5. Vektorräume

Die lineare Algebra ist die Theorie der Vektorräume. Vektoren und Vektorräume treten in vielfacher Gestalt und in vielen Anwendungen auf. Sowohl die Technik, als auch die theoretische Physik kommen ohne den Begriff des Vektorraumes nicht aus. In Anwendungen werden häufig Vektoren verwendet, um die Lage von Punkten in der Ebene, im Raum oder sogar in höher dimensionalen Räumen durch ihre Koordinaten zu beschreiben. Es gibt jedoch auch viele andere Beispiele für Vektorräume, die in den praktischen Anwendungen sehr wichtig sind. Wir nennen hier nur einige Gebiete, in denen die Vektorräume als zentrales Hilfsmittel verwendet werden: lineare Gleichungssysteme, Approximationstheorie, Kryptographie, Stochastik, Ökonomie, Spieltheorie, Computer Graphik, Statik, Genetik, Computer Tomographie, elektrische Netzwerke. Der Begriff des Vektorraumes greift also weit über die Mathematik hinaus, ist aber auch grundlegend für den Aufbau der gesamten weiteren Mathematik.

5.1 Grundbegriffe, Untervektorräume

Wir definieren in diesem Abschnitt den für uns zentralen Begriff eines Vektorraumes. Wir werden eine sehr allgemeine Definition geben. Die Definition bezieht sich dabei auf einen als fest vorausgesetzten Körper K.

Definition 5.1.1 Sei K im folgenden ein fest gewählter Körper. Seine Elemente werden auch *Skalare* genannt. Eine Menge V zusammen mit einer Addition

$$V \times V \ni (v, w) \mapsto v + w \in V$$

und einer Multiplikation mit Skalaren aus K

$$K \times V \ni (\alpha, v) \mapsto \alpha \cdot v \in V$$

heißt ein *Vektorraum* (über dem Körper K, auch K-Vektorraum), wenn die folgenden Gesetze erfüllt sind:

1. $(V, +)$ ist eine abelsche Gruppe,
2. $\forall \alpha, \beta \in K, v \in V[\alpha \cdot (\beta \cdot v) = (\alpha \cdot \beta) \cdot v]$ (Assoziativgesetz der Multiplikation),

3. $\forall \alpha \in K, v, w \in V[\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w]$ (1. Distributivgesetz),
4. $\forall \alpha, \beta \in K, v \in V[(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$ (2. Distributivgesetz),
5. für 1 in K gilt $\forall v \in V[1 \cdot v = v]$ (Gesetz von der Eins).

Die Elemente eines Vektorraumes V werden *Vektoren* genannt. Das neutrale Element der Gruppe $(V, +)$ wird mit $0 \in V$ bezeichnet und heißt *Nullvektor*. Wir werden in Zukunft grundsätzlich den Multiplikationspunkt fortlassen und schreiben $\alpha v := \alpha \cdot v$.

Beispiele 5.1.2 1. Offenbar ist der Körper K selbst mit seiner Addition und Multiplikation ein Vektorraum.
2. Der Vektorraum 0, genannt der *Nullvektorraum*, der nur aus einem Element besteht, ist ein Vektorraum, denn die Addition und Multiplikation sind vollständig festgelegt und erfüllen trivialerweise die Vektorraumgesetze.
3. Als nächstes Beispiel betrachten wir die „reelle Ebene". Die Menge V sei dabei die Menge der Paare (x, y) von reellen Zahlen, also $\mathbb{R} \times \mathbb{R}$ oder $\mathbb{R}^2$. Jeden Punkt der Ebene kann man eindeutig festlegen durch die Angabe seiner x-Koordinate und seiner y-Koordinate. Die Addition in $V = \mathbb{R}^2$ sei definiert durch die komponentenweise Addition der Paare, also durch
$$(\alpha, \beta) + (\gamma, \delta) := (\alpha + \gamma, \beta + \delta).$$
Die Multiplikation mit Skalaren aus dem Körper $\mathbb{R}$ der reellen Zahlen sei die komponentenweise Multiplikation eines Paares mit einer reellen Zahl, also
$$\alpha(\beta, \gamma) := (\alpha\beta, \alpha\gamma).$$
Weil wir dann in jeder Komponente einzeln rechnen, ergeben sich die Vektorraumgesetze unmittelbar aus den Körpergesetzen von $\mathbb{R}$.
4. Ein weiteres Beispiel ist die Menge $V := \mathbb{R}^n$ der n-Tupel reeller Zahlen. Die Addition auf V definieren wir wie bei den Paaren komponentenweise
$$(\xi_1, \ldots, \xi_n) + (\eta_1, \ldots, \eta_n) := (\xi_1 + \eta_1, \ldots, \xi_n + \eta_n).$$
Ebenso definieren wir die Multiplikation mit Skalaren komponentenweise
$$\alpha(\xi_1, \ldots, \xi_n) := (\alpha\xi_1, \ldots, \alpha\xi_n).$$
Damit lassen sich die Axiome für einen Vektorraum leicht nachrechnen. Wir wollen hier auf die Rechnung verzichten.
5. Ein fünftes Beispiel für einen Vektorraum ist die Menge der Quadrupel $(\xi_1, \ldots, \xi_4)$ mit $\xi_1 + \xi_2 + \xi_3 + \xi_4 = 0$, also
$$V := \{(\xi_1, \xi_2, \xi_3, \xi_4) | \xi_1, \xi_2, \xi_3, \xi_4 \in \mathbb{R}, \xi_1 + \xi_2 + \xi_3 + \xi_4 = 0\}.$$
Auch diese Menge bildet einen Vektorraum mit komponentenweiser Addition und Multiplikation. Diesen Vektorraum, obwohl er „dreidimensional" ist, können wir uns schlecht vorstellen. Er ist jedoch typisch für viele auf diese Weise konstruierte Vektorräume.

6. Wenn L ein Körper ist und den Körper $K \subset L$ als Unterkörper enthält, dann ist L mit der Addition eine abelsche Gruppe. Die Multiplikation von Elementen aus L mit Elementen aus K erfüllt offenbar alle Vektorraumgesetze. Also ist L ein Vektorraum über dem Körper K. Insbesondere ist $\mathbb{R}$ ein Vektorraum über $\mathbb{Q}$, weiter ist $\mathbb{C}$ ein Vektorraum über $\mathbb{R}$ und schließlich ist auch $\mathbb{C}$ ein Vektorraum über $\mathbb{Q}$.

Wir kommen zu einer allgemeinen Methode, neue Vektorräume aus schon vorhandenen zu konstruieren.

Lemma 5.1.3 *Sei V ein Vektorraum und I eine Menge. Dann ist die Menge der Abbildungen* $\mathrm{Abb}(I, V) = \prod_I V = V^I$ *mit komponentenweisen Operationen ein Vektorraum.*

Beweis. Wie bei den Gruppen in Kapitel 3 definieren wir die Addition von Familien durch $(v_i) + (w_i) := (v_i + w_i)$. Die Multiplikation mit Skalaren wird durch $\alpha(v_i) := (\alpha v_i)$ definiert. Da die Vektorraumgesetze in jeder Komponente einzeln erfüllt sind, sind sie auch für die Familien erfüllt.

Beispiele 5.1.4 1. (Hauptbeispiel für Vektorräume): Da K ein Vektorraum ist, ist auch K^I nach Lemma 5.1.3 ein Vektorraum für jede Menge I.

2. Wir definieren den Vektorraum K^n als K^I mit $I = \{1, \ldots, n\}$. Die Elemente sind n-Tupel $(\alpha_1, \ldots, \alpha_n)$ mit Koeffizienten aus K. Sie werden auch *Zeilenvektoren* der Länge n genannt. Die zuvor angegebenen Beispiele 2. und 3. fallen unter diese Konstruktion.
3. Der Vektorraum K_n soll ebenfalls als K^I definiert sein mit $I = \{1, \ldots, n\}$. Seine Elemente sollen jedoch als *Spaltenvektoren*
$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$
geschrieben werden.
4. Der Vektorraum $K_m^n = K^{m \times n}$ wird als K^I mit $I = \{1, \ldots, m\} \times \{1, \ldots, n\}$ definiert. Wir schreiben jeden Vektor als rechteckiges Schema
$$\begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \ldots & \alpha_{mn} \end{pmatrix}$$
und nennen einen solchen Vektor eine $m \times n$*-Matrix* (Plural: Matrizen). Die Zeilenvektoren der Form $(\alpha_{i1}, ..., \alpha_{in})$ heißen dabei die i-ten *Zeilen* der Matrix, die Spaltenvektoren der Form
$$\begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

heißen die j-ten *Spalten* der Matrix. m heißt die *Zeilenzahl*, n die *Spaltenzahl*. Insbesondere können Matrizen gleicher Zeilen- und Spaltenzahl addiert werden und mit Skalaren multipliziert werden.

5. Die Folgen reeller Zahlen $\mathbb{R}^{\mathbb{N}}$ bilden bei komponentenweiser Operation einen Vektorraum über den reellen Zahlen (kurz einen reellen Vektorraum), den *Folgenraum*.
6. Die reellen Funktionen $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ bilden einen reellen Vektorraum, den *Funktionenraum*, wenn man Addition und Multiplikation mit Skalaren auf den Werten der Funktionen definiert: $(f+g)(\alpha) = f(\alpha) + g(\alpha)$ und $(\alpha f)(\beta) = \alpha f(\beta)$.

Lemma 5.1.5 *(Rechengesetze in Vektorräumen) Sei V ein K-Vektorraum. Dann gelten für Elemente $\alpha, \beta \in K, v, w, \in V$ und für die Elemente $0_K, 1_K \in K$ und $0_V \in V$:*

1. $0_K v = 0_V = \alpha 0_V$,
2. $(-\alpha)v = -(\alpha v) = \alpha(-v)$,
3. $\alpha v = 0_V \Longrightarrow \alpha = 0_K \vee v = 0_V$.

Beweis. 1. Wir bezeichnen die Null in K und in V mit demselben Symbol 0. Aus $0v = (0+0)v = 0v + 0v$ erhalten wir durch Kürzen $0 = 0v$. Analog gilt $\alpha 0 = \alpha(0+0) = \alpha 0 + \alpha 0$.

2. Wir weisen nach, daß $(-\alpha)v$ das inverse Element bei der Addition zu αv ist. Es ist nämlich $\alpha v + (-\alpha)v = (\alpha + (-\alpha))v = 0v = 0$. Ebenso ist $\alpha v + \alpha(-v) = \alpha(v + (-v)) = \alpha 0 = 0$.

3. Wenn $\alpha = 0$ gilt, brauchen wir nichts zu zeigen. Wenn $\alpha v = 0$ und $\alpha \neq 0$ ist, dann gibt es ein Inverses α^{-1}, und es ist $v = 1v = (\alpha^{-1}\alpha)v = \alpha^{-1}(\alpha v) = \alpha^{-1} 0 = 0$, was zu zeigen war.

In jedem Vektorraum ($\neq 0$) wird es Teilmengen geben, die selbst wieder einen Vektorraum bilden. Ein Beispiel dafür kennen wir schon als Beispiel 5.1.2 5. Der dort definierte Vektorraum V ist Teilmenge von $\mathbb{R}^4$. Das führt uns zu dem Begriff des Untervektorraumes.

Definition 5.1.6 Sei V ein Vektorraum über dem Körper K. Eine nichtleere Teilmenge $U \subseteq V$ von V heißt *Untervektorraum*, wenn gelten

1. für alle $u, u' \in U$ ist $u + u' \in U$, d.h. U ist bezüglich der Addition von V abgeschlossen,

2. für alle $u \in U$ und alle $\alpha \in K$ ist $\alpha u \in U$, d.h. U ist unter der Multiplikation mit Skalaren abgeschlossen.

Man sieht leicht, daß jeder Untervektorraum selbst wieder ein Vektorraum wird, mit der Addition und der Multiplikation mit Skalaren wie in V definiert. Wir prüfen das hier nicht nach, bemerken aber, daß wir so eine Fülle von Beispielen für Vektorräume erhalten, ohne daß wir alle Axiome für einen Vektorraum in jedem Einzelfall nachrechnen müssen. Zur Definition von Untervektorräumen werden wir immer eine Teilmenge von Elementen mit einer gemeinsamen Eigenschaft bilden und untersuchen, ob diese Eigenschaft beim Bilden von Summen und Produkten mit Skalaren erhalten bleibt. Beispiel 5.1.2 5. ist so gewonnen worden.

Beispiel 5.1.7 Sei I eine Menge. Dann ist

$$K^{(I)} := \{(v_i) \in K^I | \text{für nur endlich viele } i \in I \text{ gilt } v_i \neq 0\}$$

ein Untervektorraum von K^I. Wenn man nämlich zwei Familien (v_i) und (w_i) in $K^{(I)}$ hat, so gilt auch für die Summe $(v_i) + (w_i) = (v_i + w_i)$, daß nur endlich viele Terme $v_i + w_i$ von Null verschieden sind. Ebenso sind auch in der Familie $\alpha(v_i) = (\alpha v_i)$ nur endlich viele Terme von Null verschieden. Dieser Vektorraum ist natürlich nur dann verschieden von K^I, wenn I eine unendliche Menge ist.

Ein weiteres Beispiel ist der Durchschnitt von mehreren Untervektorräumen.

Lemma 5.1.8 *Seien U_i mit $i \in I \neq \emptyset$ Untervektorräume des Vektorraumes V. Dann ist auch $\bigcap_{i\in I} U_i$ ein Untervektorraum von V.*

Beweis. Seien $u, u' \in \bigcap_{i\in I} U_i$ und sei $\alpha \in K$. Dann gilt $u, u' \in U_i$ für alle $i \in I$. Da die U_i Untervektorräume sind, gilt dann $u + u', \alpha u \in U_i$ für alle i, also auch $u + u', \alpha u \in \bigcap_{i\in I} U_i$. Damit sind die Bedingungen a) und b) für einen Untervektorraum erfüllt. Sicher ist auch $\bigcap_{i\in I} U_i$ nicht leer, weil der Vektor 0 in allen Untervektorräumen U_i und damit im Durchschnitt enthalten ist.

Um auch den trivialen Durchschnitt mit $I = \emptyset$ zu erfassen, definieren wir $\bigcap_{i\in\emptyset} U_i := V$.

Lemma 5.1.9 *Seien $U_1, ..., U_n$ Untervektorräume des Vektorraumes V. Dann ist auch*

$$U_1 + \ldots + U_n = \sum_{i=1}^{n} U_i$$

$:= \{u \in V|$ *es gibt Vektoren $u_i \in U_i$ für $i = 1, \ldots, n$ mit $u = u_1 + \ldots + u_n\}$ ein Untervektorraum von V.*

Beweis. Seien $u, u' \in \sum_{i=1}^n U_i$ und sei $\alpha \in K$. Dann gibt es Vektoren $u_i, u'_i \in U_i$ $(i = 1, \ldots, n)$ mit $u = u_1 + \ldots + u_n$ und $u' = u'_1 + \ldots u'_n$. Für $u + u'$ und αu erhält man dann $u + u' = (u_1 + u'_1) + \ldots + (u_n + u'_n)$ und $\alpha u = \alpha u_1 + \ldots \alpha u_n$. Damit haben wir wegen $u + u', \alpha u \in \sum_{i=1}^n U_i$ die Bedingungen für einen Untervektorraum erfüllt. Schließlich ist $\sum_{i=1}^n U_i$ trivialerweise nicht leer, weil $0 = 0 + \ldots + 0 \in \sum_{i=1}^n U_i$ gilt.

Wenn man eine unendliche Familie von Vektoren $(v_i|i \in I)$ mit einer beliebigen Indexmenge I hat und nur endlich viele Vektoren v_i von Null verschieden sind, dann möchte man genau diese Vektoren addieren. Wir bemerken zunächst, daß man zu dieser endlichen Summe $v_{i_1} + \ldots + v_{i_n}$, wobei die $v_{i_1}, \ldots, v_{i_n}$ genau diejenigen Vektoren sind, die von Null verschieden sind, auch noch beliebig endlich viele Exemplare des Nullvektors addieren kann, ohne diese Summe zu ändern. Es ist aber recht unbequem, immer darauf hinweisen zu müssen, daß man nun genau oder mindestens über alle diejenigen Vektoren v_i einer Familie $(v_i|i \in I)$ summiert, die von Null verschieden sind. Man schreibt daher $\sum_{i\in I} v_i$ für diese Summe der endlich vielen von Null verschiedenen Vektoren. Es handelt sich also genau genommen nicht um eine unendliche Summe, sondern lediglich um die Summe von endlich vielen Vektoren. Man kann zwei solche Summen addieren: $\sum_{i\in I} u_i + \sum_{i\in I} v_i = \sum_{i\in I}(u_i + v_i)$ und mit einem Skalar multiplizieren: $\alpha \sum_{i\in I} u_i = \sum_{i\in I}(\alpha u_i)$, da jeweils insgesamt nur endlich viele Terme von Null verschieden sind. Diesen Begriff wollen wir jetzt zur Definition der unendlichen Summe von Untervektorräumen verwenden.

Definition 5.1.10 Sei $I \neq \emptyset$ eine Menge und $(U_i|i \in I)$ eine unendliche Familie von Untervektorräumen des Vektorraumes V. Dann definieren wir

$$\sum_{i\in I} U_i := \{\sum_{i\in I} u_i|\forall i \in I[u_i \in U_i] \text{ und nur endliche viele } u_i \neq 0\}.$$

Man beweist wie zuvor

Lemma 5.1.11 *Sei $(U_i|i \in I)$ eine Familie von Untervektorräumen des Vektorraumes V. Dann ist auch $\sum_{i\in I} U_i$ ein Untervektorraum von V.*

Es handelt sich hier tatsächlich um eine echte unendliche Summe von Untervektorräumen, d.h. alle Untervektorräume der Summe können vom Nullvektorraum verschieden sein. Es werden aber trotzdem nur endliche Summen von Vektoren verwendet.

Man kann leicht wieder Rechenregeln für das Rechnen mit Untervektorräumen U_i von V beweisen. Wir führen hier nur einige relevante Regeln ohne Beweis an:

$$\begin{array}{l}
U_1 + U_2 = U_2 + U_1,\\
(U_1 + U_2) + U_3 = U_1 + (U_2 + U_3),\\
(U_1 \cap U_2) + (U_1 \cap U_3) \subset U_1 \cap (U_2 + U_3),\\
U_1 + (U_2 \cap U_3) \subset (U_1 + U_2) \cap (U_1 + U_3),\\
0 + U = U,\\
0 \cap U = 0,\\
V + U = V,\\
V \cap U = U,\\
\text{wenn } U_1 \subset U_3 \text{ gilt, dann ist}\\
\quad (U_1 + U_2) \cap U_3 = U_1 + (U_2 \cap U_3), \text{(modulares Gesetz).}
\end{array}$$

Übungen 5.1.12 1. Welche der folgenden Mengen sind Untervektorräume der angegebenen reellen Vektorräume?

a) $\{(\lambda, \quad 2\lambda, \quad 3\lambda)\,|\lambda \in \mathbb{R}\} \subset \mathbb{R}^3$,
b) $\{(\lambda, \quad 2\lambda, \quad 3\lambda)\,|\lambda \in \mathbb{Q}\} \subset \mathbb{R}^3$,
c) $\{(1+\lambda, \quad 1-\lambda)\,|\lambda \in \mathbb{R}\} \subset \mathbb{R}^2$,
d) $\{(\xi, \quad \eta, \quad \zeta) \in \mathbb{R}^3 | \xi^2 + \eta^2 + \zeta^2 = 1\} \subset \mathbb{R}^3$,
e) $\{(\xi, \quad \eta, \quad \zeta) \in \mathbb{R}^3 | 2\xi + 3\eta = \zeta\} \subset \mathbb{R}^3$,
f) $\{(\xi, \quad \eta, \quad \zeta) \in \mathbb{R}^3 | \xi\eta = \zeta\} \subset \mathbb{R}^3$,
g) $\{(x, \quad y, \quad z)\,| x^2 + 3xy + y^2 = 0\} \subset \mathbb{R}^3$,
h) $\{(x, \quad y, \quad z)\,| x^2 - 2xy + y^2 = 0\} \subset \mathbb{R}^3$,
i) $\{(x, \quad y, \quad z)\,| x^2 + 3xy + y^2 = 0\} \subset \mathbb{C}^3$ (reeller oder komplexer Untervektorraum).

2. Stellen Sie fest, ob die folgenden Summen definiert sind, und bestimmen Sie ihre Werte, soweit möglich:
a) $2 \cdot (0, \quad 1) + 5 \cdot (1, \quad 1) - 4 \cdot (2, \quad 1)$,
b) $2 \cdot \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + 5 \cdot (1, \quad 0, \quad 1)$,
c) $(1, \quad 0, \quad 1) + (0, \quad 1, \quad 0, \quad 1)$.

3. Entscheiden Sie, welche der folgenden Aussagen richtig sind (ja/nein).
a) Sei V ein Vektorraum über K. Dann gilt für alle $v \in V$: $0v = 0$.
b) Sei V ein Vektorraum über K und v ein Vektor aus V, der nicht der Nullvektor ist. Dann folgt aus $\alpha v = 0$, daß $\alpha = 0$ ist.

4. Sei V ein Vektorraum über dem Körper K, und sei X eine Menge. Zeigen Sie, daß die Menge aller Abbildungen $\mathrm{Abb}(X, V)$ von X nach V in Bezug auf die Verknüpfungen

$$\mathrm{Abb}(X, V) \times \mathrm{Abb}(X, V) \longrightarrow \mathrm{Abb}(X, V), (f, g) \mapsto f + g$$

mit $(f+g)(x) := f(x) + g(x)$ und

$$K \times \mathrm{Abb}(X, V) \longrightarrow \mathrm{Abb}(X, V), (\alpha, f) \mapsto \alpha f$$

mit $(\alpha f)(x) := \alpha(f(x))$ ein Vektorraum ist. (Sie sollen also die Vektorraumaxiome detailliert überprüfen.)

5. Sei

$$G := \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid \forall x \in \mathbb{R}\,[f(-x) = f(x)]\}$$

der Raum der geraden Funktionen und

$$U := \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid \forall x \in \mathbb{R}\,[f(-x) = -f(x)]\}$$

der Raum der ungeraden Funktionen.
a) Zeigen Sie, daß G und U Untervektorräume des Vektorraumes $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ sind.
b) Zeigen Sie: $\mathrm{Abb}(\mathbb{R}, \mathbb{R}) = G + U$.

