathscr{B}$ von V mit $M_{\mathscr{B}}(f)$ in Diagonalform. □

Korollar 4.6. *Sei $f \in \mathrm{End}_K(V)$. Falls das charakteristische Polynom vollständig in Linearfaktoren zerfällt als*

$$\chi_f(t) \;=\; (t-\lambda_1)\cdots(t-\lambda_n)$$

mit paarweise verschiedenen Nullstellen $\lambda_1,\ldots,\lambda_n \in K$, so ist f diagonalisierbar.

Beweis. In diesem Fall ist $1 \le d(f,\lambda_i) \le e(f,\lambda_i) = 1$ für alle i, somit gilt in beiden Ungleichungen Gleichheit. □

Über algebraisch abgeschlossenen Körpern K ist jedes Polynom ein Produkt von Linearfaktoren. Dann setzt das Korollar nur voraus, dass die Nullstellen paarweise verschieden sind — eine zufällig gewählte Matrix über den komplexen Zahlen ist mit Sicherheit diagonalisierbar:

Beispiel 4.7. Für Matrizen

$$A \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\in\; \mathrm{Mat}(2\times 2,\mathbb{C})$$

hat $\chi_A(t) = t^2 - (a+d)t + (ad-bc)$ die beiden Nullstellen

$$\lambda \;=\; \tfrac{a+d}{2} \pm \sqrt{\tfrac{(a-d)^2}{4} + bc}$$

wobei die Wurzel komplex und nur bis auf einen Faktor ± 1 eindeutig ist. Sobald hier $(a-d)^2 \neq -4bc$ gilt, ist A also über $\mathbb{C}$ diagonalisierbar nach Korollar 4.6.

5 Anwendung: Lineare Rekursionen

In vielen Anwendungen spielen Potenzen von Matrizen eine Rolle. Wenn wir eine Matrix diagonalisieren können, lassen sich ihre Potenzen sehr einfach ablesen:

Beispiel 5.1. Sei

$$A \;=\; \begin{pmatrix} a & 1 \\ 1 & a \end{pmatrix} \;\in\; \mathrm{Mat}(2\times 2,\mathbb{R}) \quad \text{mit} \quad a \in \mathbb{R}.$$

Diese Matrix hat zwei verschiedene Eigenwerte $\lambda = a+1$ und $\mu = a-1$ zu den Eigenvektoren

$$v \;=\; \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad w \;=\; \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Also ist

$$S^{-1}AS = D = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \quad \text{für} \quad S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Die Potenzen

$$A^n = (SDS^{-1})^n = (SDS^{-1})(SDS^{-1})\cdots(SDS^{-1}) = S \cdot D^n \cdot S^{-1}$$

können wir nun sehr einfach berechnen. Mit

$$D^n = \begin{pmatrix} \lambda^n & 0 \\ 0 & \mu^n \end{pmatrix} \quad \text{und} \quad S^{-1} = \frac{1}{2} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

erhalten wir

$$\begin{aligned} A^n &= \frac{1}{2} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} \lambda^n & 0 \\ 0 & \mu^n \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{2} \cdot \begin{pmatrix} \lambda^n & \mu^n \\ \lambda^n & -\mu^n \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{2} \cdot \begin{pmatrix} c_n^+ & c_n^- \\ c_n^- & c_n^+ \end{pmatrix} \quad \text{mit} \quad c_n^\pm = \lambda^n \pm \mu^n. \end{aligned}$$

Man beachte, dass diese Formel auch für $n < 0$ gültig ist!

Matrixpotenzen spielen beispielsweise eine Rolle bei der Lösung von linearen Rekursionen. Wir fangen mit einem einfachen Beispiel an:

Beispiel 5.2. Sei $(F_n)_{n\in\mathbb{N}}$ die Folge der Fibonacci-Zahlen, die rekursiv definiert ist durch

$$F_0 := 0, \quad F_1 := 1 \quad \text{und} \quad F_{n+1} := F_n + F_{n-1} \quad \text{für} \quad n \in \mathbb{N}.$$

Wir schreiben diese Rekursion als Vektorgleichung

$$v_{n+1} = A \cdot v_n \quad \text{mit} \quad A := \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad v_n := \begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix}.$$