6. Zeigen Sie: Ein Vektorraum kann nicht Vereinigung von zwei echten Untervektorräumen sein.
7. Zeigen Sie: Enthält der Grundkörper unendlich viele Elemente, so kann ein Vektorraum über diesem Körper nicht Vereinigung von endlich vielen echten Untervektorräumen sein. (Hinweis: Man schließe durch Widerspruch und verwende das Dirichletsche Schubfachprinzip: Ist V doch Vereinigung von endlichen vielen Untervektorräumen $U_1, \ldots, U_n$, so kann man n so klein wie möglich wählen. In dieser Situation gibt es $u_1 \in U_1$ und $u_2 \in U_2$, die jeweils in keinem anderen Untervektorraum liegen. (Warum?) Für $n+1$ verschiedene Zahlen $\lambda_1, \ldots, \lambda_{n+1} \in K$ betrachte man die Vektoren $u_1 + \lambda_1 u_2, \ldots, u_1 + \lambda_{n+1} u_2$.)
8. In 3.6.10 wurde die Menge der reellen Polynomfunktionen
$$P = \{f : \mathbb{R} \longrightarrow \mathbb{R} | \exists n \in \mathbb{N}\, \exists \alpha_0, \ldots, \alpha_n \in \mathbb{R} \forall x \in \mathbb{R}[f(x) = \sum_{i=0}^{n} \alpha_i x^i]\}$$
eingeführt. Zeigen Sie, daß P ein reeller Vektorraum ist.
9. In 3.6.7 wurde gezeigt, daß die Menge $\mathbb{Z}/(2)$ ein Körper (mit zwei Elementen) ist. Zeigen Sie:
   a) Die Menge $\{0, a, b, c\}$ mit der durch die Gruppentafel
$$\begin{array}{c|cccc} + & 0 & a & b & c \\ \hline 0 & 0 & a & b & c \\ a & a & 0 & c & b \\ b & b & c & 0 & a \\ c & c & b & a & 0 \end{array}$$
   beschriebenen abelschen Gruppenstruktur und einer geeeigneten Multiplikation mit Skalaren ist ein Vektorraum über dem Körper $\mathbb{Z}/(2)$.
   b) Für jeden $\mathbb{Z}/(2)$-Vektorraum V gilt $\forall v \in V[v + v = 0]$.
   c) Es gibt keine $\mathbb{Z}/(2)$-Vektorraumstruktur auf der Menge $\{0, a, b, c\}$, deren Addition durch die Gruppentafel
$$\begin{array}{c|cccc} + & 0 & a & b & c \\ \hline 0 & 0 & a & b & c \\ a & a & b & c & 0 \\ b & b & c & 0 & a \\ c & c & 0 & a & b \end{array}$$
   beschrieben wird.
10. Sei V ein $\mathbb{R}$-Vektorraum. Zeigen Sie: Mit der durch
$$\begin{aligned} (v, w) + (x, y) &= (v + x, w + y) \\ (\alpha + \beta i) \cdot (v, w) &= (\alpha v - \beta w, \alpha w + \beta v) \end{aligned}$$
für $v, w, x, y \in V$ und $\alpha, \beta \in \mathbb{R}$ beschriebenen Addition und Multiplikation mit Skalaren ist $V \times V$ ein $\mathbb{C}$-Vektorraum.

11. Beweisen Sie das modulare Gesetz: Sind V ein K-Vektorraum und U_1, U_2, $U_3 \subset V$ Untervektorräume mit $U_1 \subset U_3$, dann ist $(U_1 + U_2) \cap U_3 = U_1 + (U_2 \cap U_3)$.

5.2 Linearkombinationen, Basen, Dimension

Nachdem wir bisher nur ganze Vektorräume studiert haben, soll jetzt das eigentliche Rechnen mit Vektoren betrachtet werden. Da wir in jedem Vektorraum Vektoren addieren können, auch mehr als zwei Vektoren, und diese Vektoren zudem mit Elementen aus dem Körper multipliziert werden können, können wir allgemeinere Ausdrücke der folgenden Form bilden:

$$\sum_{i=1}^{n} \alpha_i v_i = \alpha_1 v_1 + \ldots + \alpha_n v_n.$$

Im vorhergehenden Abschnitt haben wir solche Summen sogar für Familien mit unendlich vielen Vektoren gebildet, sofern nur endlich viele der Vektoren von Null verschieden sind. Das kann in einem Ausdruck der Form $\sum_{i \in I} \alpha_i v_i$ z.B. dadurch geschehen, daß nur endlich viele Skalare α_i von Null verschieden sind, oder auch dadurch, daß nur endlich viele Vektoren v_i von Null verschieden sind. Daher setzen wir allgemein, wenn wir eine Summe der Form $\sum_{i \in I} \alpha_i v_i$ schreiben, immer stillschweigend voraus, daß nur endlich viele der Produkte $\alpha_i v_i \neq 0$ sind.

Definition 5.2.1 Eine Summe $\sum_{i \in I} \alpha_i v_i$ mit $v_i \in V$ und $\alpha_i \in K$, wobei nur endlich viele $\alpha_i v_i \neq 0$ sind, wird *Linearkombination* der Vektoren v_i mit den *Koeffizienten* α_i genannt. Für $n = 1$ definieren wir

$$\sum_{i=1}^{1} \alpha_i v_i := \alpha_1 v_1.$$

Für $n = 0$, also eine leere Summe (ohne Summanden), definieren wir

$$\sum_{i \in} \alpha_i v_i = \sum_{i=1}^{0} \alpha_i v_i := 0.$$

Definition 5.2.2 1. Seien V ein Vektorraum, $X \subset V$ eine Teilmenge und $v \in V$ ein Vektor. Wir sagen, daß v von X *linear abhängig* ist, wenn es eine Darstellung $v = \sum_{v \in X} \alpha_v v$ gibt, oder genauer, wenn es Vektoren $v_1, \ldots, v_n \in X$ und Skalare $\alpha_1, \ldots, \alpha_n \in K$ so gibt, daß $v = \sum_{i=1}^{n} \alpha_i v_i$ gilt, d.h. daß v eine Linearkombination der Vektoren $v_1, \ldots, v_n$ ist. Wir lassen dabei auch die Möglichkeiten $n = 0$ und $n = 1$ zu.
2. Die Menge der von einer Teilmenge $X \subset V$ linear abhängigen Vektoren heißt die von X *erzeugte* oder *aufgespannte Menge* $\langle X \rangle$. Ist $X = \emptyset$, so gilt $\langle \emptyset \rangle = \{0\}$.

3. Eine Teilmenge $X \subset V$ eines Vektorraumes V wird *Erzeugendenmenge* genannt, wenn jeder Vektor $v \in V$ linear abhängig von X ist, d.h. wenn jeder Vektor aus V eine Linearkombination von Vektoren aus X ist. Mit anderen Worten gilt $V = \langle X \rangle$ genau dann, wenn X eine Erzeugendenmenge von V ist.
4. Der Vektorraum V heißt *endlich erzeugt*, wenn es eine endliche Erzeugendenmenge für V gibt.
5. Eine Teilmenge $X \subset V$ eines Vektorraumes V heißt *linear unabhängig*, wenn keiner der Vektoren aus X von den übrigen Vektoren aus X linear abhängig ist. Gleichbedeutend damit ist die folgende Aussage: Sind $\alpha_v \in K$ für alle $v \in X$ so gegeben, daß $\sum_{v \in X} \alpha_v v = 0$ gilt, dann gilt $\alpha_v = 0$ für alle $v \in X$. Ist X nicht linear unabhängig, so heißt X *linear abhängig*.

Aus dieser Definition folgt unmittelbar, daß jede Teilmenge einer linear unabhängigen Menge wieder linear unabhängig ist, und daß jede Obermenge einer Erzeugendenmenge wieder eine Erzeugendenmenge ist. Außerdem ist jede Obermenge einer linear abhängigen Menge linear abhängig.

Lemma 5.2.3 *Sei $X \subset V$ eine Teilmenge des Vektorraumes V. Dann ist die von X erzeugte Menge $\langle X \rangle$ ein Untervektorraum von V.*

Beweis. Sicher ist $0 \in \langle X \rangle$ als Linearkombination ohne Summanden. Also ist $\langle X \rangle \neq \emptyset$. Seien $v_j = \sum_{x \in X} \alpha_{xj} x$ mit $j = 1, 2$ zwei beliebige Elemente aus $\langle X \rangle$. Sei weiter $\beta \in K$. Dann sind

$$v_1 + v_2 = (\sum_{x \in X} \alpha_{x1} x) + (\sum_{x \in X} \alpha_{x2} x) = \sum_{x \in X} (\alpha_{x1} + \alpha_{x2}) x$$

und

$$\beta v_1 = \beta(\sum_{x \in X} \alpha_{x1} x) = \sum_{x \in X} (\beta \alpha_{x1}) x$$

ebenfalls Elemente von $\langle X \rangle$. Also ist $\langle X \rangle$ ein Untervektorraum.

Beispiele 5.2.4 1. Im Vektorraum $\mathbb{R}^2$ gilt

$$(-7, 6) = 3(1, 2) - 5(2, 0).$$

Also ist $(-7, 6)$ von der Menge $\{(1, 2), (2, 0)\}$ linear abhängig. Ebenso ist $(5, 7)$ von der Menge $\{(1, 0), (1, 1), (0, 1)\}$ linear abhängig. Es gilt nämlich

$$(5, 7) = 2(1, 0) + 3(1, 1) + 4(0, 1).$$

Eine unserer Aufgaben wird es sein, die Koeffizienten (hier 2, 3 und 4) zu finden, die in einer Linearkombination zur Darstellung von $(5, 7)$ verwendet werden müssen. Sie werden im allgemeinen nicht eindeutig bestimmt sein, d. h. wir werden verschiedene Wahlmöglichkeiten haben.

2. Im Vektorraum $\mathbb{R}^3$ gilt

$$(-1, 2, 5) = 2(1, 2, 3) + (-1)(3, 2, 1).$$

Jedoch ist $(1, 0, 0)$ von der Menge $\{(1, 2, 3), (3, 2, 1)\}$ nicht linear abhängig. Gäbe es nämlich α und β mit

$$(1, 0, 0) = \alpha(1, 2, 3) + \beta(3, 2, 1),$$

so wären α und β Lösungen des linearen Gleichungssystems

$$\begin{aligned} 1\alpha + 3\beta &= 1, \\ 2\alpha + 2\beta &= 0, \\ 3\alpha + 1\beta &= 0. \end{aligned}$$

Dieses Gleichungssystem hat aber offenbar keine Lösung. (Die letzten beiden Gleichungen lassen sich nur durch $\alpha = \beta = 0$ erfüllen. Damit kann aber die erste Gleichung nicht erfüllt werden.)

3. Wir bezeichnen mit $e_i = (0, \ldots, 1, \ldots, 0)$ die Vektoren in K^n mit 1 an der i-ten Stelle und 0 an allen anderen Stellen. Die Menge der Vektoren $\{e_i | \, i = 1, 2, \ldots, n\}$ ist eine Erzeugendenmenge für den Vektorraum K^n, denn es gilt für alle Vektoren

$$(\xi_1, \ldots, \xi_n) = \sum_{i=1}^{n} \xi_i e_i.$$

Weiterhin bemerken wir, daß $\{e_i\}$ eine linear unabhängige Menge ist. Nehmen wir nämlich ein e_i aus dieser Menge heraus, so kann man durch Linearkombinationen der übrigen e_j nur solche Vektoren darstellen, die an der i-ten Stelle eine 0 haben, insbesondere also nicht e_i.

4. Sei I eine beliebige Menge. Wir definieren eine Abbildung $\delta : I \times I \longrightarrow \{0, 1\}$ durch

$$\delta(i, j) := \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j. \end{cases}$$

Diese Abbildung heißt auch *Kronecker*[1] *Funktion*, *Kronecker Delta* oder *Kronecker Symbol*. Man schreibt oft $\delta_{ij} := \delta(i, j)$. Die Vektoren

$$e_i := (\delta_{ij} | j \in I)$$

liegen in dem Vektorraum $K^{(I)}$, denn es ist nur eine Komponente von Null verschieden. Die Menge der Vektoren $\{e_i | i \in I\}$ ist eine Erzeugendenmenge für den Vektorraum $K^{(I)}$, denn es gilt für alle Vektoren

$$(\xi_i) = \sum_{j \in I} \xi_j e_j,$$

[1] Leopold Kronecker (1823–1891)

da in (ξ_i) nur endlich viele Komponenten von Null verschieden sind. Weiterhin bemerken wir, daß $\{e_i|i \in I\}$ eine linear unabhängige Menge ist. Nehmen wir nämlich ein e_i aus dieser Menge heraus, so kann man durch Linearkombinationen der übrigen e_j nur solche Vektoren darstellen, die an der i-ten Stelle eine 0 haben, insbesondere also nicht e_i. Man beachte, daß die e_i jedoch nicht den Vektorraum K^I erzeugen, weil es Vektoren (Familien) mit unendlich vielen von Null verschiedenen Koeffizienten geben kann, jedoch jede (endliche!) Linearkombination $\sum_{i \in I} \alpha_i e_i$ nur endlich viele von Null verschiedene Koeffizienten haben kann. Die e_i liegen aber auch in $K^{(I)}$ und sind offenbar sogar eine Erzeugendenmenge für diesen Untervektorraum.

5. Sei nun $V := \{(\xi_1, \xi_2, \xi_3, \xi_4)|\xi_1, \xi_2, \xi_3, \xi_4 \in \mathbb{R}, \xi_1 + \xi_2 + \xi_3 + \xi_4 = 0\}$. Dann liegen die Vektoren $(-1, 1, 0, 0)$, $(-1, 0, 1, 0)$ und $(-1, 0, 0, 1)$ in V und bilden eine Erzeugendenmenge für V. Weiter ist diese Menge auch linear unabhängig. Den Beweis hierfür überlassen wir dem Leser zur Übung.

Lemma 5.2.5 *Die Menge B im Vektorraum V ist genau dann linear abhängig, wenn (mindestens) ein Vektor $w \in B$ Linearkombination der übrigen Vektoren ist:*

$$w = \sum_{v \in B, v \neq w} \alpha_v v.$$

Wenn die letzte Bedingung erfüllt ist, dann sind die von B und von $B \setminus \{w\}$ erzeugten Mengen gleich:

$$\langle B \rangle = \langle B \setminus \{w\} \rangle.$$

Beweis. Die erste Behauptung ist genau die Definition der linearen Abhängigkeit. In der zweiten Behauptung ist die Inklusion $\langle B \setminus \{w\} \rangle \subset \langle B \rangle$ trivialerweise wahr. Sei also $u = \sum_{v \in B} \beta_v v$ und $w = \sum_{v \in B, v \neq w} \alpha_v v$. Setzen wir w in die erste Summe ein, so erhalten wir

$$u = \sum_{v \in B, v \neq w} (\beta_v + \beta_w \alpha_v) v,$$

also die geforderte Eigenschaft.

Wir beachten insbesondere, daß eine Menge B linear abhängig ist, wenn sie den Nullvektor enthält. Andererseits ist die leere Menge linear unabhängig. Sie ist im Nullvektorraum auch eine Erzeugendenmenge.

Definition 5.2.6 Sei V ein Vektorraum. Eine Menge $B \subset V$ heißt eine *Basis* für V, wenn die Menge B linear unabhängig und eine Erzeugendenmenge von V ist.

Satz 5.2.7 *Die Menge $\{e_i|i \in I\}$ ist eine Basis für den Vektorraum $K^{(I)}$, genannt die* kanonische Basis*. Insbesondere ist $\{e_i|i \in \{1, \ldots, n\}\}$ eine Basis für K^n.*

Beweis. Die Aussage wurde schon im Beispiel 5.2.4 4. gezeigt.

Da der Begriff der Basis für die lineare Algebra und analytische Geometrie von höchster Wichtigkeit ist, wollen wir weitere dazu äquivalente Eigenschaften studieren.

Satz 5.2.8 *Sei V ein Vektorraum und $B \subset V$ eine Teilmenge. B ist genau dann eine Basis von V, wenn es zu jedem Vektor $w \in V$ eindeutig durch w bestimmte Koeffizienten $\alpha_v \in K$ mit $v \in B$ gibt, so daß*

$$w = \sum_{v \in B} \alpha_v v.$$

Beweis. Sei B eine Basis. Dann läßt sich jedes $w \in V$ als Linearkombination der $b \in B$ darstellen, weil B eine Erzeugendenmenge bildet. Um die Eindeutigkeit der Darstellung im Satz, d.h. der Koeffizienten in der Darstellung, zu zeigen, nehmen wir an $w = \sum_{v \in B} \alpha_v v = \sum_{v \in B} \beta_v v$. Wir erhalten durch Umstellen der Gleichung

$$\sum_{v \in B} (\alpha_v - \beta_v) v = 0.$$

Da B eine linear unabhängige Menge ist, folgt $\alpha_v - \beta_v = 0$ für alle $v \in B$ oder $\alpha_v = \beta_v$ für alle $v \in B$. Also ist die Darstellung von w als Linearkombination der $v \in B$ eindeutig.

Sei umgekehrt jeder Vektor eindeutig als Linearkombination der $v \in B$ darstellbar, so ist B eine Erzeugendenmenge. Weiter ist der Nullvektor 0 eindeutig darstellbar als $0 = \sum_{v \in B} 0v$, d.h. für eine Linearkombination $0 = \sum_{v \in B} \alpha_v v$ muß notwendig $\alpha_v = 0$ für alle $v \in B$ gelten. Damit ist B aber linear unabhängig und eine Erzeugendenmenge, also eine Basis.

Wir werden unten für beliebige Vektorräume V zeigen, daß sie immer eine Basis besitzen. Dieser Beweis ist jedoch nicht konstruktiv, führt also nicht zur Angabe einer konkreten Basis. Man hat jedoch häufig das Problem, daß eine konkrete Basis zum Rechnen benötigt wird. Deshalb beweisen wir diesen Satz gesondert in einem besonders einfachen Fall, in dem wir konstruktiv eine Basis angeben können. Diese Basis wird nur endlich viele Elemente haben. Im allgemeinen gibt es aber auch Vektorräume, die lediglich eine unendliche Basis besitzen, die man nicht konkret angeben kann.

Satz 5.2.9 *Jeder endlich erzeugte Vektorraum besitzt eine endliche Basis.*

Beweis. Sei $\{b_1, \ldots, b_n\}$ eine endliche Erzeugendenmenge von V. Wenn die Erzeugendenmenge linear unabhängig ist, dann ist sie eine Basis. Ist sie jedoch linear abhängig, so können wir nach Lemma 5.2.5 ein w aus dieser Menge entfernen, so daß die verbleibende Menge wieder eine Erzeugendenmenge ist. Sie hat dann aber nur noch $n - 1$ Elemente. Nach endlich vielen Schritten muß dieser Prozeß abbrechen mit einer linear unabhängigen Erzeugendenmenge (bei geeigneter Numerierung) $\{b_1, \ldots, b_m\}$. Diese ist dann eine Basis für V.

Wir haben in dem Beweis sogar mehr bewiesen, nämlich die Aussage, daß jede endliche Erzeugendenmenge eine Basis enthält. Man kann nun mit nichtkonstruktiven Mitteln der Mengenlehre einen wesentlich allgemeineren Satz beweisen. Im Beweis wird das Zornsche Lemma (2.3.7) verwendet. Wir zeigen zunächst

Satz 5.2.10 *(Steinitzscher[2] Austauschsatz) Seien V ein Vektorraum, $E \subset V$ eine Erzeugendenmenge und $X \subset V$ eine linear unabhängige Menge. Dann gibt es eine Teilmenge $F \subset E$ mit $F \cap X = \emptyset$, so daß $X \cup F$ eine Basis von V ist.*

Beweis. Wir bilden die folgende Menge

$$\mathcal{Z} := \{D \subset E | D \cap X = \emptyset \text{ und } D \cup X \text{ linear unabhängig}\}.$$

Diese Menge ist unter der Inklusion eine geordnete Menge. (Sie ist Teilmenge der Potenzmenge von E.)

Wir zeigen, daß jede total geordnete Teilmenge von $\mathcal{Z}$ eine obere Schranke besitzt. Ist die total geordnete Teilmenge leer, so ist eine obere Schranke durch $\emptyset \subset E$, also $\emptyset \in \mathcal{Z}$ gegeben.

Sei $\mathcal{Y} \subset \mathcal{Z}$ nicht leer und total geordnet. Wir bilden $F' := \bigcup\{D \in \mathcal{Y}\}$. Offenbar gilt $D \subset F'$ für alle $D \in \mathcal{Y}$.

Um zu zeigen, daß F' in $\mathcal{Z}$ liegt, weisen wir zunächst $F' \cap X = \emptyset$ nach. Ist $x \in F' \cap X$, so ist x insbesondere in F', also in einer der Teilmengen D, deren Vereinigung F' ist. Also gibt es ein $D \in \mathcal{Y}$ mit $x \in D \cap X$ im Widerspruch zu $D \in \mathcal{Z}$.

Weiter müssen wir zeigen, daß $F' \cup X$ linear unabhängig ist. Seien also $\alpha_v \in K$ mit $v \in F' \cup X$ gegeben, so daß $\sum_{v \in F' \cup X} \alpha_v v = 0$ gilt. Wenn überhaupt Koeffizienten $\alpha_v \neq 0$ in dieser Darstellung auftreten, dann auch solche mit $v \in F'$, weil ja X allein linear unabhängig ist. Andererseits sind nur endlich viele $\alpha_v \neq 0$ möglich. Jedes zugehörige v liegt schon in einem $D \in \mathcal{Y}$, es spielen also endlich viele solche D's eine Rolle. Da die Teilmenge $\mathcal{Y} \subset \mathcal{Z}$ totalgeordnet war, gibt unter diesen endlich vielen D's ein größtes D', in dem dann alle in der Linearkombination verwendeten v's mit von Null verschiedenen Koeffizienten liegen. Da aber $D' \cup X$ linear unabhängig ist, müssen die Koeffizienten alle Null sein. Also ist $F' \cup X$ wiederum linear unabhängig. Damit ist mit F' eine obere Schranke für $\mathcal{Y}$ gefunden.

Nach dem Zornschen Lemma 2 enthält also $\mathcal{Z}$ ein maximales Element F. Nach Definition von $\mathcal{Z}$ gilt schon $F \cap X = \emptyset$ und $F \cup X$ linear unabhängig. Wir zeigen nun noch, daß $F \cup X$ eine Basis bildet. Dazu bleibt nur zu zeigen, daß $F \cup X$ eine Erzeugendenmenge ist. Zunächst bemerken wir, daß es genügt, alle Elemente aus E als Linearkombination von Elemente aus $F \cup X$ darzustellen. Denn kann man jeden Vektor aus E als Linearkombination von Vektoren

[2] Ernst Steinitz (1871–1928)

aus $F \cup X$ erhalten, so kann man auch jeden anderen Vektor aus V als Linearkombination von Vektoren aus $F \cup X$ erhalten. Ist nun $u \in E$ nicht als Linearkombination von Vektoren aus $F \cup X$ darstellbar, dann hat jede Linearkombination der Form $\alpha_u u + \sum_{v \in F \cup X} \alpha_v v = 0$ notwendigerweise die Koeffizienten $\alpha_u = \alpha_v = 0$, denn $\alpha_u \neq 0$ würde nach Division durch α_u eine Darstellung $u = -\sum_{v \in F \cup X} \alpha_u^{-1} \alpha_v v$ ergeben, und die α_v müssen auch alle Null sein, weil $F \cup X$ linear unabhängig ist. Also ist $\{u\} \cup F \cup X$ linear unabhängig. Außerdem ist sicherlich $(\{u\} \cup F) \cap X = \emptyset$ und damit ist $\{u\} \cup F \in \mathcal{Z}$. Da $u \notin F$ und F maximal in $\mathcal{Z}$ ist, haben wir einen Widerspruch erhalten. Somit kann jeder Vektor aus E als Linearkombination von Elementen aus $F \cup X$ geschrieben werden, $F \cup X$ ist also eine Basis von V. (Uff!)

Aus diesem mächtigen Satz mit sehr abstraktem Beweis gehen nun unmittelbar eine Reihe von Folgerungen hervor.

Folgerung 5.2.11 *1. Jeder Vektorraum besitzt eine Basis.*
2. In jeder Erzeugendenmenge eines Vektorraumes ist eine Basis enthalten.
3. Jede linear unabhängige Menge läßt sich zu einer Basis ergänzen.

Beweis. Alle drei Aussagen folgen unmittelbar, wenn wir feststellen, daß jeder Vektorraum eine Erzeugendenmenge und eine linear unabhängige Menge enthält. Der gesamte Vektorraum ist aber eine Erzeugendenmenge und die leere Menge ist eine linear unabhängige Menge. Die leere Menge enthält nämlich keinen Vektor, der als Linearkombination der „übrigen Vektoren“ darstellbar wäre.

Folgerung 5.2.12 *Die folgenden Aussagen für eine Teilmenge B eines Vektorraumes V sind äquivalent:*

1. *B ist eine Basis von V.*
2. *B ist eine minimale Erzeugendenmenge.*
3. *B ist eine maximal linear unabhängige Menge.*

Beweis. Wir stellen zunächst fest, daß keine echte Teilmenge E von B Erzeugendenmenge von V ist, denn ein Vektor $b \in B \setminus E$ läßt sich nicht als Linearkombination von Elementen aus E darstellen, da B linear unabhängig ist. Weiter ist jede echte Obermenge X von B linear abhängig, denn die hinzukommenden Elemente sind Linearkombinationen von Elementen aus B. Damit ist 1. $\Longrightarrow$ 2. und 1. $\Longrightarrow$ 3. gezeigt.

Wenn B eine minimale Erzeugendenmenge ist, so enthält sie eine Basis C. Diese ist auch Erzeugendenmenge, muß also wegen der Minimalität mit B übereinstimmen. Damit gilt 2. $\Longrightarrow$ 1.

Wenn B eine maximal linear unabhängige Menge ist, dann läßt sie sich zu einer Basis C ergänzen. Da C auch linear unabhängig ist, muß $B = C$ und damit 3. $\Longrightarrow$ 1. gelten.

Wir beschäftigen uns jetzt mit dem Abzählen von Elementen in linear unabhängigen Mengen und Basen. Dazu wollen wir im Rest dieses Abschnitts voraussetzen, daß die betrachteten Vektorräume endlich erzeugt sind. Einige der Aussagen würden auch aus dem Steinitzschen Austauschsatz folgen. Für sie ist jedoch auch eine explizite Konstruktion von Interesse.

Lemma 5.2.13 *Wenn der Vektorraum V von der Menge $\{b_1, \ldots, b_n\} \subset V$ erzeugt wird und die Menge $\{c_1, \ldots, c_m\}$ in V linear unabhängig ist, dann gilt $m \leq n$. Weiterhin gibt es Vektoren $b_{m+1}, \ldots, b_n$ aus der Menge der b_i (nach geeigneter Umnumerierung), so daß V von der Menge $\{c_1, \ldots, c_m, b_{m+1}, \ldots, b_n\}$ erzeugt wird.*

Beweis. Wir beweisen das Lemma durch vollständige Induktion nach m. Für $m = 0$ ist nichts zu zeigen. Gelte das Lemma für alle linear unabhängigen Mengen von m Vektoren. Sei die Menge $\{c_1, \ldots, c_{m+1}\}$ linear unabhängig. Dann ist auch $\{c_1, \ldots, c_m\}$ linear unabhängig. Nach Induktionsannahme gibt es Vektoren $b_{m+1}, \ldots, b_n$, so daß V durch $\{c_1, \ldots, c_m, b_{m+1}, \ldots, b_n\}$ erzeugt wird. Insbesondere gilt

$$c_{m+1} = \alpha_1 c_1 + \ldots + \alpha_m c_m + \alpha_{m+1} b_{m+1} + \ldots + \alpha_n b_n.$$

Wir zeigen $m+1 \leq n$. Wenn das nicht der Fall ist, dann ist nach Induktionsvoraussetzung $m = n$, also treten in der Summe keine Summanden der Form $\alpha_i b_i$ auf. Damit würde aber die Gleichung für c_{m+1} zeigen, daß die Menge $\{c_1, \ldots, c_{m+1}\}$ linear abhängig ist. Das ist ein Widerspruch zur Voraussetzung. Fügen wir jetzt den Vektor c_{m+1} zu der Liste hinzu, so erhalten wir eine Erzeugendenmenge $\{c_1, \ldots, c_{m+1}, b_{m+1}, \ldots, b_n\}$. Wegen obiger Darstellung von c_{m+1} ist die Menge linear abhängig. Wir können einen Vektor aus der Liste fortlassen, nämlich den ersten Vektor, der eine Linearkombination der vorhergehenden Vektoren ist. Da nun $\{c_1, \ldots, c_{m+1}\}$ linear unabhängig ist, ist dieser Vektor aus den $b_{m+1}, \ldots, b_n$ zu wählen. Wir erhalten schließlich eine Erzeugendenmenge von n Vektoren, die die Vektoren $c_1, \ldots, c_{m+1}$ enthält.

Man beachte, daß das vorhergehende Lemma im Gegensatz zum Steinitzschen Austauschsatz konstruktiv ist. In der Tat kann man die Beweisschritte als Algorithmus für die Berechnung einer Basis verwenden, wenn eine endliche Erzeugendenmenge gegeben ist.

Folgerung 5.2.14 *1. Sei V ein endlich erzeugter Vektorraum. Dann haben je zwei Basen von V gleich viele Elemente.*
2. Wenn V eine Basis mit unendlich vielen Elementen besitzt, dann hat auch jede andere Basis von V unendlich viele Elemente.

Beweis. Seien $\{b_1, \ldots, b_n\}$ und $\{c_i | i \in I\}$ zwei Basen des Vektorraumes V. Die Indexmenge I darf hier zunächst sogar unendlich sein. Dann ist jede endliche Teilmenge von $\{c_i\}$ linear unabhängig, besitzt also nach dem vorhergehenden Lemma höchstens n Elemente. Folglich ist auch $\{c_i|i \in I\}$ eine endliche Menge mit höchstens n Elementen. Damit ist 2. bewiesen. Vertauschen wir jetzt die Rollen der b_i und der c_i, so muß nochmals nach dem vorstehenden Lemma die Menge der c_i mindestens n Elemente besitzen. Damit ist auch 1. bewiesen.

Wenn $\{b_1, \ldots, b_n\}$ eine Basis für den Vektorraum V ist, so ist die Zahl n nur durch den Vektorraum selbst bestimmt. V kann zwar viele verschiedene Basen haben, jedoch haben alle Basen dieselbe Anzahl von Elementen. Damit ist n eine interessante Invariante für den Vektorraum V, die unabhängig von der gewählten Basis ist. Das führt uns zu der

Definition 5.2.15 Wenn der Vektorraum V eine endliche Basis besitzt, so wird die Anzahl n der Vektoren der Basis *Dimension* genannt: $\dim V = n$. Sonst sagen wir, daß die Dimension unendlich ist: $\dim V = \infty$.

Folgerung 5.2.16 *Sei V ein n-dimensionaler Vektorraum. Dann gelten:*

1. *je $n + 1$ Vektoren sind linear abhängig,*
2. *V kann nicht durch eine Menge von $n - 1$ Vektoren erzeugt werden.*

Beweis folgt unmittelbar aus 5.2.13.

Folgerung 5.2.17 *Sei V ein n-dimensionaler Vektorraum. Dann gelten:*

1. *jede linear unabhängige Menge von n Vektoren ist ein Basis,*
2. *jede Erzeugendenmenge für V von n Vektoren ist eine Basis.*

Beweis. 1. Jede linear unabhängige Menge läßt sich zu einer Basis vervollständigen, die jedoch nach 5.2.16 1. nicht mehr als n Elemente haben kann. Also ist die gegebene Menge selbst schon eine Basis.

2. Jede Erzeugendenmenge für V enthält nach 5.2.16 2. eine Basis, die jedoch nach nicht weniger als n Elemente haben kann. Also ist die gegebene Menge selbst schon eine Basis.

Satz 5.2.18 *Sei V ein n-dimensionaler Vektorraum und U ein Untervektorraum von V. Dann ist U ein endlichdimensionaler Vektorraum, und es gilt $\dim U \leq n$. Ist $\dim U = \dim V$, so ist $U = V$.*

Beweis. Eine Basis von U besteht aus höchstens n linear unabhängigen Vektoren wegen 5.2.16 1. Besitzt U eine Basis aus $n = \dim V$ Vektoren, so ist diese nach 5.2.17 1. schon eine Basis von V, also erzeugt sie ganz V. Daher folgt $U = V$.

Übungen 5.2.19 1. Im Vektorraum $\mathbb{R}^3$ betrachten wir die Vektoren $v = (1, 2, 1)$ und $w = (1, 2, -1)$.

a) Zeigen Sie, daß der Vektor $(1, 2, 5)$ eine Linearkombination von v und w ist.

b) Zeigen Sie, daß der Vektor $(3,4,0)$ keine Linearkombination von v und w ist.

2. a) Entscheiden Sie, ob die Vektoren $(1,0,2)$, $(2,1,4)$ und $(1,-1,0)$ eine Basis des $\mathbb{R}^3$ bilden.
   b) Entscheiden Sie, ob die Vektoren $(1,0,2)$, $(3,1,-1)$ und $(-2,-1,3)$ eine Basis des $\mathbb{R}^3$ bilden.
3. Sei $V = \mathbb{R}^4$ und $U = \{(2x, x, 7y, 7z) \mid x + 2y + 3z = 0\} \subset \mathbb{R}^4$. Bestimmen Sie eine Basis von V, die eine Basis von U enthält.
4. Sei V ein n-dimensionaler Vektorraum. Dann gibt es eine Teilmenge $T \subset V$, so daß für jedes $v \in T$ die Menge $T \setminus \{v\}$ eine Basis von V ist.
5. Sei V ein endlichdimensionaler Vektorraum. U_1 und U_2 seien zwei Untervektorräume mit $U_1 \cap U_2 = \{0\}$. Zeigen Sie: $\dim V \geq \dim U_1 + \dim U_2$.
6. Sei U ein Untervektorraum des endlichdimensionalen Vektorraums V mit $\dim U = \dim V$. Zeigen Sie $U = V$.
7. U_1 und U_2 seien Untervektorräume von $\mathbb{R}^3$ mit $\dim U_1 = 1$ und $\dim U_2 = 2$. Zeigen Sie
$$U_1 + U_2 = \mathbb{R}^3 \Leftrightarrow U_1 \cap U_2 = \{0\}.$$
8. Zeigen Sie, daß die Polynomfunktionen $1, x, x^2, x^3$ in $\mathbb{R}^\mathbb{R}$ linear unabhängig sind. (Hinweis: Betrachten Sie die Ableitungen von Linearkombinationen der Polynomfunktionen.)
9. Bestimmen Sie eine linear unabhängige Erzeugendenmenge für den Untervektorraum $\langle 1, x^2 + 2x, x^2 - 2x, 2x - 1\rangle \subseteq \mathbb{R}^\mathbb{R}$.
10. Bestimmen Sie eine Basis für den Untervektorraum P_3 der Polynomfunktionen vom Grad kleiner oder gleich 3 im Vektorraum aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$.
11. Sei $K = \mathbb{F}_3 = \mathbb{Z}/(3)$ der Körper mit drei Elementen. Bestimmen Sie eine Basis für den Untervektorraum aller Polynomfunktionen vom Grad kleiner oder gleich 3 im K-Vektorraum $\mathbb{F}_3^{\mathbb{F}_3}$ aller Abbildungen von $\mathbb{F}_3$ nach $\mathbb{F}_3$.
12. Wieviele Basen hat ein dreidimensionaler $\mathbb{Z}/(2)$-Vektorraum?
13. Zeigen Sie, daß $\{\sin(x), \cos(x)\}$ eine linear unabhängige Menge in $\mathbb{R}^\mathbb{R}$ bildet. (Verwenden Sie dazu Ihre Schulkenntnisse über die Funktionen $\sin(x)$ und $\cos(x)$.)
14. Sei U der Untervektorraum von $\mathbb{R}^\mathbb{R}$, der von $\sin(x)$ und $\cos(x)$ erzeugt wird. Zeigen Sie, daß für jeden Winkel φ gilt $\sin(\varphi + x) \in U$ und $\cos(\varphi + x) \in U$. Zeigen Sie außerdem, daß $\{\sin(\varphi + x), \cos(\varphi + x)\}$ eine Basis für U bildet.
15. Zeigen Sie, daß $\{\sin(x), \cos(x), \tan(x)\}$ eine linear unabhängige Menge in $\mathbb{R}^\mathbb{R}$ bildet.
16. Sei V ein endlichdimensionaler Vektorraum und U ein Untervektorraum von V. $v_1, \ldots, v_n$ seien linear unabhängige Vektoren von V. Zeigen Sie daß folgende Aussagen äquivalent sind:
   a) Ergänzt man $v_1, \ldots, v_n$ durch eine Basis von U, so ist die entstehende Menge linear unabhängig.

b) Der Nullvektor ist der einzige Vektor, der gleichzeitig Linearkombination von $v_1, \ldots, v_n$ ist und in U liegt.

17. Entscheiden Sie, welche der folgenden Aussagen richtig sind (ja/nein).
    a) Sei X eine Menge von linear unabhängigen Vektoren in einem Vektorraum. Dann ist auch jede Teilmenge von X linear unabhängig.
    b) Sei X ein Erzeugendensystem des Vektorraumes V. Dann ist auch jede Teilmenge von X ein Erzeugendensystem von V.
    c) Jede Menge von Vektoren in einem Vektorraum läßt sich zu einer Basis ergänzen.
18. Wir betrachten die reellen Zahlen als Vektorraum über den rationalen Zahlen (vgl. Beispiel 5.1.2 6.). Zeigen Sie:
    a) die Vektoren 1, $\sqrt{2}$ und $\sqrt{3}$ sind linear unabhängig (Hinweis: Binomische Formel, Primfaktorzerlegung: Kann die Quadratwurzel einer Primzahl rational sein? Man sollte u.a. zeigen, daß $\sqrt{6}$ irrational ist.)
    b) $\sqrt{7}$ ist keine Linearkombination von $\sqrt{3}$ und $\sqrt{5}$.
19. Seien $v_1, \ldots, v_n$ linear abhängige Vektoren in dem Vektorraum V über dem Körper K mit der Eigenschaft, daß je $n-1$ von diesen Vektoren linear unabhängig sind. Seien

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$$

und

$$\mu_1 v_1 + \ldots + \mu_n v_n = 0$$

zwei Linearkombinationen, bei denen jeweils nicht alle Koeffizienten Null sind. Zeigen Sie, daß es eine Zahl $\lambda \in K$ mit der Eigenschaft

$$\mu_j = \lambda \lambda_j$$

für alle $j = 1, \ldots, n$ gibt.

5.3 Direkte Summen

In Abschnitt 1 haben wir schon Summen von Untervektorräumen diskutiert. Hier wollen wir einen besonders schönen Spezialfall davon betrachten.

Definition 5.3.1 Seien U_1 und U_2 Untervektorräume des Vektorraumes V. Wenn $U_1 \cap U_2 = 0$ und $U_1 + U_2 = V$ gelten, dann heißt V eine *direkte Summe* der beiden Untervektorräume. Wir schreiben $V = U_1 \oplus U_2$. Weiter heißen U_1 und U_2 *direkte Summanden* von V. U_2 heißt *direktes Komplement* zu U_1 in V.

Der folgende Satz verhilft uns zu einer Vielzahl von Beispielen. Er zeigt in der Tat, daß jeder Untervektorraum als direkter Summand in einer direkten Summe auftritt. Das ist eine ganz besondere Eigenschaft von Vektorräumen, die aus der Existenz einer Basis für jeden Vektorraum folgt.

Satz 5.3.2 *Sei U ein Untervektorraum des Vektorraumes V. Dann gibt es ein direktes Komplement zu U.*

Beweis. Sei $\{u_i | i \in I\}$ eine Basis für U. Da diese Menge linear unabhängig ist, kann sie nach 5.2.11 3. zu einer Basis $\{u_i | i \in I\} \dot{\cup} \{v_j | j \in J\}$ fortgesetzt werden. Sei U' der von $\{v_j | j \in J\}$ (als Basis) erzeugte Untervektorraum. Wir zeigen, daß U' ein direktes Komplement zu U ist. Sei $v \in V$ gegeben. v hat eine Basisdarstellung

$$v = (\sum_{i \in I} \alpha_i u_i) + (\sum_{j \in J} \beta_j v_j).$$

Die beiden Klammerausdrücke liegen jedoch in U bzw. in U'. Damit ist $V = U + U'$. Um $U \cap U' = 0$ zu zeigen, wählen ein $v \in U \cap U'$. Dann hat v zwei Basisdarstellungen

$$v = (\sum_{i \in I} \alpha_i u_i) + 0 = 0 + (\sum_{j \in J} \beta_j v_j),$$

weil v sowohl in U als auch in U' liegt. Wir haben bei den Darstellungen die Basen von U bzw. U' verwendet und noch eine Linearkombination der restlichen Basiselemente, die Null ergibt, hinzugefügt. Wegen 5.2.8 stimmen die Koeffizienten überein: $\alpha_i = 0$ für alle $i \in I$ und $\beta_j = 0$ für alle $j \in J$. Damit ist aber $v = 0$ und $U \cap U' = 0$.

Es ist nicht nur die direkte Summe von zwei Untervektorräumen interessant, sondern auch die direkte Summe beliebig vieler Untervektorräume. Die Definition hierfür ist komplizierter, als die Definition einer direkten Summe von zwei Vektorräumen. Deswegen haben wir den einfachen Fall zunächst gesondert betrachtet.

Definition 5.3.3 Sei $(U_i | i \in I)$ eine Familie von Untervektorräumen des Vektorraumes V. Wenn

$$\sum_{i \in I} U_i = V$$

und

$$\forall j \in I [U_j \cap \sum_{i \in I, i \neq j} U_i = 0]$$

gelten, dann heißt V eine *(innere) direkte Summe* der Untervektorräume U_i. Wir schreiben $V = \bigoplus_{i \in I} U_i$. Weiter heißen die Untervektorräume U_i *direkte Summanden* von V. Wenn für die Indexmenge $I = \{1, \ldots, n\}$ gilt, so schreiben wir auch $V = U_1 \oplus \ldots \oplus U_n$.

Weil der Begriff der direkten Summe beinahe genauso stark ist, wie der einer Basis, wollen wir eine Reihe von Eigenschaften finden, die dazu äquivalent sind.

Satz 5.3.4 *Sei $(U_i|i \in I)$ eine Familie von Untervektorräumen des Vektorraumes V. V ist genau dann direkte Summe der U_i, wenn sich jeder Vektor $v \in V$ auf genau eine Weise als Summe einer Familie von Vektoren $(u_i \in U_i|i \in I)$ schreiben läßt: $v = \sum_{i \in I} u_i$.*

Beweis. Sei $V = \bigoplus_{i \in I} U_i$ und habe $v \in V$ zwei Summendarstellungen

$$v = \sum_{i \in I} u_i = \sum_{i \in I} u'_i.$$

Dann ist $u_j - u'_j = \sum_{i \in I, i \neq j} u'_i - u_i = 0$, also $u_j = u'_j$ für alle $j \in I$. Sei umgekehrt jeder Vektor $v \in V$ eindeutig als Summe darstellbar. Zunächst ist dann sicher jeder Vektor $v \in V$ als Summe darstellbar, also gilt $V = \sum_{i \in I} U_i$. Ist aber $v \in U_j \cap \sum_{i \in I, i \neq j} U_i$, so gilt $v = u_j + 0 = 0 + \sum_{i \in I, i \neq j} u_i$ für geeignete Vektoren $u_j (= v) \in U_j$ und $u_i \in U_i$. Wegen der Eindeutigkeit folgt daraus $u_j = 0 = v$, also gilt $U_j \cap \sum_{i \in I, i \neq j} U_i = 0$.

Folgerung 5.3.5 *Sei B eine Basis für den Vektorraum V. Dann ist V eine direkte Summe der eindimensionalen Untervektorräume $(Kb|b \in B)$:*

$$V = \bigoplus_{b \in B} Kb.$$

Beweis folgt mit dem vorhergehenden Satz unmittelbar aus der Eigenschaft, daß jeder Vektor eine eindeutige Basisdarstellung hat.

Satz 5.3.6 *Sei $(U_i|i = 1, \ldots, n)$ eine Familie von Untervektorräumen des endlichdimensionalen Vektorraumes V. V ist genau dann direkte Summe der U_i, wenn $\sum_{i=1}^n U_i = V$ und $\sum_{i=1}^n \dim U_i = \dim V$ gelten.*

Beweis. Seien $B_i = \{b_{i1}, \ldots, b_{ik_i}\}$ Basen für die Untervektorräume U_i (für $i = 1, \ldots, n$). Dann gilt $\dim U_i = k_i$. Sei $B := \bigcup_{i=1}^n B_i$. Da V Summe der U_i ist und die U_i von den Mengen B_i erzeugt werden, ist B offenbar eine Erzeugendenmenge für V.

„$\Longrightarrow$“: Sei V eine direkte Summe der U_i. Dann ist die Vereinigung $B := \bigcup_{i=1}^n B_i$ eine disjunkte Vereinigung. Ist nämlich $i \neq j$ und $b \in B_i \cap B_j \subset U_i \cap U_j \subset \sum_{i=1, i \neq j}^n U_i \cap U_j = 0$, so ist $b = 0$, kann also kein Basiselement sein.

Wir zeigen, daß B eine Basis für V ist. Wenn wir das gezeigt haben, ist nämlich $\sum_{i=1}^n \dim U_i = \sum_{i=1}^n k_i = |B| = \dim V$. Sei $\sum_{i=1}^n \sum_{k=1}^{k_i} \alpha_{ik} b_{ik} = 0$, dann ist wegen der eindeutigen Darstellbarkeit 5.3.4 $\sum_{k=1}^{k_i} \alpha_{ik} b_{ik} = 0$ für alle $i = 1, \ldots, n$. Da B_i Basis ist, folgt also $\alpha_{ik} = 0$ für alle i und k. Damit ist B linear unabhängig und folglich eine Basis.