Dann gilt

$$v_{n+1} = A \cdot v_n = A^2 \cdot v_{n-1} = \cdots = A^n \cdot v_1 \quad \text{mit} \quad v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Die Potenzen A^n bestimmen wir, indem wir die Matrix diagonalisieren: Zunächst berechnen wir das charakteristische Polynom $\chi_A(t) = t^2 - t - 1$. Die Eigenwerte von A sind die Nullstellen des charakteristischen Polynoms, in unserem Fall sind dies

$$\lambda_\pm = \frac{1 \pm \sqrt{5}}{2}.$$

Man beachte $\lambda_+ + \lambda_- = 1$ und $\lambda_+ \cdot \lambda_- = -1$, was $\lambda_\pm - 1 = 1/\lambda_\pm$ zeigt. Für die Eigenräume folgt

$$E(A,\lambda_\pm) \;=\; \ker\begin{pmatrix}\lambda_\pm & -1\\ -1 & \lambda_\pm - 1\end{pmatrix} \;=\; \ker\begin{pmatrix}\lambda_\pm & -1\\ -1 & 1/\lambda_\pm\end{pmatrix} \;=\; \left\langle \begin{pmatrix}1\\ \lambda_\pm\end{pmatrix} \right\rangle_{\mathbb{R}}$$

und wir erhalten

$$S^{-1}AS \;=\; \begin{pmatrix}\lambda_+ & 0\\ 0 & \lambda_-\end{pmatrix} \quad \text{mit} \quad S \;=\; \begin{pmatrix}1 & 1\\ \lambda_+ & \lambda_-\end{pmatrix}.$$

Eine kurze Rechnung liefert

$$S^{-1} \;=\; \frac{1}{\sqrt{5}} \cdot \begin{pmatrix}-\lambda_- & 1\\ \lambda_+ & -1\end{pmatrix}.$$

Somit ist

$$A^n \;=\; S \cdot \begin{pmatrix}\lambda_+^n & 0\\ 0 & \lambda_-^n\end{pmatrix} \cdot S^{-1} \;=\; \frac{1}{\sqrt{5}} \cdot \begin{pmatrix}\lambda_+^n & \lambda_-^n\\ * & *\end{pmatrix} \cdot \begin{pmatrix}-\lambda_- & 1\\ \lambda_+ & -1\end{pmatrix} \;=\; \frac{1}{\sqrt{5}} \cdot \begin{pmatrix}* & \lambda_+^n - \lambda_-^n\\ * & *\end{pmatrix}$$

Die erste Komponente der Vektorgleichung $v_{n+1} = A^n \cdot v_1$ mit dem vorgegebenen Anfangswert $v_1 = e_2$ liefert nun die bekannte Formel

$$\boxed{F_n \;=\; \frac{1}{\sqrt{5}} \cdot \left(\lambda_+^n - \lambda_-^n\right).}$$

Man beachte, dass wir aus der berechneten Matrixpotenz A^n auch die Lösungen der Rekursion mit beliebigen anderen Anfangswerten erhalten! Ein solches Verfahren funktioniert allgemeiner auch für n-stufige lineare Rekursionen:

Satz 5.3. *Seien $c_0, c_1, \ldots, c_n \in K$, und sei V die Menge aller Folgen $x_0, x_1, \ldots$ in K mit*

$$x_{k+1} \;=\; c_n x_k + c_{n-1} x_{k-1} + \cdots + c_0 x_{k-n}, \quad \textit{für alle } k \geq n. \tag{$\dagger$}$$

Dann gilt:

a) Die Menge V bildet im Vektorraum aller Folgen einen Untervektorraum der Dimension

$$\dim_K V \;=\; n+1.$$

b) Für jede Nullstelle $\lambda \in K$ des Polynoms

$$p(t) \;:=\; t^{n+1} - c_n t^n - \cdots - c_0 \;\in\; K[t]$$

ist die Folge $x_0, x_1, \ldots$ der Potenzen $x_k := \lambda^k$ eine Lösung von ($\dagger$).

c) Die so gefundenen Lösungen sind als Vektoren von V linear unabhängig.

Beweis. Wir schreiben (†) als Vektorgleichung

$$v_{k+1} = A \cdot v_k \quad \text{mit} \quad A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 \\ c_0 & c_1 & c_2 & \cdots & c_n \end{pmatrix} \quad \text{und} \quad v_k = \begin{pmatrix} x_{k-n} \\ x_{k-n+1} \\ \vdots \\ x_{k-1} \\ x_k \end{pmatrix}$$