„$\Longleftarrow$“: Gelte nun $\sum_{i=1}^n U_i = V$ und $\sum_{i=1}^n \dim U_i = \dim V$. Da B eine Erzeugendenmenge ist, gilt $|B| \geq \dim V = \sum_{i=1}^n \dim U_i \geq |B|$, also ist $|B| = \dim V$ und B eine disjunkte Vereinigung der B_i. Insbesondere ist B eine Basis von V nach 5.2.17 2. Sei nun $u_j = \sum_{i=1,i\neq j}^n u_i \in U_j \cap \sum_{i=1,i\neq j}^n U_i$, so erhalten wir als Basisdarstellung

$$u_j = \sum_{k=1}^{k_j} \alpha_{jk} b_{jk} + 0 = 0 + \sum_{i=1,i\neq j}^{n} \sum_{l=1}^{k_i} \alpha_{il} b_{il}.$$

Da B eine Basis ist und wir zwei Basisdarstellungen haben, verschwinden alle Koeffizienten. Also ist $u_j = 0$. Das gilt für alle $j = 1, \ldots, n$. Daher ist $U_j \cap \sum_{i=1,i\neq j}^n U_i = 0$. Es gilt somit $V = U_1 \oplus \ldots \oplus U_n$.

Folgerung 5.3.7 *Seien U_1, U_2 Untervektorräume des endlichdimensionalen Vektorraumes V. Äquivalent sind*

1. $V = U_1 \oplus U_2$.
2. $V = U_1 + U_2$ *und* $\dim V = \dim U_1 + \dim U_2$.
3. $0 = U_1 \cap U_2$ *und* $\dim V = \dim U_1 + \dim U_2$.

Beweis. Es ist nur zu zeigen, daß aus 3. folgt $V = U_1 + U_2$. Basen B_1 von U_1 und B_2 von U_2 sind wegen $0 = U_1 \cap U_2$ sicher disjunkt. Es folgt sogar, daß $B := B_1 \dot\cup B_2$ linear unabhängig ist. Wegen der Bedingung $\dim V = \dim U_1 + \dim U_2$ ist B dann aber eine Basis und damit $V = U_1 + U_2$.

Satz 5.3.8 *(Dimensionssatz für Untervektorräume) Seien U_1 und U_2 endlichdimensionale Untervektorräume des Vektorraumes V. Dann gilt*

$$\dim U_1 + \dim U_2 = \dim(U_1 + U_2) + \dim(U_1 \cap U_2).$$

Beweis. Sei $C = \{c_1, \ldots, c_k\}$ eine Basis von $U_1 \cap U_2$. Wir setzen C zu Basen von U_1 bzw. U_2 fort: $B_1 = \{c_1, \ldots, c_k, a_{k+1}, \ldots, a_i\}$ und $B_2 = \{c_1, \ldots, c_k, b_{k+1}, \ldots, b_j\}$. Wir zeigen, daß dann $B := \{c_1, \ldots, c_k, a_{k+1}, \ldots, a_i, b_{k+1}, \ldots, b_j\}$ eine Basis für $U_1 + U_2$ ist. Zunächst ist B eine Erzeugendenmenge, denn alle Vektoren aus U_1 bzw. U_2 lassen sich als Linearkombinationen darstellen, also auch alle Vektoren aus $U_1 + U_2$. Außerdem ist B linear unabhängig, denn wenn $\sum_{r=1}^k \gamma_r c_r + \sum_{s=1}^i \alpha_s a_s + \sum_{t=1}^j \beta_t b_t = 0$ ist, dann ist $v := \sum_{r=1}^k \gamma_r c_r + \sum_{s=1}^i \alpha_s a_s = -\sum_{t=1}^j \beta_t b_t \in U_1 \cap U_2$, also sind alle $\alpha_s = 0$, weil C Basis von $U_1 \cap U_2$ ist. Ebenso sieht man, daß alle $\beta_t = 0$. Dann ist aber in der ursprünglichen Summe $\sum_{r=1}^k \gamma_r c_r = 0$. Damit sind auch die $\gamma_r = 0$. Das zeigt, daß B eine Basis von $U_1 + U_2$ ist. Jetzt gilt $\dim(U_1 + U_2) = k + (i - k) + (j - k) = \dim(U_1 \cap U_2) + \dim U_1 - \dim(U_1 \cap U_2) + \dim U_2 - \dim(U_1 \cap U_2)$. Daraus folgt die behauptete Dimensionsformel.

Beispiele 5.3.9 1. Sei I eine Menge mit mindestens zwei Elementen. Die $U_i := \langle e_i \rangle = K e_i$ sind Unterräume von $K^{(I)}$. Da die Menge $\{e_i\}$ eine Basis von $K^{(I)}$ bildet, sieht man sofort, daß gilt $K^{(I)} = \bigoplus_{i \in I} K e_i$.

2. Wir schließen diesen Abschnitt mit dem uns schon bekannten Beispiel 5.1.2 5. des Vektorraumes

$$U := \{(\xi_1, \xi_2, \xi_3, \xi_4) | \xi_1, \xi_2, \xi_3, \xi_4 \in \mathbb{R}, \xi_1 + \xi_2 + \xi_3 + \xi_4 = 0\}.$$

Er ist ein Untervektorraum des Vektorraumes

$$V := \{(\xi_1, \xi_2, \xi_3, \xi_4) | \xi_1, \xi_2, \xi_3, \xi_4 \in \mathbb{R}\},$$

denn wenn zwei Vektoren $(\xi_1, \xi_2, \xi_3, \xi_4), (\eta_1, \eta_2, \eta_3, \eta_4) \in U$ gegeben sind, wenn also $\xi_1+\xi_2+\xi_3+\xi_4 = 0$ und $\eta_1+\eta_2+\eta_3+\eta_4 = 0$ gelten und wenn $\alpha \in K$ gegeben ist, so gelten auch $(\xi_1+\eta_1)+(\xi_2+\eta_2)+(\xi_3+\eta_3)+(\xi_4+\eta_4) = 0$ und $\alpha\xi_1 + \alpha\xi_2 + \alpha\xi_3 + \alpha\xi_4 = 0$, d.h. $(\xi_1, \xi_2, \xi_3, \xi_4) + (\eta_1, \eta_2, \eta_3, \eta_4) \in U$ und $\alpha(\xi_1, \xi_2, \xi_3, \xi_4) \in U$. U hat nach 5.2.4 5. eine Basis $(-1, 1, 0, 0)$, $(-1, 0, 1, 0)$, $(-1, 0, 0, 1)$, hat also die Dimension 3. Jeder Vektor $v \in U$ läßt sich eindeutig in der Form

$$v = \alpha_1(-1, 1, 0, 0) + \alpha_2(-1, 0, 1, 0) + \alpha_3(-1, 0, 0, 1)$$

darstellen, kann also allein durch die Angabe des Tripels $(\alpha_1, \alpha_2, \alpha_3)$ eindeutig beschrieben werden. Bei festgehaltener Basis des Vektorraumes U können damit dessen Elemente allein durch die Angabe ihrer Koeffiziententripel angegeben werden. Die Summe zweier Vektoren und das Produkt mit einem Skalar ergibt für die Koeffiziententripel komponentenweise Addition bzw. Multiplikation. So gilt z.B. für die Vektoren bzw. deren Summe:

$$(-6, 1, 2, 3) = 1(-1, 1, 0, 0) + 2(-1, 0, 1, 0) + 3(-1, 0, 0, 1)$$

mit dem Koeffiziententripel $(1, 2, 3)$ und

$$(-4, 2, 0, 2) = 2(-1, 1, 0, 0) + 0(-1, 0, 1, 0) + 2(-1, 0, 0, 1)$$

mit dem Koeffiziententripel $(2, 0, 2)$ ergibt

$$\begin{aligned}(-6, 1, 2, 3) + (-4, 2, 0, 2) = (-10, 3, 2, 5) = \\ = 3(-1, 1, 0, 0) + 2(-1, 0, 1, 0) + 5(-1, 0, 0, 1)\end{aligned}$$

mit $(1, 2, 3) + (2, 0, 2) = (3, 2, 5)$.

Ein direktes Komplement zu U in $\mathbb{R}^4$ können wir finden, wenn wir die Basis von U zu einer Basis von $\mathbb{R}^4$ vervollständigen. Das kann auf sehr vielfältige Weise geschehen. Insbesondere sind die Vektoren

$$(-1, 1, 0, 0), (-1, 0, 1, 0), (-1, 0, 0, 1), e_i$$

für jede Wahl von $i = 1, 2, 3, 4$ eine Basis von $\mathbb{R}^4$, also ist jeder der Untervektorräume $\mathbb{R}e_i$ ein direktes Komplement von U. Wir empfehlen dem Leser, noch weitere direkte Komplemente von U zu suchen.

Übungen 5.3.10 1. Seien V und W zwei Vektorräume über dem Körper K.

a) Zeigen Sie, daß das (kartesische) Produkt $V \times W$ mit der Addition

$$(v, w) + (v', w') := (v + v', w + w')$$

und der Multiplikation mit Skalaren

$$\lambda(v, w) := (\lambda v, \lambda w)$$

ein Vektorraum ist.

b) Zeigen Sie: Hat V eine Basis aus n Vektoren und W eine Basis aus m Vektoren, so hat $V \times W$ eine Basis aus $n + m$ Vektoren.

2. Zeigen Sie
   a) $\mathrm{Span}(1 + x, 1 - x) + \mathrm{Span}(x^2 - x, x^2 - 1) = \mathrm{Span}(1, x, x^2)$.
   b) Ist die Summe eine direkte Summe?
   c) Berechnen Sie die Dimension des Durchschnitts.
3. Sei P der Untervektorraum von $\mathbb{R}^{\mathbb{R}}$, der von den Funktionen $\{1, x, x^2, \ldots\}$ aufgespannt wird. Wir nennen P den Raum der Polynomfunktionen. Zeigen Sie, daß P die direkte Summe der Räume $P_g := \langle 1, x^2, x^4, x^6, \ldots \rangle$ und $P_u := \langle x, x^3, x^5, \ldots \rangle$ der Polynome mit geraden bzw. ungeraden Exponenten ist.
4. Zeigen Sie, daß $\mathbb{R}_5 = U \oplus V$, wobei

$$U := \{x \in \mathbb{R}_5 \left| \begin{pmatrix} 1 & 2 & 1 & 0 & 1 \\ 2 & 3 & 1 & 1 & 1 \end{pmatrix} x = 0 \right.\}$$

und

$$V := \{x \in \mathbb{R}_5 | x_1 = 2t, x_2 = s + t, x_3 = 3s + t, x_4 = s, x_5 = 2t, s, t \in \mathbb{R}\}.$$

5. Seien $V = \{(\,\lambda, \quad 2\lambda, \quad 2\mu, \quad \mu\,) | \lambda, \mu \in \mathbb{R}\}$ und $W = \{(\,\lambda, \quad \lambda, \quad \lambda, \quad \lambda\,) | \lambda \in \mathbb{R}\}$. Bestimmen Sie Unterräume $U_1, U_2, U_3 \subset \mathbb{R}^4$, so daß $U_i \oplus V = \mathbb{R}^4$ für $i = 1, 2, 3$ gilt, $U_1 \cap U_2 = 0$ und $U_1 \cap U_3 = W$.

5.4 Lineare Abbildungen

Wie bei der Diskussion der algebraischen Grundstrukturen Halbgruppe, Monoid und Gruppe wollen wir jetzt auch für Vektorräume den Begriff eines Homomorphismus einführen.

Wenn wir irgendwelche Punkte, Geraden oder beliebige Punktmengen in einem Vektorraum haben, so wollen wir diese sinnvoll in dem vorgegebenen Vektorraum bewegen können oder sie sogar in einen anderen Vektorraum „hinüberbringen“ können, z.B. durch eine Projektion oder durch eine Streckung. Der dafür geeignete mathematische Begriff ist der des Homomorphismus von Vektorräumen oder der linearen Abbildung. Lineare Abbildungen zwischen endlichdimensionalen Vektorräumen können in sehr einfacher Weise mit Matrizen beschrieben werden. Wir werden diese Technik ausführlich in Abschnitt 5.5 darstellen.

Definition 5.4.1 Sei K ein beliebiger Körper. Seien V und W zwei Vektorräume über K. Sei schließlich $f : V \longrightarrow W$ eine Abbildung von V in W. f heißt eine *lineare Abbildung* oder ein *Homomorphismus*, wenn die folgenden Gesetze erfüllt sind:

$f(v_1 + v_2) = f(v_1) + f(v_2)$ für alle v_1, v_2 in V,
$f(\lambda v) = \lambda f(v)$ für alle $\lambda \in K$ und $v \in V$.

Wie bei anderen algebraischen Strukturen (vgl. 3.2.4 und 3.2.5) sprechen wir auch bei Homomorphismen von Vektorräumen $f : V \longrightarrow W$ von einem *Isomorphismus*, wenn f bijektiv ist (und damit einen Umkehrhomomorphismus besitzt). Zwei Vektorräume heißen *isomorph*, wenn es einen Isomorphismus zwischen ihnen gibt.
Ein injektiver Homomorphismus von Vektorräumen $f : V \longrightarrow W$ heißt *Monomorphismus*.
Ein surjektiver Homomorphismus von Vektorräumen $f : V \longrightarrow W$ heißt *Epimorphismus*.
Ein Homomorphismus von Vektorräumen $f : V \longrightarrow V$ mit $V = \mathrm{Qu}(f) = \mathrm{Zi}(f)$ heißt *Endomorphismus*.
Ein Isomorphismus von Vektorräumen $f : V \longrightarrow V$ mit $V = \mathrm{Qu}(f) = \mathrm{Zi}(f)$ heißt *Automorphismus*.

Wegen der ersten Bedingung ist f auch ein Homomorphismus von abelschen Gruppen. Insbesondere erhält man $f(0) = 0$. Offensichtlich läßt sich diese Definition auch verwenden, um allgemein zu zeigen

$$f(\sum_{i=1}^{n} \lambda_i v_i) = \sum_{i=1}^{n} \lambda_i f(v_i).$$

Damit werden alle wesentlichen Operationen in einem Vektorraum von linearen Abbildungen respektiert. Es gelten im übrigen sinngemäß wieder die Ausführungen von Kapitel 3, insbesondere

Lemma 5.4.2 *Seien $f : V \longrightarrow W$ und $g : W \longrightarrow Z$ zwei lineare Abbildungen. Dann ist auch $gf : V \longrightarrow Z$ eine lineare Abbildung. Weiterhin ist die identische Abbildung $\mathrm{id}_V : V \longrightarrow V$ eine lineare Abbildung.*

Beweis. Wir brauchen die beiden Bedingungen nur auszuschreiben:

$$gf(v_1 + v_2) = g(f(v_1) + f(v_2)) = gf(v_1) + gf(v_2)$$

und

$$gf(\lambda v) = g(\lambda f(v)) = \lambda gf(v).$$

Für die identische Abbildung ist die Aussage noch trivialer.

Lineare Abbildungen oder Homomorphismen haben oft eine weitreichende geometrische Bedeutung, wie die folgenden Beispiele zeigen.

Beispiele 5.4.3 1. Wir wollen ein erstes Beispiel für eine lineare Abbildung angeben. Dazu verwenden wir die Vektorräume $V = \mathbb{R}^3$ und $W = \mathbb{R}^2$. Die Abbildung $f : V \longrightarrow W$ sei definiert durch

$$f(\xi, \eta, \zeta) := (\xi, \zeta),$$

also die Projektion des dreidimensionalen Raumes auf die x-z-Ebene (entlang der y-Achse). Wir rechnen schnell nach, daß

$$\begin{aligned} f((\xi_1,\eta_1,\zeta_1) + (\xi_2,\eta_2,\zeta_2)) &= f(\xi_1+\xi_2, \eta_1+\eta_2, \zeta_1+\zeta_2) \\ &= (\xi_1+\xi_2, \zeta_1+\zeta_2) = (\xi_1,\zeta_1) + (\xi_2,\zeta_2) \\ &= f(\xi_1,\eta_1,\zeta_1) + f(\xi_2,\eta_2,\zeta_2) \end{aligned}$$

und auch

$$\begin{aligned} f(\lambda(\xi,\eta,\zeta)) &= f(\lambda\xi, \lambda\eta, \lambda\zeta) = \\ &= (\lambda\xi, \lambda\zeta) = \lambda(\xi,\zeta) = \lambda f(\xi,\eta,\zeta) \end{aligned}$$

gelten. Damit ist f eine lineare Abbildung. Entsprechend können wir die Projektionen auf die x-y-Ebene bzw. auf die y-z-Ebene als lineare Abbildungen auffassen. Ja sogar die Projektionen auf die einzelnen Achsen, die x-Achse, die y-Achse und die z-Achse, sind lineare Abbildungen.

2. Die Abbildung

$$f : \mathbb{R}^2 \ni (\xi,\eta) \mapsto (\frac{1}{2}\sqrt{2}\xi + \frac{1}{2}\sqrt{2}\eta, -\frac{1}{2}\sqrt{2}\xi + \frac{1}{2}\sqrt{2}\eta) \in \mathbb{R}^2$$

stellt die Drehung der reellen Ebene um den Nullpunkt und den Winkel von 45^o dar (im mathematischen Drehsinn – der technische Drehsinn wird rechtsdrehend oder im Uhrzeigersinn gerechnet). Zur Abkürzung sei $a := \frac{1}{2}\sqrt{2}$. Dann gilt

$$\begin{aligned} f((\xi_1,\eta_1) + (\xi_2,\eta_2)) &= f(\xi_1+\xi_2, \eta_1+\eta_2) \\ &= (a(\xi_1+\xi_2) + a(\eta_1+\eta_2), -a(\xi_1+\xi_2) + a(\eta_1+\eta_2)) \\ &= (a\xi_1 + a\eta_1, -a\xi_1 + a\eta_1) + (a\xi_2 + a\eta_2, -a\xi_2 + a\eta_2) \\ &= f(\xi_1,\eta_1) + f(\xi_2,\eta_2), \end{aligned}$$

$$\begin{aligned} f(\lambda(\xi,\eta)) &= f(\lambda\xi, \lambda\eta) = (a\lambda\xi + a\lambda\eta, -a\lambda\xi + a\lambda\eta) = \\ &= \lambda(a\xi + a\eta, -a\xi + a\eta) = \lambda f(\xi,\eta). \end{aligned}$$

Damit ist f ein Homomorphismus von Vektorräumen.

Man kann eine lineare Abbildung $f : V \longrightarrow W$ auf einen Untervektorraum $U \subset V$ einschränken und erhält so eine lineare Abbildung $f|_U : U \longrightarrow W$, denn die Verträglichkeit mit der Addition von Vektoren und der Multiplikation mit Skalaren bleibt natürlich erhalten. Die Frage, ob auch eine Einschränkung in der Bildmenge W möglich ist, beantwortet das folgende Lemma.

Lemma 5.4.4 *Sei $f : V \longrightarrow W$ eine lineare Abbildung und $U \subset V$ ein Untervektorraum. Dann ist das Bild $f(U) \subset W$ von U ein Untervektorraum von W. Den Vektorraum $f(V)$ nennt man allgemein das* Bild von f*, oder einfach* $\mathrm{Bi}(f)$.

Beweis. Seien $f(u), f(u') \in f(U)$ mit $u, u' \in U$ gegeben, und sei $\lambda \in K$. Dann gelten $f(u) + f(u') = f(u + u') \in f(U)$ und $\lambda f(u) = f(\lambda u) \in f(U)$, weil U ein Untervektorraum ist, also $u + u' \in U$ und $\lambda u \in U$ gelten. Damit ist $f(U)$ ein Untervektorraum.

Wir können also auch den Zielvektorraum W einer linearen Abbildung $f : V \longrightarrow W$ einschränken, jedoch nicht beliebig wie das beim Quellvektorraum V der Fall war, sondern nur auf einen Untervektorraum, der den Untervektorraum $f(V)$ enthält. Schränken wir V weiter ein, so ist auch eine weitere Einschränkung von W möglich.

Andere wichtige Eigenschaften von Vektorräumen werden bei linearen Abbildungen nicht erhalten, wie etwa die Dimension oder die Eigenschaft einer Menge von Vektoren, Basis zu sein. Hier gelten kompliziertere Zusammenhänge. Besonders die Basiseigenschaft spielt eine ausgezeichnete Rolle für lineare Abbildungen. Sie hängt nämlich eng mit dem Begriff des freien Vektorraumes zusammen, wie er in 3.3.1 für die algebraischen Strukturen Halbgruppe, Monoid bzw. Gruppe schon definiert worden ist. Wir bemerken zunächst, daß 3.3.3 auch im Falle von Vektorräumen gilt:

Satz 5.4.5 *1. Ist V mit $\iota : B \longrightarrow V$ ein freier Vektorraum über der Menge B, so ist ι injektiv.*

2. Sind V und V' mit $\iota : B \longrightarrow V$ und $\iota' : B \longrightarrow V'$ freie Vektorräume, so gibt genau einen Homomorphismus $f : V \longrightarrow V'$ mit

$$\begin{tikzcd} B \arrow[r, "\iota"] \arrow[dr, "\iota'"'] & V \arrow[d, "f"] \\ & V' \end{tikzcd}$$

kommutativ (d.h. $f\iota = \iota'$), und f ist ein Isomorphismus.

Beweis. Der Beweis von 3.3.3 kann wörtlich übernommen werden, wenn man die Gruppe $\{1, -1\}$ durch den eindimensionalen Vektorraum K ersetzt.

Der Zusammenhang des Begriffes der Basis und eines freien Vektorraumes wird durch den folgenden Satz hergestellt.

Satz 5.4.6 *1. Seien V ein Vektorraum und B eine Basis von V. Dann ist V zusammen mit der Inklusionsabbildung $\iota : B \longrightarrow V$ ein (äußerer) freier Vektorraum über der Menge B (im Sinne von 3.3.1).*

2. Seien B eine Menge, V ein Vektorraum und $\iota : B \longrightarrow V$ eine Abbildung, so daß (V, ι) ein freier Vektorraum über B ist. Dann ist das Bild $\iota(B) \subset V$ eine Basis von V.

3. Sei B eine Menge. Dann gibt es einen freien Vektorraum (V, ι) über B.

Beweis. 1. Gegeben seien ein Vektorraum W und eine Abbildung $\alpha : B \longrightarrow W$. Zu zeigen ist, daß es genau einen Homomorphismus $f : V \longrightarrow W$ gibt mit $f\iota = \alpha$:

$$\begin{tikzcd} B \arrow[r, "\iota"] \arrow[dr, "\alpha"'] & V \arrow[d, "f"] \\ & W. \end{tikzcd}$$

Wir zeigen zunächst die Eindeutigkeit, weil dann schon ersichtlich wird, wie die Existenz zu zeigen ist. Seien f und g lineare Abbildungen von V nach W mit $\alpha(b_i) = f(b_i) = g(b_i)$ für alle i. Sei $v \in V$ beliebig gewählt. Dann läßt sich v darstellen als $v = \sum_{i=1}^n \lambda_i b_i$. Da f und g linear sind, gilt

$$\begin{aligned} f(v) &= f(\textstyle\sum_{i=1}^n \lambda_i b_i) = \sum_{i=1}^n \lambda_i f(b_i) \\ &= \textstyle\sum_{i=1}^n \lambda_i g(b_i) = g(\sum_{i=1}^n \lambda_i b_i) = g(v). \end{aligned}$$

Also gilt $f = g$.

Ist nur α gegeben und $v = \sum_{i=1}^n \lambda_i b_i \in V$, so konstruieren wir eine lineare Abbildung f durch

$$f(v) := \sum_{i=1}^n \lambda_i \alpha(b_i).$$

Da die Basisdarstellung von v eindeutig ist, d.h. die Koeffizienten λ_i durch v eindeutig bestimmt sind, ist mit dieser Definition eine Abbildung $f : V \longrightarrow W$ gegeben. Durch eine leichte Rechnung zeigt man jetzt $f(v+v') = f(v)+f(v')$ und $f(\lambda v) = \lambda f(v)$. Damit ist f eine lineare Abbildung und erfüllt offenbar $\alpha(b_i) = f(b_i) = f\iota(b_i)$ für alle i, also $\alpha = f\iota$.

2. Wir bezeichnen die Elemente von B mit $b_i \in B$, $i \in I$ und die Bildelemente mit $b'_i = \iota(b_i) \in \iota(B) = B'$. Die Menge B' ist linear unabhängig. Sei nämlich $\sum_i \lambda_i b'_i = 0$. Wir definieren für $i \in I$ eine Abbildung $\alpha_i : B \longrightarrow K$ mit $\alpha_i(b_j) := \delta_{ij}$, das Kronecker Symbol. Der induzierte Homomorphismus $f_i : V \longrightarrow K$, der existiert, weil (V, ι) frei ist, hat also die Eigenschaft $f_i(b'_j) = f_i\iota(b_j) = \alpha_i(b_j) := \delta_{ij}$. Damit folgt $0 = f_i(0) = f_i(\sum_j \lambda_j b'_j) = \sum_j \lambda_j f_i(b'_j) = \lambda_i$, alle Koeffizienten sind also Null, die Menge B' ist linear unabhängig. Wir ergänzen nun B' durch eine Menge $C \subset V$ zu einer Basis von V. Sei U der von C erzeugte Untervektorraum. Wir definieren einen Homomorphismus $p : V \longrightarrow U$ durch $p(b'_i) := 0$ und $p(c) = c$ für alle $c \in C$. Der Homomorphismus p ist nach Teil 1. wohldefiniert und eindeutig bestimmt. Er macht das Diagramm

$$\begin{tikzcd} B \arrow[r, "\iota"] \arrow[dr, "\alpha"'] & V \arrow[d, "p"] \\ & U \end{tikzcd}$$

mit $\alpha(b_i) = 0$ für alle $i \in I$ kommutativ. Dann muß aber p die Nullabbildung sein (wiederum nach Teil 1.), und es gilt $p(c) = 0$ für alle $c \in C$. Da C Teil einer Basis ist, kann das für keinen Vektor aus C gelten. Also ist $C = \emptyset$ und damit B' eine Basis von V.

3. Wir definieren $V := K^{(B)}$ und $\iota : B \longrightarrow V$ durch $\iota(b) := e_b$ (vgl. 5.2.4 4.). Sei W ein Vektorraum und $\alpha : B \longrightarrow W$ eine Abbildung. Dann definieren wir $f : V \longrightarrow W$ durch $f(\xi_b) := \sum_{b\in B} \xi_b \alpha(b)$. Man rechnet sofort nach, daß f ein Homomorphismus ist. Weiter gilt $f\iota(b) = f(e_b) = \sum_{b'} \delta_{bb'} \alpha(b') = \alpha(b)$, also $f\iota = \alpha$. Gilt auch $g\iota = \alpha$ für einen Homomorphismus $g : V \longrightarrow W$, so ist $g(\xi_b) = g(\sum_{b\in B} \xi_b e_b) = \sum_{b\in B} \xi_b g(e_b) = \sum_{b\in B} \xi_b g\iota(b) = \sum_{b\in B} \xi_b \alpha(b) = f(\xi_b)$, also ist $f = g$.

Die in 1. bewiesene Eigenschaft einer Basis in bezug auf lineare Abbildungen ist äußerst wichtig. Sie besagt zunächst, daß eine lineare Abbildung nur auf einer Basis vorgeschrieben werden muß. Darauf kann sie zudem noch beliebig gewählt werden. Dann gibt es eine eindeutig bestimmte lineare Fortsetzung.

Andrerseits bedeutet der Satz aber auch, daß zwei lineare Abbildungen schon gleich sind, wenn sie nur auf einer Basis übereinstimmen. Wir haben damit eine leichte Methode, um für beliebige lineare Abbildungen feststellen zu können, ob sie gleich sind.

Teil 1. dieses Satzes zusammen mit der Tatsache, daß jeder Vektorraum eine Basis besitzt (5.2.11 1.), sind der eigentliche Grund dafür, daß die Theorie der Vektorräume einfacher ist als andere algebraische Theorien. Damit ist jeder Vektorraum frei. Im allgemeinen ist es aber nicht wahr, daß jedes Modell für eine algebraische Struktur (also jede Gruppe, jede Halbgruppe oder jedes Monoid) frei ist.

Im Zusammenhang mit einer linearen Abbildung kommt einem bestimmten Untervektorraum eine besondere Bedeutung zu. Er wird im folgenden Lemma eingeführt.

Lemma 5.4.7 *Sei $f : V \longrightarrow W$ eine lineare Abbildung. Dann ist die Menge*

$$\mathrm{Ke}(f) := \{v \in V | f(v) = 0\}$$

ein Untervektorraum von V, der sogenannte Kern *von f.*

Beweis. Da offensichtlich $0 \in \mathrm{Ke}(f)$ wegen $f(0) = 0$, genügt es zu zeigen, daß mit $v, v' \in \mathrm{Ke}(f)$ und $\lambda \in K$ auch $v + v', \lambda v \in \mathrm{Ke}(f)$ gilt. Aus $f(v) = f(v') = 0$ folgt aber $f(v + v') = 0$ und $f(\lambda v) = \lambda f(v) = 0$.

Lemma 5.4.8 *Sei $f : V \longrightarrow W$ eine lineare Abbildung. f ist genau dann injektiv, wenn* $\mathrm{Ke}(f) = 0$ *gilt.*

Beweis. Ist f injektiv, so ist $0 \in V$ offenbar der einzige Vektor, der auf $0 \in W$ abgebildet wird, also ist $\mathrm{Ke}(f) = 0$. Ist umgekehrt $\mathrm{Ke}(f) = 0$ und $f(v) = f(v')$, so folgt $0 = f(v) - f(v') = f(v - v')$, also $v - v' \in \mathrm{Ke}(f) = 0$. Damit ist aber $v = v'$ und f injektiv.

Der Kern der linearen Abbildung $f : V \longrightarrow W$ ist ein Spezialfall des schon bekannten Begriffes des Urbilds eines Vektors $w \in W$ oder sogar des Urbilds einer Menge $M \subset W$ von Vektoren.

$$f^{-1}(w) := \{v \in V | f(v) = w\},$$

$$f^{-1}(M) := \{v \in V | f(v) \in M\}.$$

Lemma 5.4.9 *Ist $f : V \longrightarrow W$ eine lineare Abbildung und ist $v \in f^{-1}(w)$, insbesondere also $f^{-1}(w)$ nicht leer, so gilt*

$$f^{-1}(w) = v + \mathrm{Ke}(f) = \{v + v' | v' \in \mathrm{Ke}(f)\}.$$

Beweis. Sei $v' \in \mathrm{Ke}(f)$. Dann gilt $f(v + v') = f(v) + f(v') = w + 0 = w$, also ist $v + v' \in f^{-1}(w)$. Ist nun umgekehrt $v'' \in f^{-1}(w)$, so gilt $f(v'' - v) = f(v'') - f(v) = w - w = 0$, also ist $v' := v'' - v \in \mathrm{Ke}(f)$. Damit erhält man aber $v'' = v + (v'' - v) = v + v' \in v + \mathrm{Ke}(f)$. Das war zu zeigen.

Wir haben also insbesondere gesehen, daß $\mathrm{Ke}(f) = f^{-1}(0)$ gilt. Weiterhin sieht man, daß der Vektor 0 nur in $f^{-1}(0)$ liegt und in keiner der anderen Mengen $f^{-1}(w)$, mit $w \neq 0$, denn es gilt immer $f(0) = 0$.

Über die Berechnung des Kerns einer linearen Abbildung werden wir in Kapitel 6 mehr erfahren. Hier können wir jedoch schon eine fundierte Aussage über die Größe des Kerns oder genauer über seine Dimension machen. Es gilt nämlich

Satz 5.4.10 *(Dimensionssatz für Homomorphismen) Seien V ein n-dimensionaler Vektorraum und $f : V \longrightarrow W$ eine lineare Abbildung. Dann gilt*

$$\dim(\mathrm{Ke}(f)) + \dim(\mathrm{Bi}(f)) = \dim(V).$$

Beweis. Wir wählen für $\mathrm{Ke}(f)$ zunächst eine Basis $b_1, \ldots, b_k$. Diese ist eine linear unabhängige Menge in V und kann daher zu einer Basis $b_1, \ldots, b_k, b_{k+1}, \ldots, b_n$ fortgesetzt werden. Wir behaupten nun, daß die Vektoren $f(b_{k+1}), \ldots, f(b_n)$ alle paarweise verschieden sind und eine Basis von $\mathrm{Bi}(f)$ bilden. Ist nämlich $f(b_i) = f(b_j)$ mit $k \leq i, j$, so ist $f(b_i - b_j) = 0$, also $b_i - b_j \in \mathrm{Ke}(f)$. Damit gibt es eine Linearkombination $b_i - b_j = \sum_{r=1}^{k} \alpha_r b_r$. Wegen der linearen Unabhängigkeit der $b_1, \ldots, b_n$ sind damit die $\alpha_r = 0$, und es gilt $b_i = b_j$. Ist weiter $\sum_{r=k+1}^{n} \beta_r f(b_r) = 0$, also $f(\sum_{r=k+1}^{n} \beta_r b_r) = 0$, so gilt $\sum_{r=k+1}^{n} \beta_r b_r \in \mathrm{Ke}(f)$ und $\sum_{r=k+1}^{n} \beta_r b_r = \sum_{r=1}^{k} \alpha_r b_r$. Wegen der linearen Unabhängigkeit der $b_1, \ldots, b_n$ ergibt sich wieder $\beta_r = 0$, also sind die $f(b_{k+1}), \ldots, f(b_n)$ linear unabhängig. Um zu zeigen, daß sie eine Basis für $\mathrm{Bi}(f)$ bilden, sei $f(v)$ mit $v = \sum_{r=1}^n \beta_r b_r$ ein beliebiger Vektor in $\mathrm{Bi}(f)$. Dann gilt

$$f(v) = f(\sum_{r=1}^n \beta_r b_r) = \sum_{r=1}^n \beta_r f(b_r) = \sum_{r=k+1}^n \beta_r f(b_r),$$

weil die Vektoren $f(b_1) = \ldots = f(b_k) = 0$ sind. Damit ist gezeigt, daß die Vektoren $f(b_{k+1}), \ldots, f(b_n)$ eine Basis des Bildes $\mathrm{Bi}(f)$ ist. Die Dimension des Bildes ist also $\dim(\mathrm{Bi}(f)) = n-k = \dim(V) - \dim(\mathrm{Ke}(f))$, wie die Formel im Satz behauptet.

Definition 5.4.11 Die Dimension $\dim(\mathrm{Bi}(f))$ einer linearen Abbildung f heißt auch *Rang* der linearen Abbildung und wird mit $\mathrm{rg}(f)$ bezeichnet.

Folgerung 5.4.12 *Sei $f : V \longrightarrow W$ eine lineare Abbildung und seien V, W endlichdimensional. f ist genau dann bijektiv oder ein Isomorphismus, wenn* $\mathrm{Ke}(f) = 0$ *und* $\dim(V) = \dim(W)$ *gelten. In diesem Falle stimmen der Rang von f und die Dimension von V überein.*

Beweis. Sei zunächst f bijektiv. Dann ist nach Lemma 5.4.8 $\mathrm{Ke}(f) = 0$. Wegen 5.4.10 gilt dann $\dim(V) = \dim(\mathrm{Bi}(f))$. Da f surjektiv ist, ist $\mathrm{Bi}(f) = W$, also gilt auch $\dim(V) = \dim(W)$.

Um die Umkehrung zu zeigen, beachten wir zunächst, daß nach 5.4.8 f schon injektiv ist. Dann folgt aber $\dim(\mathrm{Bi}(f) = \dim(V) = \dim(W)$. Es gibt also in $\mathrm{Bi}(f)$ eine Basis von $n = \dim(W)$ Vektoren. Nach 5.4.10 ist diese auch eine Basis für W, also $\mathrm{Bi}(f) = W$. Damit ist f auch surjektiv, also bijektiv.

Die Aussage über den Rang folgt unmittelbar aus der Gleichung $\dim(V) = \dim(W)$.

Folgerung 5.4.13 *Sei $f : V \longrightarrow W$ eine lineare Abbildung und sei V endlichdimensional. Dann gilt*

$$\dim \mathrm{Ke}(f) + \mathrm{rg}(f) = \dim V.$$

Beweis. folgt unmittelbar aus dem Dimensionssatz 5.4.10.

Folgerung 5.4.14 *Sei $f : V \longrightarrow W$ eine lineare Abbildung und sei* $\dim V = \dim W$ *endlich. Dann sind äquivalent*

1. f ist ein Isomorphismus.
2. f ist ein Epimorphismus.
3. f ist ein Monomorphismus.

Beweis. Die Äquivalenz von 1. und 3. folgt unmittelbar aus 5.4.12. Außerdem folgt 2. aus 1. Ist 2. gegeben, so folgt nach dem Dimensionssatz, daß $\dim \mathrm{Ke}(f) = 0$, also auch $\mathrm{Ke}(f) = 0$, gilt und daher mit 5.4.8 die Behauptung 3.

Wir betrachten jetzt noch den Zusammenhang zwischen Homomorphismen und direkten Summen.

Satz 5.4.15 *1. Sei $V = \bigoplus_{i \in I} U_i$ eine (innere) direkte Summe von Untervektorräumen. Seien W ein Vektorraum und $(f_i : U_i \longrightarrow W | i \in I)$ eine Familie von Homomorphismen. Dann gibt es genau einen Homomorphismus $f : V \longrightarrow W$, so daß die Diagramme*

$$\begin{tikzcd} U_i \arrow[r, "\iota_i"] \arrow[dr, "f_i"'] & V \arrow[d, "f"] \\ & W \end{tikzcd}$$

für alle $i \in I$ kommutieren.

2. Seien U_i, $i \in I$, Vektorräume. Dann gibt es einen Vektorraum V und eine Familie von Homomorphismen $(j_i : U_i \longrightarrow V)$, so daß für jeden Vektorraum W und jede Familie von Homomorphismen $(f_i : U_i \longrightarrow W)$ genau ein Homomorphismus $f : V \longrightarrow W$ existiert, so daß die Diagramme

$$\begin{tikzcd} U_i \arrow[r, "j_i"] \arrow[dr, "f_i"'] & V \arrow[d, "f"] \\ & W \end{tikzcd}$$

für alle $i \in I$ kommutieren.
In diesem Falle sind die Homomorphismen $j_i : U_i \longrightarrow V$ injektiv. Weiter ist V (innere) direkte Summe der Bilder $j_i(U_i)$. V zusammen mit der Familie $(j_i : U_i \longrightarrow V | i \in I)$ heißt auch (äußere) direkte Summe der Vektorräume U_i.

Beweis. 1. Wie schon bei früheren Beweisen zeigen wir zunächst die Eindeutigkeit und dann die Existenz des Homomorphismus f. Seien f und g mit $fj_i = gj_i = f_i$ (für alle i) gegeben. Für einen Vektor $v = \sum_{i \in I} u_i \in V = \bigoplus_{i \in I} U_i$ ist dann $f(v) = f(\sum_i u_i) = \sum_i f(u_i) = \sum_i fj_i(u_i) = \sum_i f_i(u_i) = \sum_i gj_i(u_i) = \sum_i g(u_i) = g(\sum_i u_i) = g(v)$. Also gilt $f = g$. Die Rechnung ist mit unserer Definition der Summe (vgl. Bemerkung vor 5.1.10) $\sum_i u_i$ durchführbar, weil nur endlich viele Terme von Null verschieden sind.

Um jetzt die Existenz von f zu zeigen, definieren wir $f(v) := \sum_i f_i(u_i)$. Damit ist eine wohldefinierte Abbildung $f : V \longrightarrow W$ gegeben, weil die Darstellung $v = \sum_i u_i$ nach 5.3.4 eindeutig ist. Man rechnet sofort nach, daß f ein Homomorphismus ist und die Bedingung $fj_i = f_i$ für alle $i \in I$ erfüllt.

2. Die Konstruktion von V ist analog zur Konstruktion des n-fachen Produkts eines Vektorraumes V mit sich selbst (5.1.3 und 5.1.7). Wir definieren

$$V := \{(u_i|i \in I)|\forall i \in I[u_i \in U_i], u_i \neq 0 \text{ für nur endlich viele } i \in I\}.$$

Man rechnet sofort nach, daß V mit komponentenweiser Addition und Multiplikation mit Skalaren ein Vektorraum ist. Für $i \in I$ seien $j_i : U_i \longrightarrow V$ definiert durch $j_i(u_i) := (u'_j|j \in I)$ mit $u'_i = u_i$ und $u'_j = 0$ für $j \neq i$, also die Familie, die an der i-ten Stelle den Eintrag u_i hat und an allen anderen Stellen Null ist. Es ist auch unmittelbar klar, daß die j_i Homomorphismen sind.

Sei nun ein Vektorraum W und eine Familie von Homomorphismen $(f_i : U_i \longrightarrow W)$ gegeben. Wir müssen die Existenz und Eindeutigkeit eines Homomorphismus $f : V \longrightarrow W$ mit $fj_i = f_i$ für alle $i \in I$ zeigen. Wieder zeigen wir zunächst die Eindeutigkeit. Seien f und g Homomorphismen von V in W mit $fj_i = f_i = gj_i$. Dann ist $g((u_i)) = g(\sum_{i\in I} j_i(u_i)) = \sum_{i\in I} gj_i(u_i) = \sum_{i\in I} f_i(u_i) = \sum_{i\in I} fj_i(u_i) = f(\sum_{i\in I} j_i(u_i)) = f((u_i))$, also $f = g$. Um die Existenz zu zeigen, definieren wir $f((u_i)) := \sum_{i\in I} f_i(u_i)$. Man rechnet sofort nach, daß $f : V \longrightarrow W$ ein Homomorphismus ist. Weiter ist $fj_i(u_i) = f((u'_j)) = \sum_{j\in I} f_j(u'_j) = f_i(u_i)$, weil alle anderen Terme Null sind, also ist $fj_i = f_i$.

Wir zeigen nun die beiden letzten Aussagen von 2. Erfülle V und $(j_i : U_i \longrightarrow V)$ die Bedingungen von Teil 2. des Satzes. (Sie müssen nicht genauso konstruiert sein, wie wir das oben angegeben haben.) Sei $i \in I$ fest gewählt. Wähle Homomorphismen $f_j : U_j \longrightarrow U_i$ für alle $j \in I$ mit $f_i = \mathrm{id}_{U_i}$ und $f_j = 0$ für alle $j \neq i$. Dann können wir nach 2. den Homomorphismus $f : V \longrightarrow U_i$ bestimmen und erhalten $fj_i = \mathrm{id}_{U_i}$. Daher ist j_i injektiv.

Wir zeigen jetzt $V = \bigoplus_{i\in I} j_i(U_i)$. Sei $W := \sum_{i\in I} j_i(U_i)$ und seien $(f_i : U_i \longrightarrow W) = (j_i : U_i \longrightarrow W)$ für alle $i \in I$. Dann gibt es genau einen Homomorphismus $f : V \longrightarrow W$ mit $fj_i = j_i$. Sei $\iota : W \longrightarrow V$ die Einbettungsabbildung. Dann ist $\iota j_i = j_i$ und $\iota fj_i = j_i$. Wegen der Eindeutigkeit folgt $\iota f = \mathrm{id}_V$. Damit ist ι surjektiv (und nach Konstruktion injektiv), also ist $W = V$.

Sei $x \in U_i$ mit $j_i(x) = x' \in j_i(U_i) \cap \sum_{j\in I, j\neq i} j_j(U_j)$. Wir definieren wie oben $W := U_i$ und $f_j := \mathrm{id}_{U_i}$ für $j = i$ und $f_j = 0$ sonst. Wieder erhalten wir einen eindeutig bestimmten Homomorphismus $f : V \longrightarrow U_i$ mit $fj_j = f_j$. Wegen $x' = \sum_{j\neq i} j_j(u_j)$ ist $x = fj_i(x) = f(x') = \sum_{j\neq i} fj_j(u_j) = 0$, also $x' = 0$ und damit $j_i(U_i) \cap \sum_{j\in I, j\neq i} j_j(U_j) = 0$.

Beispiel 5.4.16 Der Begriff des Homomorphismus (von Vektorräumen) führt uns zu einem neuen Vektorraum. Wenn V und W Vektorräume sind, so sei

$$\mathrm{Hom}_K(V, W) := \{f : V \longrightarrow W | f \text{ Homomorphismus}\}$$

die Menge aller Homomorphismen von V nach W. Dann ist $\mathrm{Hom}_K(V,W)$ ein Untervektorraum von $\mathrm{Abb}(V,W)$. Es ist nämlich $0 \in \mathrm{Hom}_K(V,W)$. Wenn $f,g \in \mathrm{Hom}_K(V,W)$, dann gelten $(f+g)(v+v') = f(v+v')+g(v+v') = f(v)+f(v')+g(v)+g(v') = f(v)+g(v)+f(v')+g(v') = (f+g)(v)+(f+g)(v')$ und $(f+g)(\lambda v) = f(\lambda v)+g(\lambda v) = \lambda f(v)+\lambda g(v) = \lambda(f(v)+g(v)) = \lambda(f+g)(v)$, also ist $f+g$ wieder ein Homomorphismus. In gleicher Weise zeigt man, daß mit f auch μf wieder ein Homomorphismus ist.

Der Vektorraum $V^* := \mathrm{Hom}_K(V,K)$ heißt *dualer Vektorraum.* Die Elemente von V^* heißen *lineare Funktionale.*

Wenn V ein endlichdimensionaler Vektorraum ist, dann haben V und V^* dieselbe Dimension und sind daher isomorph zueinander. Wenn nämlich $b_1,\ldots,b_n$ eine Basis von V ist, dann ist $f_1,\ldots,f_n$ mit $f_i(b_j) := \delta_{ij}$ eine Basis von V^*, wie man leicht nachrechnet.

Ist V ein Vektorraum mit einer Basis B, so ist die Abbildung

$$h_B : V \ni v = \sum_{b\in B} \alpha_b b \mapsto (\alpha_b | b \in B) \in K^{(B)}$$

ein bijektiver Homomorphismus, also ein Isomorphismus, wegen der eindeutigen Basisdarstellung jedes Vektors $v \in V$.

Definition 5.4.17 Sei B eine Basis des Vektorraumes V. Der Isomorphismus $h_B : V \longrightarrow K^{(B)}$ heißt *Koordinatensystem* für V zur Basis B.