Dann folgt $v_{n+k} = A^k \cdot v_n$ mit beliebigem Anfangsvektor $v_n \in K^{n+1}$ und somit bilden die Lösungen der Rekursion (†) im Vektorraum aller Folgen einen Unterraum V mit $\dim_K(V) = n+1$ wie in (a) behauptet. Die Eigenwerte von A sind genau die Nullstellen von

$$\begin{aligned} \chi_A(t) &= \det \begin{pmatrix} t & -1 & & & \\ & t & -1 & & \\ & & \ddots & \ddots & \\ & & & t & -1 \\ -c_0 & -c_1 & \cdots & -c_{n-1} & -c_n \end{pmatrix} \\ &= t \cdot \det \begin{pmatrix} t & -1 & & & \\ & \ddots & \ddots & & \\ & & t & -1 & \\ -c_1 & \cdots & -c_{n-1} & -c_n & \end{pmatrix} + (-1)^n \cdot c_0 \cdot \det \begin{pmatrix} -1 & & & \\ t & -1 & & \\ & \ddots & \ddots & \\ & & t & -1 \end{pmatrix} \\ &= t \cdot (t^n - c_n t^{n-1} - \cdots - c_1) - c_0 \\ &= t^{n+1} - c_n t^n - \cdots - c_1 t - c_0, \end{aligned}$$

wobei wir im ersten Schritt nach der ersten Spalte entwickelt haben und im zweiten Schritt Induktion über n benutzt haben. Wenn wir als Anfangsvektor $v_n \in K^{n+1}$ einen Eigenvektor von A zum Eigenwert $\lambda \in K$ wählen, erhalten wir genau die Lösungen in (b). Die lineare Unabhängigkeit in (c) folgt, weil Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind. □

Korollar 5.4. *Falls das Polynom $p(t)$ im obigen Satz vollständig in Linearfaktoren zerfällt mit paarweise verschiedenen Nullstellen $\lambda_0, \ldots, \lambda_n \in K$, dann besitzt jede Lösung $x = (x_k)_{k \in \mathbb{N}_0}$ von (†) die Form*

$$x_k = \alpha_0 \lambda_0^k + \cdots + \alpha_n \lambda_n^k \quad \textit{mit eindeutigen} \quad \alpha_0, \ldots, \alpha_n \in K.$$

Beweis. Wenn $p(t)$ in paarweise verschiedene Linearfaktoren zerfällt, liefert (b) ein System von $n+1$ verschiedenen Lösungen der Rekursion. Nach (c) sind diese linear unabhängig, bilden also wegen $\dim_K(V) = n+1$ eine Basis des Lösungsraums. □

Wir haben hier die Matrix A aus dem obigen Beweis diagonalisiert: Man rechnet sofort $AS = SD$ nach für

$$S = \begin{pmatrix} 1 & \cdots & 1 \\ \lambda_0 & \cdots & \lambda_n \\ \vdots & & \vdots \\ \lambda_0^n & \cdots & \lambda_n^n \end{pmatrix} \quad \text{und} \quad D = \begin{pmatrix} \lambda_0 & & & \\ & \lambda_1 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

und für paarweise verschiedene $\lambda_0, \ldots, \lambda_n$ ist S invertierbar (Vandermonde).

6 Anwendung: Systeme linearer DGL

Eine weitere Anwendung, in der Potenzen von Matrizen eine Rolle spielen, sind sog. Systeme von linearen Differentialgleichungen erster Ordnung mit konstanten Koeffizienten. Hierunter versteht man Gleichungssysteme der Form

$$\begin{aligned} y_1'(t) &= a_{11}y_1(t) + a_{12}y_2(t) + \cdots + a_{1n}y_n(t) \\ y_2'(t) &= a_{21}y_1(t) + a_{22}y_2(t) + \cdots + a_{2n}y_n(t) \\ &\vdots \\ y_n'(t) &= a_{n1}y_1(t) + a_{n2}y_2(t) + \cdots + a_{nn}y_n(t) \end{aligned}$$

Die $a_{ij} \in \mathbb{R}$ sind dabei gegeben. Gesucht sind differenzierbare Funktionen $y_i(t)$ auf einem reellen Intervall, deren Ableitungen $y_i'(t)$ das obige System von Gleichungen erfüllen. Für $n = 1$ reduziert sich dies auf eine Gleichung $y'(t) = ay(t)$ mit $a \in \mathbb{R}$, deren allgemeine Lösung

$$y(t) = e^{at} \cdot y(0)$$

ist. Für $n \geq 1$ beliebig schreiben wir das obige DGL-System kurz als $Y'(t) = A \cdot Y(t)$ mit

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad Y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} \quad \text{und} \quad Y'(t) = \begin{pmatrix} y_1'(t) \\ \vdots \\ y_n'(t) \end{pmatrix}.$$