Folgerung 5.4.18 *Für jeden Vektorraum V gibt es einen Isomorphismus $V \cong K^{(I)}$ für eine geeignet gewählte Menge I. Ist V endlichdimensional, so ist $V \cong K^n$, wobei $n = \dim V$.*

Beweis. Man rechnet sofort nach, daß h_B ein Homomorphismus ist. Es ist auch klar, daß h_B bijektiv ist. Ist $B = \{b_1,\ldots,b_n\}$ eine endliche Menge, so schreiben wir h_B auch in der Form

$$h_B : V \ni v = \sum_{i=1}^{n} \alpha_i b_i \mapsto (\alpha_i | i = 1,\ldots,n) \in K_n.$$

Definition 5.4.19 Eine lineare Abbildung $p : V \longrightarrow V$ mit $p^2 = p$ wird eine *Projektion* oder *idempotent* genannt.

Satz 5.4.20 *Sei $p : V \longrightarrow V$ eine Projektion. Dann gilt*

$$V = \mathrm{Ke}(p) \oplus \mathrm{Bi}(p).$$

Beweis. $\mathrm{Ke}(p)$ und $\mathrm{Bi}(p)$ sind Untervektorräume von V. Für $v \in V$ gilt $v = (v-p(v))+p(v)$ und es ist $p(v-p(v)) = p(v)-p^2(v) = 0$, also $v-p(v) \in \mathrm{Ke}(p)$, und $p(v) \in \mathrm{Bi}(p)$. Damit ist $V = \mathrm{Ke}(p) + \mathrm{Bi}(p)$. Wenn $v \in \mathrm{Ke}(p) \cap \mathrm{Bi}(p)$, dann ist $v = p(w) = p^2(w) = p(v) = 0$, also ist $\mathrm{Ke}(p) \cap \mathrm{Bi}(p) = 0$.

Satz 5.4.21 *Sei $V = V' \oplus V''$. Dann gibt es genau eine Projektion $p : V \longrightarrow V$ mit $V' = \mathrm{Ke}(p)$ und $V'' = \mathrm{Bi}(p)$.*

Beweis. Sei $v \in V$ mit der eindeutigen Darstellung $v = v' + v''$ gegeben. Wir definieren $p(v) := v''$. Man sieht leicht, daß p eine lineare Abbildung und eine Projektion ist und daß $V' = \mathrm{Ke}(p)$ und $V'' = \mathrm{Bi}(p)$ gilt. Wenn $p' : V \longrightarrow V$ eine Projektion mit $V' = \mathrm{Ke}(p')$ und $V'' = \mathrm{Bi}(p')$ ist, dann ist $p'(v') = 0$ für alle $v' \in V'$. Für $v'' \in V''$ gilt $v'' = p(w'') = p^2(w'') = p(v'')$ für ein $w'' \in V$, also ist $p(v' + v'') = v'' = p(v' + v'')$.

Übungen 5.4.22 1. Sei K ein Körper. Definieren Sie

$$f : K^2 \longrightarrow K^2, (x, y) \mapsto (ax + by, cx + dy)$$

für $a, b, c, d \in K$.

a) Zeigen Sie, daß f eine lineare Abbildung ist.

b) Definieren Sie

$$g : K^2 \longrightarrow K^2, (x, y) \mapsto (dx - by, -cx + ay)$$

Zeigen Sie:

$$(f \circ g)(x, y) = (ad - bc)(x, y) = (g \circ f)(x, y)$$

c) Schließen Sie: f ist genau dann bijektiv, wenn $ad - bc \neq 0$ ist.

2. Definieren Sie für $\varphi \in \mathbb{R}$:

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, (x, y) \mapsto (x \cos\varphi - y \sin\varphi, x \sin\varphi + y \cos\varphi)$$

Zeichnen Sie die Bilder der kanonischen Basisvektoren in ein Koordinatensystem ein. Erklären Sie anhand dieses Bildes, daß die Abbildung f eine Drehung beschreibt. Bestimmen Sie den Drehsinn und den Drehwinkel. Bestimmen Sie mit Hilfe der vorhergehenden Aufgabe die Umkehrabbildung und erläutern Sie, daß auch die Umkehrabbildung eine Drehung ist. Bestimmen Sie Drehsinn und Drehwinkel der Umkehrabbildung.

3. Sei $f : V \longrightarrow W$ eine lineare Abbildung zwischen zwei Vektorräumen. Zeigen Sie: Ist $U \subset W$ ein Untervektorraum von W, so ist $f^{-1}(U)$ ein Untervektorraum von V.

4. Sei $f : V \longrightarrow W$ eine bijektive lineare Abbildung. Zeigen Sie, daß auch die Umkehrabbildung linear ist.

5. Seien V und W zwei Vektorräume und $v_1, \ldots, v_n$ Vektoren aus V, $w_1, \ldots, w_n$ Vektoren aus W.

a) Zeigen Sie: Sind $v_1, \ldots, v_n$ linear unabhängig, so gibt es mindestens eine lineare Abbildung $f : V \longrightarrow W$ mit $f(v_j) = w_j$ für alle $j = 1, \ldots, n$.

b) Zeigen Sie: Erzeugen $v_1, \ldots, v_n$ den Vektorraum V, so gibt es höchstens eine lineare Abbildung $f : V \longrightarrow W$ mit $f(v_j) = w_j$ für alle $j = 1, \ldots, n$.

6. Sei $f : V \longrightarrow W$ eine lineare Abbildung zwischen zwei Vektorräumen. Sei $v_1, \dots, v_n$ eine Basis von V.
   a) Zeigen Sie, daß f genau dann injektiv ist, wenn $f(v_1), \dots, f(v_n)$ linear unabhängig sind.
   b) Zeigen Sie, daß f genau dann surjektiv ist, wenn $f(v_1), \dots, f(v_n)$ ein Erzeugendensystem von W ist.
7. Seien V und W Vektorräume und $f : V \longrightarrow W$ ein Homomorphismus. Zeigen Sie:
   a) Äquivalent sind:
      i. f ist injektiv.
      ii. Für jeden Vektorraum U und alle Homomorphismen $s, t : U \longrightarrow V$ mit $fs = ft$ gilt $s = t$.
      iii. Es gibt einen Homomorphismus $g : W \longrightarrow V$ mit $gf = \mathrm{id}_V$.
   b) Äquivalent sind:
      i. f ist surjektiv.
      ii. Für jeden Vektorraum X und alle Homomorphismen $s, t : W \longrightarrow X$ mit $sf = tf$ gilt $s = t$.
      iii. Es gibt einen Homomorphismus $g : W \longrightarrow V$ mit $fg = \mathrm{id}_W$.
8. Sei V ein Vektorraum über K. U_1 und U_2 seien Untervektorräume von V mit $V = U_1 \oplus U_2$. Zeigen Sie, daß die Abbildung

$$f : U_1 \times U_2 \longrightarrow V, (u_1, u_2) \mapsto u_1 + u_2$$

   ein Isomorphismus ist, wobei $U_1 \times U_2$ die in der Aufgabe 1 erklärte Vektorraumstruktur trägt.
9. Definieren Sie

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, (x, y) \mapsto (x/3 + 2y/3, x/3 + 2y/3)$$

   a) Zeigen Sie, daß $f \circ f = f$ ist.
   b) Bestimmen Sie Kern und Bild von f. Zeigen Sie:

$$\mathbb{R}^2 = \mathrm{Ke}(f) \oplus \mathrm{Bi}(f)$$

   c) Zeichnen Sie $\mathrm{Ke}(f)$ und $\mathrm{Bi}(f)$ in ein Koordinatenkreuz ein. Beschreiben Sie anhand dieses Bildes die Abbildung f.
   d) Entscheiden Sie, ob f bijektiv ist.
10. Sei V ein endlichdimensionaler Vektorraum und $f : V \longrightarrow V$ eine lineare Abbildung. Zeigen Sie:
   a) Es gelten

$$\begin{array}{l} 0 \subset \mathrm{Ke}(f) \subset \mathrm{Ke}(f^2) \subset \mathrm{Ke}(f^3) \subset \dots \\ V \supset \mathrm{Bi}(f) \supset \mathrm{Bi}(f^2) \supset \mathrm{Bi}(f^3) \supset \dots \end{array}$$

   b) Wenn $\mathrm{Ke}(f^r) = \mathrm{Ke}(f^{r+1})$ gilt, dann gilt auch $\mathrm{Ke}(f^r) = \mathrm{Ke}(f^{r+i})$ und $\mathrm{Bi}(f^r) = \mathrm{Bi}(f^{r+i})$ für alle $i > 0$.

c) (*Fitting-Lemma:*)[3] Es gibt ein $n > 0$, so daß $V = \mathrm{Ke}(f^n) \oplus \mathrm{Bi}(f^n)$. (Hinweis: Man zeige, daß es ein $n = r$ wie in Teil b) gibt. Dann kann man $\mathrm{Ke}(f^n) \cap \mathrm{Bi}(f^n) = 0$ beweisen und mit Dimensionsargumenten die Aussage zeigen.)

d) (*Fitting-Zerlegung:*)f ist direkte Summe eines Automorphismus $f|_{\mathrm{Bi}(f^n)} : \mathrm{Bi}(f^n) \longrightarrow \mathrm{Bi}(f^n)$ und eines *nilpotenten* Endomorphismus (eines Endomorphismus g mit $g^r = 0$) $f|_{\mathrm{Ke}(f^n)} : \mathrm{Ke}(f^n) \longrightarrow \mathrm{Ke}(f^n)$.

11. Sei $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ die (eindeutig bestimmte) $\mathbb{R}$-lineare Abbildung mit

$$\begin{aligned} f(0,\ 1,\ 2,\ 3) &= (3,\ 5,\ 2)\\ f(1,\ 2,\ 3,\ 0) &= (0,\ 3,\ 3)\\ f(2,\ 3,\ 0,\ 1) &= (1,\ 1,\ 0)\\ f(3,\ 0,\ 1,\ 2) &= (2,\ 3,\ 1). \end{aligned}$$

Bestimmen Sie eine Basis von $\mathrm{Ke}(f)$.

12. Sei V ein Vektorraum und $p : V \longrightarrow V$ eine Projektion. Zeigen Sie, daß es genau eine Projektion $q : V \longrightarrow V$ gibt, so daß gilt:

$$pq = qp = 0 \qquad p + q = \mathrm{id}_V$$

Drücken Sie Kern und Bild von q durch Kern und Bild von p aus.

13. Sei V ein endlichdimensionaler Vektorraum. Es gibt genau dann einen Homomorphismus $f : V \longrightarrow V$ mit $\mathrm{Ke}(f) = \mathrm{Bi}(f)$, wenn die Dimension von V gerade ist.

14. Sei V ein Vektorraum und $v_1, \ldots, v_n$ eine Basis von V. Für $i = 1, \ldots, n$ sei $f_i : V \longrightarrow K$ diejenige lineare Abbildung, die auf der gegebenen Basis durch

$$f_i(v_j) = \delta_{ij}$$

bestimmt ist (vgl. Beispiel 5.4.16). Zeigen Sie:

a) Für alle $v \in V$ gilt: $v = \sum_{i=1}^n f_i(v)v_i$

b) Für alle $f \in V^* := \mathrm{Hom}(V, K)$ gilt: $f = \sum_{i=1}^n f(v_i)f_i$

15. Sei V ein endlichdimensionaler Vektorraum. Zeigen Sie, daß V^* ebenfalls endlichdimensional ist und dieselbe Dimension wie V hat.

16. Sei K ein Körper und $V = K^{(\mathbb{N})}$. Zeigen Sie, daß es Untervektorräume $U_1, U_2 \subset V$ gibt mit $V \cong U_1 \cong U_2$ und $V = U_1 \oplus U_2$.

5.5 Die darstellende Matrix

Besonders einfach lassen sich Homomorphismen $f : V \longrightarrow W$ darstellen, wenn die Vektorräume V und W in der Form K_n gegeben sind. Dazu definieren wir zunächst eine Multiplikation von Matrizen (zur Definition vgl. Abschnitt 1).

[3] Hans Fitting (1906-1938)

Definition 5.5.1 Seien M eine $m \times n$-Matrix und N eine $n \times r$-Matrix. Wir definieren das Produkt der beiden Matrizen M und N, genannt *Matrizenprodukt*, als eine $m \times r$-Matrix durch

$$\begin{aligned} M \cdot N &= \left(\sum_{j=1}^n \alpha_{ij} \cdot \beta_{jk} | i = 1, \ldots, m; k = 1, \ldots, r\right) \\ &= \begin{pmatrix} \sum_{j=1}^n \alpha_{1j} \cdot \beta_{j1} & \ldots & \sum_{j=1}^n \alpha_{1j} \cdot \beta_{jr} \\ \vdots & & \vdots \\ \sum_{j=1}^n \alpha_{mj} \cdot \beta_{j1} & \ldots & \sum_{j=1}^n \alpha_{mj} \cdot \beta_{jr} \end{pmatrix}. \end{aligned}$$

Man kann also zwei Matrizen genau dann miteinander multiplizieren, wenn die Anzahl der Spalten der ersten Matrix mit der Anzahl der Zeilen der zweiten Matrix übereinstimmt.

Die Kronecker Funktion (vgl. 5.2.4 4.) gibt Anlaß zur Definition der *Einheitsmatrix* $E_n = (\delta_{ij} | i, j \in \{1, \ldots, n\}) \in K_n^n$. Es ist also

$$E_n = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Einige Rechenregeln lassen sich leicht nachrechnen und werden hier nur angegeben. Bei Matrizenprodukten nehmen wir immer an, daß sich die Matrizen nach der vorherigen Bemerkung miteinander multiplizieren lassen. Es gelten im einzelnen:

$$\begin{aligned} &M \cdot (N \cdot P) = (M \cdot N) \cdot P, && \text{(M1)} \\ &M \cdot (N + P) = M \cdot N + M \cdot P, && \text{(M2)} \\ &(M + N) \cdot P = M \cdot P + N \cdot P, && \text{(M3)} \\ &M \cdot (\lambda \cdot N) = (\lambda \cdot M) \cdot N = \lambda \cdot (M \cdot N), && \text{(M4)} \\ &E_m \cdot M = M = M \cdot E_n \text{ für } M \in K_m^n. && \text{(M5)} \end{aligned}$$

Beispiel 5.5.2 Mit diesen Hilfsmitteln können wir jetzt das bekannte Beispiel der linearen Abbildung $f : V \longrightarrow W$ mit $f\begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} := \begin{pmatrix} \xi \\ \zeta \end{pmatrix}$ auch anders beschreiben. Für $\begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} \in \mathbb{R}_3$ gilt nämlich

$$f\begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} \xi \\ \zeta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}.$$

Wegen der oben angegebenen Rechenregeln (M3) und (M4) folgt direkt, daß die Multiplikation der Spaltenvektoren $\begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}$ von links mit der angegebenen Matrix eine lineare Abbildung ist. Allgemein halten wir den folgenden Satz fest.

Satz 5.5.3 *Sei M eine $m \times n$-Matrix. Die Multiplikation auf Spaltenvektoren aus K_n von links mit M ist eine lineare Abbildung $f : K_n \longrightarrow K_m$ mit* $f \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = M \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix}$, *wobei*

$$\eta_i = \left(\sum_{j=1}^{n} \alpha_{ij}\xi_j\right) \tag{D1}$$

gilt. Wir bezeichnen diese lineare Abbildung f auch mit $\widehat{M} : K_n \longrightarrow K_m$. Zu jeder linearen Abbildung $f : K_n \longrightarrow K_m$ gibt es genau eine Matrix M mit $f = \widehat{M}$, genannt darstellende Matrix.

Beweis. Die erste Aussage ist wie schon gesagt eine Folge der Rechenregeln (M3) und (M4). Ist umgekehrt $f : K_n \longrightarrow K_m$ eine lineare Abbildung, so erhalten wir n Vektoren $f(e_j), j = 1, \ldots, n$ in K_m, die eine eindeutig bestimmte Basisdarstellung

$$f(e_j) = \sum_{i=1}^{m} \alpha_{ij} \cdot e_i \tag{D2}$$

haben. Wir erhalten also eine Matrix $M = (\alpha_{ij})$ und behaupten $f = \widehat{M}$. Da sowohl f als auch $\widehat{M}$ lineare Abbildungen sind, genügt es nach 5.4.6, ihre Operation auf den Basisvektoren zu vergleichen. Es ist

$$f(e_j) = \sum_{i=1}^{m} \alpha_{ij} e_i = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} = M \cdot e_j,$$

also ist $f = \widehat{M}$. Insbesondere ergibt $\widehat{M}(e_j) = M \cdot e_j$ die j-te Spalte der Matrix M. Deshalb ist M durch $\widehat{M} = f$ eindeutig bestimmt. Außerdem wird das Bild von f von den Vektoren $f(e_j)$, also von den Spaltenvektoren von M aufgespannt.

Folgerung 5.5.4 *Sei $f : K_m \longrightarrow K_n$ eine lineare Abbildung. Dann ist der i-te Spaltenvektor der darstellenden Matrix M mit $\widehat{M} = f$ gegeben durch $f(e_i)$. Weiter ist das Bild von $\widehat{M}$ der von den Spaltenvektoren von M aufgespannte Untervektorraum von K_n.*

Dieses Ergebnis werden wir später viel verwenden. Man beachte jedoch, daß das Ergebnis nur für Vektorräume der Form K_n gilt, nicht aber für andere Vektorräume, z.B. Untervektorräume von K_n. Aus der Folgerung ergibt sich weiter, daß der Rang von $\widehat{M}$ mit der Dimension des von den Spaltenvektoren von M aufgespannten Untervektorraumes übereinstimmt. Wir definieren daher

Definition 5.5.5 Der *Spaltenrang* $\operatorname{rg}(M)$ einer Matrix M ist die Dimension des von den Spaltenvektoren von M aufgespannten Untervektorraumes.

Der *Zeilenrang* einer Matrix M ist die Dimension des von den Zeilenvektoren von M aufgespannten Untervektorraumes.

In Lemma 5.4.2 haben wir gezeigt, daß die Verknüpfung zweier linearer Abbildungen wieder eine lineare Abbildung ergibt. Für Vektorräume der Form K_n erhalten wir den folgenden Zusammenhang.

Folgerung 5.5.6 *Seien $f : K_r \longrightarrow K_m$ und $g : K_m \longrightarrow K_n$ zwei lineare Abbildungen mit den darstellenden Matrizen M bzw. N. Dann ist $N \cdot M$ die darstellende Matrix der linearen Abbildung $gf : K_r \longrightarrow K_n$, d.h. es gilt*

$$\widehat{N}\widehat{M} = \widehat{N \cdot M}.$$

Weiterhin hat die identische Abbildung $\mathrm{id} : K_n \longrightarrow K_n$ *die darstellende Matrix E_n.*

Beweis. Wenn $f = \widehat{M}$ und $g = \widehat{N}$ gelten, so gilt für jeden Vektor $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix} \in K_m$ die Gleichung $gf \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix} = N \cdot (M \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix}) = (N \cdot M) \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix}$ wegen (M1). Das ist die erste Behauptung. Die zweite Behauptung folgt durch (M5).

Definition 5.5.7 Wir nennen eine Matrix M *invertierbar* oder *regulär*, wenn es Matrizen N und N' gibt mit $M \cdot N = E_m$ und $N' \cdot M = E_n$. N (bzw. N') heißt dann auch *inverse* Matrix zu M. Ist eine Matrix nicht regulär, so heißt sie auch *singulär*.

Es ist leicht zu sehen, daß die inverse Matrix zu M eindeutig bestimmt ist, denn gilt $N' \cdot M = E_n$ und $M \cdot N = E_m$, so folgt $N' = N' \cdot E_m = N' \cdot M \cdot N = E_n \cdot N = N$. Wir werden daher diese eindeutig bestimmte inverse Matrix zu M auch mit M^{-1} bezeichnen. Auf Methoden der Berechnung von inversen Matrizen werden wir im nächsten Kapitel eingehen. Wegen der oben dargestellten Zusammenhänge zwischen linearen Abbildungen und Matrizen folgt nun unmittelbar

Folgerung 5.5.8 *Für eine invertierbare Matrix M gilt $\widehat{M^{-1}} = (\widehat{M})^{-1}$.*

Beispiele 5.5.9 1. Die neu eingeführten Begriffe sollen nun an einigen Beispielen erläutert werden. Zunächst betrachten wir die lineare Abbildung $f = \widehat{M} : \mathbb{R}_2 \longrightarrow \mathbb{R}_3$ mit $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. Dann gilt $f\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ 0 \end{pmatrix}$. Diese lineare Abbildung ist injektiv, aber nicht surjektiv. Sie bettet die (ξ_1, ξ_2)-Ebene in den $\mathbb{R}_3$ ein.

2. Die lineare Abbildung $f = \widehat{M}$ mit $M = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ bildet die reelle Gerade $\mathbb{R}$ auf die Raumdiagonale des $\mathbb{R}_3$ ab. Sie ist ebenfalls injektiv und nicht surjektiv.
3. Die lineare Abbildung $f = \widehat{M}$ mit $M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ projiziert den $\mathbb{R}_3$ auf die (ξ_1, ξ_2)-Ebene (,die wir uns wohl in den $\mathbb{R}_3$ eingebettet vorstellen können, die aber genau genommen der $\mathbb{R}_2$ ist). Sie ist surjektiv, aber nicht injektiv.
4. Verwenden wir $M = (1/3, 1/3, 1/3)$, so erhalten wir eine Projektion des $\mathbb{R}_3$ auf die Gerade $\mathbb{R}$. Sie ist surjektiv, aber nicht injektiv. Diese lineare Abbildung ergibt eine Projektion auf die Raumdiagonale, wenn wir noch die oben besprochene Abbildung mit $N = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ nachschalten. Die komponierte lineare Abbildung $\widehat{N \cdot M}$ läßt die Elemente der Raumdiagonalen fest. Die darstellende Matrix ist
$$N \cdot M = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}.$$
Damit haben wir eine lineare Abbildung, die weder injektiv noch surjektiv ist.
5. Ein Beispiel für eine bijektive lineare Abbildung ist die Drehung der Ebene $\mathbb{R}_2$ um 30° mit der Matrix
$$M = \begin{pmatrix} 1/2 \cdot \sqrt{3} & -1/2 \\ 1/2 & 1/2 \cdot \sqrt{3} \end{pmatrix}.$$
Durch sie wird der Basisvektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in der x_1-Richtung auf den Vektor $\begin{pmatrix} 1/2 \cdot \sqrt{3} \\ 1/2 \end{pmatrix}$ abgebildet, d.h. um 30° nach links gedreht, und der Basisvektor $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in der x_2-Richtung auf den Vektor $\begin{pmatrix} -1/2 \\ 1/2 \cdot \sqrt{3} \end{pmatrix}$. Die Drehung ist also eine Linksdrehung. Im mathematischen Sinn wird eine Drehung immer nach links gerechnet, während im technischen Sinn eine Drehung immer nach rechts gerechnet wird.