Dabei ist die gesuchte Lösung Y ein Element des Vektorraumes

$$\mathscr{D}_n(\mathbb{R}) := \{\text{differenzierbare Funktionen } f\colon \mathbb{R} \to \mathbb{R}^n\}.$$

mit der punktweisen Vektorraumstruktur. Die Lösungsmenge des DGL-Systems ist ein Untervektorraum

$$\mathscr{L}(A) = \{Y \in \mathscr{D}_n(\mathbb{R}) \mid \forall t \in \mathbb{R}\colon Y'(t) = A \cdot Y(t)\} \subset \mathscr{D}_n(\mathbb{R}),$$

denn die Nullfunktion $Y \equiv 0$ ist eine Lösung, und für je zwei Lösungen $Y_i \in \mathscr{L}(A)$ und $\alpha_i \in \mathbb{R}$ ist auch $\alpha_1 Y_1 + \alpha_2 Y_2 \in \mathscr{L}(A)$ wegen

$$(\alpha_1 Y_1 + \alpha_2 Y_2)' \;=\; \alpha_1 Y_1' + \alpha_2 Y_2' \;=\; \alpha_1 A Y_1 + \alpha_2 A Y_2 \;=\; A \cdot (\alpha_1 Y_1 + \alpha_2 Y_2).$$

Eine Basis von $\mathscr{L}(A)$ nennt man ein *Fundamentalsystem* der gegebenen DGL. Für Diagonalmatrizen lässt sich ein solches leicht angeben:

Beispiel 6.1. Für Diagonalmatrizen

$$A \;=\; \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

mit $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ lautet das zugehörige DGL-System:

$$\begin{aligned} y_1'(t) \;&=\; \lambda_1 y_1(t) \\ &\vdots \\ y_n'(t) \;&=\; \lambda_n y_n(t). \end{aligned}$$

Die Gleichungen sind entkoppelt, wir erhalten somit sämtliche Lösungen in der Form

$$y_i(t) \;=\; e^{\lambda_i t} \cdot v_i \quad \text{für beliebige Anfangswerte } v_i \in \mathbb{R},\ i = 1, \ldots, n.$$

Insbesondere ist hier $\dim_{\mathbb{R}} \mathscr{L}(A) = n$. Dasselbe geht für diagonalisierbare Matrizen:

Satz 6.2. *Es gilt:*

a) Ist $v \in \mathbb{R}^n$ ein Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{R}$, so ist

$$Y(t) := e^{\lambda t} \cdot v \;\in\; \mathscr{L}(A).$$

b) Ist A diagonalisierbar und wählen wir eine Basis aus Eigenvektoren $v_i \in \mathbb{R}^n$ zu Eigenwerten $\lambda_i \in \mathbb{R}$, dann bilden die Funktionen

$$Y_i(t) \;:=\; e^{\lambda_i t} \cdot v_i \quad \textit{für } i = 1, \ldots, n$$

ein Fundamentalsystem für die Differentialgleichung $Y'(t) = AY(t)$.

Beweis. *a)* Für $A \cdot v = \lambda \cdot v$ mit $\lambda \in \mathbb{R}$ liegt die Funktion $Y(t) := e^{\lambda t} \cdot v$ in $\mathscr{L}(A)$ wegen

$$Y'(t) \;=\; e^{\lambda t} \cdot \lambda v \;=\; e^{\lambda t} \cdot Av \;=\; A \cdot e^{\lambda t} v \;=\; A \cdot Y(t).$$

b) Sei nun $v_1,\dots,v_n$ eine Basis aus Eigenvektoren zu Eigenwerten $\lambda_1,\dots,\lambda_n$, also

$$D := S^{-1}AS = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad \text{für} \quad S = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} \in \mathrm{GL}_n(\mathbb{R}).$$

Für $Z(t) := S^{-1} \cdot Y(t)$ gilt dann:

$$\begin{aligned} Y'(t) &= S \cdot Z'(t) \cdot S^{-1} \\ A \cdot Y(t) &= S \cdot D \cdot Z(t) \cdot S^{-1}. \end{aligned}$$

Also ist $Y'(t) = AY(t)$ genau für $Z'(t) = DZ(t)$. Wir erhalten einen Isomorphismus von Vektorräumen

$$\mathscr{L}(D) \xrightarrow{\sim} \mathscr{L}(A), \quad Z(t) \mapsto Y(t) = S \cdot Z(t),$$

und die Behauptung folgt aus dem vorigen Beispiel. □

Beispiel 6.3. Gesucht seien alle Lösungen des DGL-Systems

$$\begin{aligned} y_1' &= ay_1 + y_2 \\ y_2' &= y_2 + ay_1. \end{aligned}$$

In Matrixform lautet dieses

$$Y'(t) = A \cdot Y(t) \quad \text{mit} \quad A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Wir haben im vorigen Abschnitt gesehen, dass die Matrix A diagonalisierbar ist. Genauer gilt

$$S^{-1}AS = \begin{pmatrix} a+1 & 0 \\ 0 & a-1 \end{pmatrix} \quad \text{mit} \quad S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Die allgemeine Lösung des DGL-Systems hat somit die Form

$$Y(t) = \alpha \cdot e^{a+1} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \beta \cdot e^{a-1} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

mit beliebigen Konstanten $\alpha, \beta \in \mathbb{R}$. Für die Komponenten bedeutet dies

$$\begin{aligned} y_1(t) &= \alpha \cdot e^{a+1} + \beta \cdot e^{a-1} \\ y_2(t) &= \alpha \cdot e^{a+1} - \beta \cdot e^{a-1} \end{aligned}$$

Wer's nicht glaubt, sollte an dieser Stelle nachrechnen, dass diese beiden Funktionen tatsächlich eine Lösung des gegebenen DGL-Systems bilden!

Der obige Satz bleibt sinngemäß auch im Fall komplexer Eigenwerte gültig, wenn man mit komplexwertigen Funktionen $y\colon \mathbb{R} \to \mathbb{C}$ arbeitet. Die Ableitung ist erklärt durch

$$y'(t) \;:=\; \frac{d}{dt}\mathrm{Re}(y)(t) + i\frac{d}{dt}\mathrm{Im}(y)(t),$$

also durch die separate Ableitung des Real- und Imaginärteils der Funktion. Damit gilt die Identität

$$\tfrac{d}{dt}e^{\lambda t} \;=\; \lambda e^{\lambda t}$$

auch für $\lambda \in \mathbb{C}$, sodass der obige Satz auch für komplexwertige Lösungen linearer DGL-Systeme angewendet werden kann. Das ist nützlich, auch wenn man sich nur für reelle Lösungen interessiert:

Beispiel 6.4. Gesucht seien alle Lösungen des DGL-Systems

$$\begin{aligned} y_1' &= y_1 - y_3, \\ y_2' &= y_2, \\ y_3' &= y_1 + y_3 \end{aligned}$$

Wir schreiben das DGL-System zunächst als

$$Y'(t) \;=\; A \cdot Y(t) \quad \text{mit} \quad A \;=\; \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Die Eigenwerte von A sind 1 und $1 \pm i$, tatsächlich gilt

$$S^{-1}AS \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1+i & 0 \\ 0 & 0 & 1-i \end{pmatrix} \quad \text{für} \quad S \;=\; \begin{pmatrix} 0 & i & -i \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Somit hat jede komplexwertige Lösung des DGL-Systems die Form

$$Y(t) \;=\; \alpha \cdot e^t \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \beta \cdot e^{(1+i)t} \cdot \begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix} + \gamma \cdot e^{(1-i)t} \cdot \begin{pmatrix} -i \\ 0 \\ 1 \end{pmatrix}$$

mit $\alpha, \beta, \gamma \in \mathbb{C}$. Um reelle Lösungen zu erhalten, müssen wir $\alpha \in \mathbb{R}$ und $\gamma = \overline{\beta}$ wählen. Wenn wir speziell $(\alpha, \beta, \gamma) = (1,0,0), \frac{1}{2}(0,1,1), \frac{1}{2}(0,i,-i)$ einsetzen, so erhalten wir für den $\mathbb{R}$-Vektorraum der reellwertigen Lösungen eine Basis bestehend aus den drei Vektoren

$$e^t \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e^t \cdot \begin{pmatrix} -\sin t \\ 0 \\ \cos t \end{pmatrix}, \quad e^t \cdot \begin{pmatrix} \cos t \\ 0 \\ \sin t \end{pmatrix}.$$

Die allgemeine reellwertige Lösung des DGL-Systems erhalten wir als $\mathbb{R}$-Linearkombination dieser drei Vektoren. Indem wir wieder zu den Komponenten übergehen, erhalten

wir die reellen Lösungen

$$\begin{aligned} y_1(t) &= (c \cdot \cos t - b \cdot \sin t) \cdot e^t \\ y_2(t) &= a \cdot e^t \\ y_3(t) &= (b \cdot \cos t + c \cdot \sin t) \cdot e^t \end{aligned}$$

mit $a, b, c \in \mathbb{R}$, die durch die Anfangswerte $(y_1(0), y_2(0), y_3(0))$ festgelegt sind.