Wir hatten oben schon bemerkt, daß lineare Abbildungen sich nur dann durch Multiplikation mit Matrizen darstellen lassen, wenn der Vektorraum von der Form K_n ist. Für Vektorräume von anderer Form, z.B. Untervektorräume von K_n, Vektorräume der Form $\mathrm{Hom}(V, W)$ oder K_m^n haben wir jedoch in Satz 5.4.6 ein gutes Hilfsmittel zur Beschreibung von linearen Abbildungen, das sogar auch zu einer Beschreibung mit Hilfe von Matrizen führt. Wir werden in den folgenden Betrachtungen Basen immer mit Indexmengen der Form $\{1, \ldots, n\}$ indizieren, damit eine gewisse Reihenfolge bei den Basisvektoren, den Matrixeinträgen und der Summation festgelegt ist. Wir nennen eine Familie $(b_i | i = 1, \ldots, n)$ von Vektoren in einem Vektorraum V eine *Basisfamilie*, wenn die Menge $\{b_i | i = 1, \ldots, n\}$ eine Basis von V ist und die Vektoren mit verschiedenen Indizes paarweise verschieden sind.

Definition und Lemma 5.5.10 *Seien V und W zwei Vektorräume mit den Basisfamilien $(b_1, \ldots, b_m)$ von V und $(c_1, \ldots, c_n)$ von W. Sei $g : V \longrightarrow W$ eine lineare Abbildung. Dann ist g durch die Vorgabe der Werte $g(b_i) \in W, i = 1, \ldots, m$ schon eindeutig bestimmt. Die Bildvektoren haben eine eindeutige Basisdarstellung*

$$g(b_j) = \sum_{i=1}^{n} \alpha_{ij} \cdot c_i. \qquad (D2)$$

Also ist durch g eine $n \times m$-Matrix $M = (\alpha_{ij})$ bestimmt. Die Matrix M heißt darstellende Matrix *der linearen Abbildung g bezüglich der Basisfamilien $(b_1, \ldots, b_m)$ von V und $(c_1, \ldots, c_n)$ von W. Die Kenntnis von M allein (und die Kenntnis der Basisfamilien von V und W) bestimmt nach Satz 5.4.6 den Homomorphismus g schon vollständig.*

Ist eine beliebige $m \times n$-Matrix M gegeben, so wird durch die oben angegebene Formel genau eine lineare Abbildung g bestimmt, da die Bildvektoren der Basisvektoren b_j beliebig gewählt werden dürfen.

Ist $V = K_m$, $W = K_n$, $f : V \longrightarrow W$ eine lineare Abbildung, so stimmt die darstellende Matrix M von $f = \widehat{M}$ mit der soeben definierten Matrix überein, wenn wir als Basisfamilien in den beiden Vektorräumen die kanonischen Basisfamilien (e_i) verwenden. Der Leser kann das leicht nachrechnen.

Lemma 5.5.11 *Seien drei Vektorräume U, V und W gegeben zusammen mit Basisfamilien (b_i), (c_i) beziehungsweise (d_i), seien weiter lineare Abbildungen $f : U \longrightarrow V$ und $g : V \longrightarrow W$ mit darstellenden Matrizen M bzw. N gegeben, so ist $N \cdot M$ die darstellende Matrix von $gf : U \longrightarrow W$.*

Beweis. Seien $f(b_k) = \sum_j \alpha_{jk} \cdot c_j$ und $g(c_j) = \sum_i \beta_{ij} \cdot d_i$. Dann ist

$$gf(b_k) = g(\sum_{j=1}^{m} \alpha_{jk} \cdot c_j) = \sum_{j=1}^{m} \alpha_{jk} \cdot g(c_j)$$

$$= \sum_{j=1}^{m} \alpha_{jk} \cdot \sum_{i=1}^{r} \beta_{ij} \cdot d_i = \sum_{i=1}^{r} \left(\sum_{j=1}^{m} \beta_{ij} \alpha_{jk} \right) \cdot d_i.$$

Wir betrachten einen Spezialfall des vorhergehenden Lemmas. Sei $f : U \longrightarrow V$ eine bijektive lineare Abbildung. Sei $U = W$. Wir wählen dieselben Basisfamilien (b_i) und (d_i). Sei $g = f^{-1}$. Dann ergibt die Zusammensetzung gf der beiden Abbildungen die identische Abbildung $\mathrm{id}: U \longrightarrow U$. Für diese gilt aber $\mathrm{id}(b_i) = \sum \delta_{ij} \cdot b_j$ mit dem Kronecker Symbol (δ_{ij}). Die darstellende Matrix ist also die Einheitsmatrix E_n. Wir erhalten daher für das Matrizenprodukt $(\alpha_{ij}) \cdot (\beta_{jk}) = E_n$ und symmetrisch $(\beta_{jk}) \cdot (\alpha_{ij}) = E_n$. Die Matrix (α_{ij}) ist damit invertierbar. Offenbar lassen sich diese Schlüsse auch umkehren. Wir haben gezeigt

Folgerung 5.5.12 *Wird die lineare Abbildung $f : V \longrightarrow W$ durch die Matrix (α_{ij}) (bezüglich zweier beliebiger Basisfamilien von V beziehungsweise W) dargestellt, so ist f genau dann bijektiv, wenn die Matrix (α_{ij}) invertierbar ist.*

Folgerung 5.5.13 *Eine reguläre Matrix $M \in K_m^n$ ist quadratisch, d.h. es gilt $m = n$.*

Beweis. Nach 5.4.12 ist $\widehat{M}$ eine lineare Abbildung zwischen den Räumen K_m und K_n gleicher Dimension, also $m = n$.

Folgerung 5.5.14 *Eine Matrix ist genau dann regulär oder invertierbar, wenn sie quadratisch ist und ihr Rang mit ihrer Zeilenzahl übereinstimmt.*

Beweis. Ist eine Matrix M regulär, so folgt die Behauptung nach 5.4.12. Ist die Matrix M quadratisch und stimmen Rang und Zeilenzahl überein, so verschwindet der Kern der linearen Abbildung $\widehat{M}$ nach 5.4.13. Weiter stimmen die Dimensionen von Quelle und Ziel der Abbildung $\widehat{M}$ überein. Nach 5.4.12 ist dann $\widehat{M}$ invertierbar, also auch M.

Ist eine lineare Abbildung $f : K_m \longrightarrow K_n$ durch Multiplikation mit einer Matrix M gegeben, also $f = \widehat{M}$, so kann man den Rang von f leicht aus der Matrix berechnen. Es ist nämlich die Menge der Spaltenvektoren von M eine Erzeugendenmenge des Bildes von f, da $f(e_i) = M \cdot e_i$ der i-te Spaltenvektor der Matrix M ist, und das Bild der Basisfamilie $(e_1, \ldots, e_m)$ unter f eine Erzeugendenmenge von $f(K_m)$ ist. Um also den Rang von f zu berechnen, genügt es, die maximale Anzahl, genannt *(Spalten-) Rang der Matrix M*, von linear unabhängigen Spaltenvektoren von M zu bestimmen. Diese bilden gerade eine Basisfamilie von $f(K_m)$. Die Dimension des Kernes dieser linearen Abbildung ergibt sich dann aus dem Satz 5.4.13.

Lemma 5.5.15 *Der Rang der linearen Abbildung $\widehat{M}$ ist die maximale Anzahl der linear unabhängigen Spaltenvektoren von M.*

Wir werden später zeigen, daß die maximale Anzahl von linear unabhängigen Spaltenvektoren mit der maximalen Anzahl von linear unabhängigen Zeilenvektoren von M übereinstimmt, so daß man gelegentlich eine vereinfachte Berechnung des Ranges von M durchführen kann.

Wir betrachten jetzt einen Vektorraum V mit zwei Basisfamilien $(b_1, \ldots, b_n)$ und $(c_1, \ldots, c_n)$. Da sich jeder Vektor eindeutig als Linearkombination der Basiselemente schreiben läßt, gibt es eindeutig bestimmte Koeffizienten α_{ij} mit $b_i = \sum_{j=1}^n \alpha_{ij} c_j$.

Definition 5.5.16 Für den Vektorraum V seien zwei Basisfamilien $(b_1, \ldots, b_n)$ und $(c_1, \ldots, c_n)$ gegeben mit

$$b_j = \sum_{i=1}^n \alpha_{ij} c_i \tag{D3}$$

Die Matrix $T = (\alpha_{ij})$ heißt *Transformationsmatrix für die Basistransformation von* (b_i) *nach* (c_i).

Sei $\mathrm{id}(b_j) = b_j = \sum_{i=1}^n \alpha_{ij} \cdot c_i$. Dabei seien die Koeffizienten dieselben wie oben bestimmt. Nach Folgerung 5.5.12 ist diese Matrix invertierbar, da die identische Abbildung invertierbar ist. Daraus folgt

Lemma 5.5.17 *Jede Transformationsmatrix für eine Basistransformation ist invertierbar.*

Offenbar ist auch jede Familie $(b_j | j = 1, \ldots, n)$ von Vektoren mit $b_j = \sum_{i=1}^n \alpha_{ij} \cdot c_i$ eine Basis, wenn die Matrix (α_{ij}) invertierbar ist.

Wir leiten jetzt noch die Transformationsformel für die Transformation der Koeffizienten eines Vektors bei einer Basistransformation her. Seien wie oben V ein Vektorraum mit zwei Basisfamilien $(b_1, \ldots, b_n)$ und $(c_1, \ldots, c_n)$. Sei (α_{ij}) die Transformationsmatrix. Sei schließlich $v = \sum_{j=1}^n \beta_j \cdot b_j = \sum_{i=1}^n \gamma_i \cdot c_i$ ein beliebiger Vektor in V. Dann gilt $v = \sum_{j=1}^n \beta_j \cdot b_j = \sum_{j=1}^n \beta_j \cdot \sum_{i=1}^n \alpha_{ij} \cdot c_i = \sum_{i=1}^n \left(\sum_{j=1}^n \alpha_{ij} \cdot \beta_j \right) \cdot c_i$. Wegen der Eindeutigkeit der Basisdarstellung erhalten wir also

$$\gamma_i = \sum_{j=1}^n \alpha_{ij} \cdot \beta_j. \tag{D4}$$

Es besteht ein enger Zusammenhang zwischen den darstellenden Matrizen von Homomorphismen und Koordinatensystemen.

Satz 5.5.18 *Seien V und W Vektorräume mit den Basisfamilien $B = (b_j | j = 1, \ldots, m)$ bzw. $C = (c_i | i = 1, \ldots, n)$. Sei $f : V \longrightarrow W$ eine lineare Abbildung. Die Matrix $M = (\alpha_{ij})$ ist genau dann darstellende Matrix von f bezüglich der Basisfamilien B und C, wenn das Diagramm*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "h_B"'] & W \arrow[d, "h_C"] \\ K_m \arrow[r, "\widehat{M}"] & K_n \end{tikzcd}$$

kommutiert.

Beweis. Sei $N = (\beta_{ij})$ die darstellende Matrix von f. Es ist $h_C f = \widehat{M} h_B$ genau dann, wenn für jeden Basisvektor b_j gilt

$$(\beta_{ij}|i = 1, \ldots, n) = h_C(\sum_{i=1}^{n} \beta_{ij} c_i) = h_C f(b_j)$$

$$= \widehat{M} h_B(b_j) = \sum_{k=1}^{m} \alpha_{ik}\delta_{kj} = (\alpha_{ij}|i = 1, \ldots, n),$$

genau dann, wenn $M = N$.

Satz 5.5.19 *Seien $B = (b_i)$ und $C = (c_i)$ Basisfamilien des Vektorraumes V. Die Matrix T ist genau dann die Transformationsmatrix für die Basistransformation von (b_i) nach (c_i), wenn das Diagramm der Koordinatensysteme*

$$\begin{tikzcd} & V \arrow[dl, "h_B"'] \arrow[dr, "h_C"] & \\ K_n \arrow[rr, "\widehat{T}"] & & K_n \end{tikzcd}$$

kommutiert.

Beweis. Ist T die Transformationsmatrix, so gilt $b_j = \mathrm{id}(b_j) = \sum_{i=1}^{n} \alpha_{ij} c_i$. Damit kommutiert das Diagramm

$$\begin{tikzcd} V \arrow[r, "\mathrm{id}"] \arrow[d, "h_B"'] & V \arrow[d, "h_C"] \\ K_n \arrow[r, "\widehat{T}"] & K_n \end{tikzcd}$$

und damit auch das Diagramm im Satz. Es ist klar, daß auch die Umkehrung gilt.

Hieraus ergibt sich nun die wichtige Formel über die Änderung der darstellenden Matrix bei Koordinatentransformationen.

Satz 5.5.20 *Seien V ein Vektorraum mit den Basisfamilien $B = (b_j|j = 1, \ldots, m)$ und $B' = (b'_j|j = 1, \ldots, m)$ und W ein Vektorraum mit den Basisfamilien $C = (c_i|i = 1, \ldots, n)$ und $C' = (c'_i|i = 1, \ldots, n)$. Seien S die Transformationsmatrix für die Basistransformation von B nach B' und T die Transformationsmatrix für die Basistransformation von C nach C'. Sei $f : V \longrightarrow W$ eine lineare Abbildung mit der darstellenden Matrix M bezüglich der Basisfamilien B und C. Dann ist TMS^{-1} die darstellende Matrix von f bezüglich der Basisfamilien B' und C'.*

Beweis. Das Diagramm

$$\begin{tikzcd} V \arrow[r, "\mathrm{id}"] \arrow[d, "h_{B'}"'] & V \arrow[r, "f"] \arrow[d, "h_B"] & W \arrow[r, "\mathrm{id}"] \arrow[d, "h_C"'] & W \arrow[d, "h_{C'}"] \\ K_m \arrow[r, "\widehat{S}^{-1}"] & K_m \arrow[r, "\widehat{M}"] & K_n \arrow[r, "\widehat{T}"] & K_n \end{tikzcd}$$

kommutiert. Die untere Gesamtabbildung ist dabei $(\widehat{TMS^{-1}})$. Nach 5.5.18 ist dann TMS^{-1} die darstellende Matrix für f.

Eine weitere Operation, die auf Matrizen ausgeübt werden kann, ist die *Transposition.* Ist $M = (\alpha_{ij})$ eine $m \times n$-Matrix, so ist $M^t := (\beta_{kl})$ mit $\beta_{kl} := \alpha_{lk}$ mit $k = 1, \ldots, n$ und $l = 1, \ldots, m$ eine $n \times m$-Matrix, die *Transponierte* oder *transponierte Matrix* von M genannt wird.

Man stelle sich diese Operation so vor, daß das Koeffizientenschema an der Diagonalen $\alpha_{11}, \ldots, \alpha_{ii}, \ldots$ gespiegelt wird, d.h. daß die Terme links unterhalb dieser Diagonalen rechts oberhalb der Diagonalen zu stehen kommen und umgekehrt, insbesondere, daß die Diagonale fest bleibt. Die Transponierte zur Matrix

$$M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

wird also

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

Diese Operation kann daher auch auf nichtquadratische Matrizen angewendet werden. Damit werden aus Zeilenvektoren Spaltenvektoren und umgekehrt.

Man sieht nun leicht ein, daß die Transposition ein Homomorphismus zwischen Vektorräumen von Matrizen ist, allgemeiner daß

$$\begin{aligned} &(\alpha M + \beta N)^t = \alpha M^t + \beta N^t, \\ &(M^t)^t = M, \\ &(M \cdot N)^t = N^t \cdot M^t \end{aligned}$$

gelten.

Übungen 5.5.21 1. Zeigen Sie die folgenden Rechenregeln für die Matrizenmultiplikation:
   a) Für $M \in K_m^n$, $N \in K_n^k$ und $P \in K_k^p$ gilt: $(M \cdot N) \cdot P = M \cdot (N \cdot P)$
   b) Für $M \in K_m^n$, $N \in K_m^n$ und $P \in K_n^k$ gilt: $(M + N) \cdot P = (M \cdot P) + (N \cdot P)$
   c) Für $M \in K_m^n$, $N \in K_n^k$ und $P \in K_n^k$ gilt: $M \cdot (N + P) = (M \cdot N) + (M \cdot P)$

d) Für $M \in K_m^n$, $N \in K_n^k$ und $\lambda \in K$ gilt: $(\lambda M) \cdot N = M \cdot (\lambda N) = \lambda(M \cdot N)$

e) Für $M \in K_m^n$ gilt: $E_m \cdot M = M \cdot E_n = M$

2. Sei V der von den Funktionen $\sin(x)$ und $\cos(x)$ erzeugte Untervektorraum des Vektorraumes $\mathbb{R}^{\mathbb{R}} = \mathrm{Abb}(\mathbb{R}, \mathbb{R})$. Bestimmen Sie eine darstellende Matrix der linearen Abbildung
$$\frac{d}{dx} : V \longrightarrow V.$$
3. Sei V ein endlichdimensionaler Vektorraum und $f : V \longrightarrow V$ eine lineare Abbildung mit $f \circ f = \mathrm{id}_V$. Im Grundkörper K sei $1 + 1 \neq 0$.
   a) Sei $V_+ := \{v \in V \mid f(v) = v\}$ und $V_- := \{v \in V \mid f(v) = -v\}$. Zeigen Sie: $V = V_+ \oplus V_-$. (Hinweis: Erklären Sie, wo die Voraussetzung über den Grundkörper eingeht.)
   b) Zeigen Sie, daß es eine Basis von V gibt, für die die darstellende Matrix von f auf der Diagonalen nur die Einträge 1 und -1 und außerhalb der Diagonalen nur den Eintrag 0 hat.
4. Sei
$$M = \begin{pmatrix} 0 & 2 & -1 \\ 1 & -1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$
Berechnen Sie mit Hilfe der Transformationsformel aus Satz 5.5.20 die darstellende Matrix von $\widehat{M}$ bezüglich der Basis
$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$
von $\mathbb{Q}_3$.
5. Betrachten Sie die Basis $(1, -1, 0), (1, 0, -1), (1, 1, 1)$ des Vektorraumes $\mathbb{Q}^3$. Bestimmen Sie die Transformationsmatrix des Basiswechsels von dieser Basis auf die kanonische Basis und die Transformationsmatrix des Basiswechsels von der kanonischen Basis auf diese Basis. Zeigen Sie, daß beide Transformationsmatrizen zueinander invers sind.
6. Seien $v_1, \ldots, v_n$ und $w_1, \ldots, w_n$ Basen des Vektorraumes V. T sei die Transformationsmatrix des Basiswechsels von (v_i) nach (w_i). Zeigen Sie, daß T^{-1} die Transformationsmatrix des Basiswechsels von (w_i) nach (v_i) ist.
7. Seien $b_1 = (0, 1, 1)$, $b_2 = (1, 0, 1)$, $b_3 = (1, 1, 0)$, $c_1 = (2, 1, 1)$, $c_2 = (1, 2, 1)$ und $c_3 = (1, 1, 2)$. Bestimmen Sie die Transformationsmatrix S von der Basis (b_i) auf die Basis (c_i) von $\mathbb{R}^3$ und die Transformationsmatrix von der Basis (c_i) auf die Basis (b_i).
8. Sei $v_1, \ldots, v_n$ eine Basis des Vektorraumes V. Sei $T = (\alpha_{ij}) \in K_n^n$ eine invertierbare Matrix. Für $j = 1, \ldots, n$ setze

$$w_j := \sum_{i=1}^{n} \alpha_{ij} v_i$$

Zeigen Sie, daß $w_1, \ldots, w_n$ ebenfalls eine Basis von V ist.

9. Entscheiden Sie, ob die folgende Aussage richtig ist (ja/nein).
   Sei $A \in K_m^n$ und $B \in K_n^m$. Es gelte $AB = E_m$. Dann gilt auch $BA = E_n$.
10. Sei

$$V = \left\{ \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} \in \mathbb{R}_4 \,\middle|\, w + x + y + z = 0 \right\}$$

und sei $f : V \longrightarrow V$ der durch

$$f \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix}$$

definierte Homomorphismus. Bestimmen Sie die darstellende Matrix von f bezüglich der Basis von V

$$\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

5.6 Restklassenräume, affine Räume

Da die Vektorraumstruktur eine algebraische Grundstruktur ist, kann man wie in Kapitel 3 Abschnitt 4 Kongruenzrelationen und Restklassen studieren. Wie in 3.4 nennen wir eine Äquivalenzrelation $R \subset V \times V$ eine Kongruenzrelation für den Vektorraum V, wenn R ein Untervektorraum von $V \times V$ ist. Wenn R eine Kongruenzrelation für V ist, dann gilt wieder die Aussage von 3.4.5, daß V/R genau eine Struktur eines Vektorraumes trägt, so daß die Restklassenabbildung $\nu : V \longrightarrow V/R$ eine lineare Abbildung ist.

Satz 5.6.1 *Die Zuordnung, die jedem Untervektorraum $U \subset V$ die Kongruenzrelation $R_U \subset V \times V$ mit*

$$R_U := \{(v, w) \in V \times V | v - w \in U\}$$

zuordnet, ist bijektiv. Die Umkehrabbildung ordnet jeder Kongruenzrelation $R \subset V \times V$ den Untervektorraum

$$U_R := \{v - w | (v, w) \in R\}$$

zu.

Beweis. Da jeder Vektorraum auch eine (abelsche) Gruppe ist, ist $R_U \subset V \times V$ Untergruppe. Für $\lambda \in K$ und $(v, w) \in R_U$ ist $\lambda(v, w) = (\lambda v, \lambda w) \in R_U$, denn $\lambda v - \lambda w = \lambda(v - w) \in U$. Damit ist R_U tatsächlich eine Kongruenzrelation. Ist R gegeben, so ist wiederum U wie in 3.5.3 eine Untergruppe. Für $\lambda \in K$ und $v - w \in U_R$ mit $(v, w) \in R$ gilt $\lambda(v, w) \in R$, also $\lambda(v - w) = \lambda v - \lambda w \in U_R$. Damit ist U_R ein Untervektorraum.

Wir schreiben $V/U := V/R_U$. Nach 3.5.4 ist $v + U = \{v + u | u \in U\}$ dann das allgemeine Element des Restklassenraumes V/U.

Definition 5.6.2 Eine Teilmenge $v + U$ in einem Vektorraum V bezüglich eines Untervektorraumes U heißt auch *affiner Unterraum von V mit Translationsraum U.*

Satz 5.6.3 *(Faktorisierungssatz für Vektorräume) Sei $f : V \longrightarrow W$ eine lineare Abbildung und $U \subset V$ ein Untervektorraum. Wenn $f(U) = 0$, dann gibt es genau eine lineare Abbildung $\overline{f} : V/U \longrightarrow W$, so daß*

$$\begin{tikzcd} V \arrow[r, "\nu"] \arrow[dr, "f"'] & V/U \arrow[d, "\overline{f}"] \\ & W \end{tikzcd}$$

kommutiert.

Beweis. verläuft analog zum Beweis von 3.5.5. Einzig neu ist die Behauptung, daß $\overline{f}$ auch mit der Multiplikation mit Skalaren verträglich ist: $\overline{f}(\lambda\overline{v}) = \overline{f}(\overline{\lambda v}) = \overline{f}\nu(\lambda v) = f(\lambda v) = \lambda f(v) = \lambda\overline{f}(\overline{v})$.

Satz 5.6.4 *Sei $U \subset V$ ein Untervektorraum und U' ein direktes Komplement zu U. Dann ist $f : U' \stackrel{\iota}{\longrightarrow} V \stackrel{\nu}{\longrightarrow} V/U$ ein Isomorphismus.*

Beweis. Es ist $V = U \oplus U'$. Wir definieren die Umkehrabbildung $\bar{g} : V/U \longrightarrow U'$ zu f durch die Teilabbildungen auf den direkten Summanden $(g_1 : U \longrightarrow U') = 0$ und $(g_2 : U' \longrightarrow U') = \mathrm{id}_{U'}$. Nach 5.4.15 ist die lineare Abbildung $g : V \longrightarrow U'$ dadurch eindeutig bestimmt. Da $g(U) = g_1(U) = 0$, folgt nach 5.6.3, daß g durch eine eindeutig bestimmte lineare Abbildung $\bar{g} : V/U \longrightarrow U'$ faktorisiert werden kann, daß also insbesondere $\bar{g}\nu = g$ gilt. Dann ist $f\bar{g}(\bar{v}) = fg(v) = f(g_1(u) + g_2(u')) = fg_{(}u') = f(u') = \nu(u') = \nu(v) = \bar{v} = \mathrm{id}(\bar{v})$, wenn $v \in V$ die Zerlegung $v = u + u'$ bezüglich $V = U \oplus U'$ hat. Weiter ist $\bar{g}f(u') = \bar{g}\nu(u') = g(u') = u' = \mathrm{id}(u')$. Also gelten $f\bar{g} = \mathrm{id}$ und $\bar{g}f = \mathrm{id}$.

Die im vorhergehenden Satz gezeigte Eigenschaft, daß der Restklassenraum V/U auch zu einem Untervektorraum U' von V isomorph ist, ist eine sehr seltene Eigenschaft von algebraischen Strukturen. Der Untervektorraum U' ist ja auch durch V/U oder durch U keineswegs eindeutig bestimmt, wie wir schon bei der Diskussion von direkten Komplementen gesehen haben. Bei Ringen ist diese Aussage z.B. falsch. Keiner der Ringe $\mathbb{Z}/n\mathbb{Z}$ mit $n > 1$ ist zu einem Unterring von $\mathbb{Z}$ isomorph, weil in $\mathbb{Z}$ kein Element der additiven Ordnung n existiert, d.h. es gibt kein Element $x \in \mathbb{Z}$, $x \neq 0$, mit $nx = 0$, was aber in $\mathbb{Z}/n\mathbb{Z}$ z.B. für die Eins gilt. Man sollte daher vermeiden, V/U mit irgendeinem Untervektorraum von V zu identifizieren oder sich auch nur eine solche Identifizierung vorzustellen.

Mit den Elementen von V/U kann man jedoch in sehr natürlicher Weise eine andere Vorstellung verbinden. Ein solches Element, ein affiner Raum, hat die Form $\bar{v} = v + U = \{v + u | u \in U\}$. Er entsteht also durch „Verschiebung" des Untervektorraumes U um den Vektor $v \in V$. Da die Elemente von V/U eine Partition bilden, werden sich zwei solche affinen Unterräume mit demselben U nicht schneiden (es sei denn, sie stimmen überein). Man betrachtet sie dann auch als parallele affine Unterräume. Weiter überdecken sie den gesamten Vektorraum V, wegen $V = \bigcup\{v+U|v \in V\}$. Damit ist V/U ein interessantes weiteres Beispiel für einen Vektorraum, nämlich die Menge der zu U in V parallelen affinen Unterräume mit Translationsraum U. Das ist daher auch ein Grund für die folgende Definition.

Definition 5.6.5 Sei $A = v + U$ ein affiner Unterraum von V mit dem Translationsraum U. Dann ist die Dimension $\dim(A)$ von A definiert als die Dimension von U.

Eigentlich müßte man hier zunächst zeigen, daß U durch A eindeutig bestimmt ist. Man sieht aber leicht $U = \{v - w | v, w \in A\}$.

Wenn man einen Vektor in einem affinen Unterraum $v' \in v + U =: A$ hat, so kann man zu diesem Vektor beliebige Vektoren $u \in U$ addieren, ohne den gegebenen affinen Unterraum A zu verlassen, denn zu v' existiert ein $u' \in U$ mit $v' = v + u'$. Für $u \in U$ gilt dann $v' + u = v + (u' + u) \in v + U = A$. Damit hat eine Operation, die *Translationsoperation*

$$A \times U \ni (v', u) \mapsto v' + u \in A.$$

Die Vektoren aus U verschieben also A „in sich".

Satz 5.6.6 *Sei $f : V \longrightarrow W$ eine lineare Abbildung und sei $U := \mathrm{Ke}(f)$. Wenn $f^{-1}(w) \neq \emptyset$, dann ist $A := f^{-1}(w)$ ein affiner Unterraum von V mit Translationsraum U.*

Beweis. folgt unmittelbar aus 5.4.9.

Übungen 5.6.7 1. Sei $V = \mathbb{R}^4$ und

$$U = \langle (1 \quad 1 \quad 0 \quad 0), (1 \quad -1 \quad 1 \quad 1) \rangle.$$

a) Bestimmen Sie eine Basis für V/U.

b) Bestimmen Sie ein direktes Komplement U' zu U.
c) Bestimmen Sie eine darstellende Matrix für den Isomorphismus $U' \longrightarrow V \longrightarrow V/U$.

2. Sei $U := \mathrm{Span}((1,1,0,0),(0,0,1,1)) \subset \mathbb{Q}^4$. Bestimmen Sie eine Basis von $\mathbb{Q}^4/U$.
3. Sei V ein Vektorraum. U_1 und U_2 seien zwei Untervektorräume von V. Zeigen Sie, daß die Abbildung

$$U_1/(U_1 \cap U_2) \longrightarrow (U_1 + U_2)/U_2, \bar{v} \mapsto \bar{v}$$

ein Isomorphismus ist. (1. Isomorphiesatz)

4. Seien $x, y, z \in \mathbb{R}^5$ linear unabhängig. Zeigen Sie, daß

$$A := \{\alpha x + \beta y + \gamma z | \alpha + \beta + \gamma = 1\}$$

ein affiner Unterraum von $\mathbb{R}^5$ ist und bestimmen Sie den Translationsraum U zu A. Welche Dimension hat A?

6. Matrizen und lineare Gleichungssysteme

Eine der schönsten Anwendungen der Theorie der linearen Abbildungen ist die Theorie der linearen Gleichungssysteme. Ist eine $m \times n$-Matrix M und ein Vektor $b \in K_m$ gegeben, so möchte man gern alle Vektoren $x \in K_n$ bestimmen, die die Gleichung

$$M \cdot x = b$$

oder in Komponentenschreibweise die linearen Gleichungen

$$\begin{array}{ccccc} \alpha_{11} \cdot \xi_1 & + \ldots + & \alpha_{1n} \cdot \xi_n & = & \beta_1 \\ \vdots & & \vdots & & \vdots \\ \alpha_{m1} \cdot \xi_1 & + \ldots + & \alpha_{mn} \cdot \xi_n & = & \beta_m \end{array} \tag{6.1}$$

erfüllen. Wir werden sehen, daß ein enger Zusammenhang mit der Matrizenrechnung besteht. Außerdem liefert uns die bisher entwickelte Theorie schöne Methoden zum Auffinden der Lösungen linearer Gleichungssysteme.

6.1 Lineare Gleichungssysteme

Wir befassen uns in diesem ersten Abschnitt mit den theoretischen Grundlagen der linearen Gleichungssysteme. Insbesondere machen wir weitreichende abstrakte Aussagen über mögliche Lösungen. Erst im nächsten Abschnitt geben wir Methoden zur Berechnung von Lösungen an. Wir wissen dann schon, was wir an Lösungen zu erwarten haben und welche Eigenschaften sie haben werden.

Definition 6.1.1 Eine Gleichung (6.1) mit bekannten Größen α_{ij} und β_k und Unbekannten ξ_i nennt man ein *lineares Gleichungssystem* für die ξ_i. Die Menge

$$\{x \in K_n | M \cdot x = b\}$$

heißt *Lösungsmenge* des linearen Gleichungssystems.

Wir wissen, daß die Multiplikation von links mit einer Matrix M eine lineare Abbildung von K_n nach K_m ist. Daher können wir alle im vorhergehenden Kapitel erworbenen Kenntnisse auf lineare Gleichungssysteme anwenden. Bezeichnen wir wie im vorhergehenden Kapitel $f = \widehat{M}$, so ist

$$f^{-1}(b) = \{x \in K^m | M \cdot x = b\}$$

die Lösungsmenge des linearen Gleichungssystems.

Definition 6.1.2 Das lineare Gleichungssystem $M \cdot x = b$ heißt *homogen*, wenn $b = 0$ gilt, sonst heißt es *inhomogen*.

Ein erster wichtiger Satz über Lösungen von linearen Gleichungssystemen ergibt sich unmittelbar aus einigen schon bewiesenen Behauptungen.

Satz 6.1.3 *Die Lösungsmenge L eines linearen Gleichungssystems $M \cdot x = b$ ist ein affiner Unterraum von K_n. Ist $b = 0$, so ist die Lösungsmenge U des homogenen Gleichungssystems ein Untervektorraum von K_n. Ist x_0 eine Lösung des inhomogenen Gleichungssystems $M \cdot x = b$ und U der Lösungsraum des homogenen Gleichungssystems $M \cdot x = 0$, so gilt für die Lösungsmenge des inhomogenen Gleichungssystems $L = x_0 + U$.*

Beweis. Nach Satz 5.6.6 ist $L = f^{-1}(b) = x_0 + \mathrm{Ke}(f)$ ein affiner Unterraum von K_n, denn $U = \{x \in K_n | M \cdot x = 0\} = \mathrm{Ke}(f)$.

Es gilt auch die Umkehrung des vorstehenden Satzes. Man kann sogar allgemein affine Unterräume mit Hilfe von linearen Gleichungssystemen beschreiben.

Satz 6.1.4 *Sei $B = x_0 + U$ ein affiner Unterraum des K_n. Dann gibt es ein lineares Gleichungssystem, dessen Lösungsmenge B ist.*

Beweis. Man wähle eine lineare Abbildung $f : K_n \longrightarrow K_m$ mit $\mathrm{Ke}(f) = U$ und $f = \widehat{M}$. Das ist immer möglich, wenn $m + \dim(U) \geq n$ ist. Denn dann kann man eine Basis von U zu einer Basis von K_n verlängern, die $\dim(U)$ Basisvektoren von U auf Null in K_m abbilden und die restlichen $n - \dim(U)$ Basisvektoren auf eine entsprechende Anzahl verschiedener Basisvektoren von K_m abbilden. Damit wird nach 5.4.6 1. eine lineare Abbildung $f : K_n \longrightarrow K_m$ mit $\mathrm{Ke}(f) \supset U$ und einem $(n - \dim(U))$-dimensionalen Bildraum definiert. Wegen des Dimensionssatzes 5.4.10 muß dann sogar $\mathrm{Ke}(f) = U$ gelten. Weiter setze man $b := f(x_0)$. Dann ist B die Lösungsmenge von $M \cdot x = b$.

Die Einsicht, nach der ein lineares Gleichungssystem als lineare Abbildung aufgefaßt werden kann, läßt auch Aussagen über die Dimension des Lösungsraumes zu.

Satz 6.1.5 *Sei M eine $m \times n$-Matrix und $M \cdot x = b$ ein lineares Gleichungssystem, das mindestens eine Lösung besitzt. Dann ist die Dimension des Lösungsraumes $n - \mathrm{rg}(M)$.*

Beweis. folgt aus dem Dimensionssatz 5.4.10: $\mathrm{Ke}(\widehat{M}) + \mathrm{rg}(M) = n$.

Mit Hilfe des Ranges von Matrizen können wir nun genau angeben, unter welchen Bedingungen ein lineares Gleichungssystem Lösungen besitzt und wann Lösungen eindeutig bestimmt sind.

Satz 6.1.6 *Sei $M \cdot x = b$ ein lineares Gleichungssystem mit einer $m \times n$-Matrix. Dann gelten:*

1. *Das Gleichungssystem ist genau dann lösbar, wenn gilt* $\mathrm{rg}(M, b) = \mathrm{rg}(M)$.
2. *Das Gleichungssystem ist genau dann für alle $b \in K_m$ lösbar, wenn gilt* $\mathrm{rg}(M) = m$.
3. *Ist das Gleichungssystem lösbar, so ist die Lösung genau dann eindeutig, wenn gilt* $\mathrm{rg}(M) = n$.

Beweis. 1. Die Matrix (M, b) entsteht aus der Matrix M dadurch, daß der Vektor b als Spaltenvektor zur Matrix M rechts hinzugefügt wird. Der Rang von (M, b) ist die Dimension des durch b und alle Spaltenvektoren von M aufgespannten Untervektorraumes von K_m. Dieser Untervektorraum enthält den lediglich von den Spaltenvektoren von M aufgespannten Untervektorraum. Diese beiden Untervektorräume haben gleiche Dimension genau dann, wenn b linear abhängig von den Spaltenvektoren von M ist. Genau dann ist aber das Gleichungssystem lösbar, denn die gesuchten Werte für $\xi_1, \ldots, \xi_m$ sind die Koeffizienten der erforderlichen Linearkombination der Zeilenvektoren von M zur Darstellung von b, also $\sum_{i=1}^m \xi_i a_i = b$ mit Spaltenvektoren a_i von M, was nur eine andere Schreibweise des linearen Gleichungssystems ist.

2. Das Gleichungssystem ist genau dann für alle b lösbar, wenn die Abbildung $\widehat{M}$ surjektiv ist, genau dann, wenn die Spaltenvektoren von M den ganzen Raum K_m aufspannen, genau dann, wenn $\mathrm{rg}(M) = m$ gilt.

3. Es ist wegen 5.4.10 $\mathrm{rg}(M) = n$ genau dann, wenn der Kern von $\widehat{M}$ Null ist. Das bedeutet aber die Eindeutigkeit der Lösungen aller lösbaren Gleichungen $M \cdot x = b$.

Folgerung 6.1.7 *Sei M eine quadratische $(n \times n)$-Matrix. Dann sind die folgenden Aussagen äquivalent:*

1. *M ist regulär oder invertierbar.*
2. *Für alle $b \in K_m$ ist $M \cdot x = b$ lösbar.*
3. *Das lineare Gleichungssystem $M \cdot x = 0$ hat nur die triviale Lösung.*

Beweis. Wir betrachten die durch M dargestellte lineare Abbildung $\widehat{M} : K_n \longrightarrow K_n$. Sie ist nach 5.4.14 genau dann bijektiv, wenn sie injektiv ist und dieses genau dann, wenn sie surjektiv ist. Die Bedingung 2. ist äquivalent zur Surjektivität der linearen Abbildung $\widehat{M}$, die Bedingung 3. zur Injektivität.

Bemerkung 6.1.8 Wenn die Matrix M regulär ist, dann erhält man durch $x_0 := M^{-1} \cdot b$ die (eindeutig bestimmte) Lösung des linearen Gleichungssystems $M \cdot x = b$. Es ist nämlich $M \cdot x_0 = M \cdot M^{-1} \cdot b = E_n \cdot b = b$.

Übungen 6.1.9 1. Entscheiden Sie, ob das lineare Gleichungssystem

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

lösbar ist. Bestimmen Sie gegebenenfalls die Lösungen.

2. Entscheiden Sie, ob das lineare Gleichungssystem

$$\begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

lösbar ist. Bestimmen Sie gegebenenfalls die Lösungen.

3. Entscheiden Sie, ob das lineare Gleichungssystem

$$\begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

lösbar ist. Bestimmen Sie gegebenenfalls die Lösungen.

4. Geben Sie notwendige und hinreichende Bedingungen für $\{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ an, so daß gilt

$$\begin{aligned} \{(x_1, x_2) | \alpha_1 x_1 + \alpha_2 x_2 = 0\} \cap \{(x_1, x_2) | \beta_1 x_1 + \beta_2 x_2 = 0\} \\ = \{(0, 0)\}. \end{aligned}$$

5. Zeigen Sie: Das lineare Gleichungssystem $M \cdot x = b$ ist genau dann lösbar, wenn der Vektor b eine Linearkombination der Spaltenvektoren von A ist.
6. Entscheiden Sie, ob die folgende Aussage richtig ist (ja/nein).
Betrachten Sie ein inhomogenes lineares Gleichungssystem $M \cdot x = b$, das genauso viele Gleichungen wie Unbestimmte umfaßt. Wenn das Gleichungssystem eine eindeutige Lösung besitzt, so ist es auch für jede andere rechte Seite b' lösbar.
7. Bestimmen Sie ein $\alpha \in \mathbb{R}$, so daß gilt

$$5 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + 4 \cdot \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} - \alpha \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 16 \\ 25 \\ 19 \end{pmatrix}.$$

8. Finden Sie α und β mit

$$\alpha \cdot \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} + \beta \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 11 \end{pmatrix}$$

6.2 Das Gaußsche Eliminationsverfahren

Die bisher gefundenen Aussagen über lineare Gleichungssysteme sind vorwiegend theoretischer Art. Wir wenden uns jetzt praktischen Lösungswegen zu. Eines der bekanntesten und wirkungsvollsten Verfahren ist das Gaußsche Eliminationsverfahren. Dazu wandelt man die *erweiterte Koeffizientenmatrix* (M, b) nach einem genau vorgeschriebenen Algorithmus in eine Matrix besonders einfacher Gestalt, in eine sogenannte Stufenmatrix, um. Aus dieser läßt sich die Lösungsmenge einfach ablesen. Gleichzeitig läßt sich auch der Rang des Gleichungssystems direkt ablesen und die Tatsache, ob das Gleichungssystem überhaupt Lösungen besitzt.

Es gibt zwei im wesentlichen äquivalente Wege, die einzelnen Schritte des Gaußschen Eliminationsverfahrens durchzuführen, die Multiplikation der erweiterten Koeffizientenmatrix mit geeigneten Elementarmatrizen von links oder die Durchführung elementarer Zeilenumformungen der erweiterten Koeffizientenmatrix. Beide Verfahren sind leicht als Algorithmen auf dem Computer zu implementieren. Wir beschreiben hier (zunächst) das Verfahren mit Hilfe der elementaren Zeilenumformungen.

Bemerkung 6.2.1 Sei ein lineares Gleichungssystem $M \cdot x = b$ gegeben. Offenbar ändern wir an der Lösungsmenge nichts, wenn wir eine der Gleichungen mit einem Skalarfaktor $\lambda \neq 0$ multiplizieren. Das läuft auf die Multiplikation der entsprechenden Zeile von (M, b) mit $\lambda \neq 0$ hinaus. Insbesondere können wir einen solchen Prozeß rückgängig machen, indem wir dieselbe Gleichung mit dem inversen Faktor λ^{-1} multiplizieren. Ebenso ändern wir an der Lösungsmenge nichts, wenn wir ein Vielfaches einer Gleichung zu einer anderen Gleichung addieren. Auch diesen Prozeß können wir nämlich rückgängig machen, indem wir das gleiche Vielfache abziehen. Er läuft auf die Addition eines Vielfachen einer Zeile zu einer anderen Zeile der erweiterten Matrix (M, b) hinaus. Schließlich können wir in demselben Sinne auch zwei Gleichungen bzw. zwei Zeilen miteinander vertauschen. Das führt uns zu der folgenden Definition.

Definition 6.2.2 Sei N eine Matrix. Eine *elementare Zeilenumformung erster Art* Z_1 ist die Multiplikation einer Zeile von N mit einem Faktor $\lambda \neq 0$. Eine *elementare Zeilenumformung zweiter Art* Z_2 ist die Addition eines Vielfachen einer Zeile zu einer anderen Zeile. Eine *elementare Zeilenumformung dritter Art* Z_3 ist die Vertauschung zweier Zeilen.

Ist N die erweiterte Matrix eines linearen Gleichungssystems, so ändern elementare Zeilenumformungen die Lösungsmenge des linearen Gleichungssystems nicht, d.h. die Gleichungssysteme $M \cdot x = b$ und $M' \cdot x = b'$ haben dieselben Lösungsmengen, wenn (M', b') aus (M, b) durch Anwendung von endlich vielen elementaren Zeilenumformungen hervorgeht.

Elementare Zeilenumformungen an der Matrix M und an der Matrix $(M, 0)$ bewirken eingeschränkt auf M sicherlich dasselbe, weil sie auf der Nullspalte gar keine Änderung hervorrufen. Der (Spalten-)Rang der Matrix M ist aber nach 6.1.5 gleich der Zeilenzahl von M minus Dimension des Lösungsraumes. Da der Lösungsraum von $M \cdot x = 0$ sich bei elementaren Zeilenumformungen von $(M, 0)$ nicht ändert und die Zeilenzahl von $(M, 0)$ ebenfalls konstant bleibt, ändert sich insbesondere auch der Rang der Matrix M nicht. Wir haben also erhalten:

Folgerung 6.2.3 *Elementare Zeilenumformungen einer Matrix lassen deren (Spalten-)Rang invariant.*

Durch Anwendung geeigneter elementarer Zeilenumformungen läßt sich nun eine Matrix wesentlich vereinfachen, nämlich auf die Form einer Stufenmatrix.

Definition 6.2.4 Eine Matrix S ist eine *Stufenmatrix*, wenn für je zwei aufeinanderfolgende Zeilen a_i und a_{i+1} von S folgendes gilt: wenn die linken k Koeffizienten von a_i Null sind, so sind die linken $k+1$ Koeffizienten von a_{i+1} Null (oder a_{i+1} ist der Nullvektor):

$$\forall i = 1, \ldots, m,\ k = 1, \ldots, n$$
$$[(\forall j = 1, \ldots, k[\alpha_{ij} = 0]) \Rightarrow (\forall j = 1, \ldots, k+1[\alpha_{(i+1)j} = 0])].$$

Damit hat eine Stufenmatrix die folgende Form

$$\begin{pmatrix} 0, \ldots, 0, \alpha_{1j_1} & \ldots & & \ldots & \alpha_{1n} \\ & \alpha_{2j_2} & \ldots & \ldots & \alpha_{2n} \\ & & & \vdots & \vdots \\ & & \alpha_{kj_k} & \ldots & \alpha_{kn} \\ & & & & 0 \\ & & & & \vdots \\ & & & & 0 \end{pmatrix} \tag{6.2}$$

Ein lineares Gleichungssystem $M \cdot x = b$ hat genau dann eine Stufenmatrix als erweiterte Koeffizientenmatrix, wenn es die folgende Form hat:

$$\begin{array}{rcrl} \alpha_{1j_1} \cdot \xi_{j_1} + \ldots & & \ldots + \alpha_{1n} \cdot \xi_n = & \beta_1 \\ & \alpha_{2j_2} \cdot \xi_{j_2} + \ldots & \ldots + \alpha_{2n} \cdot \xi_n = & \beta_2 \\ & \vdots & & \vdots \\ & & \alpha_{kj_k} \cdot \xi_{j_k} + \ldots + \alpha_{kn} \cdot \xi_n = & \beta_k \\ & & 0 = & \beta_{k+1} \\ & & 0 = & 0 \\ & & \vdots & \vdots \\ & & 0 = & 0 \end{array} \tag{6.3}$$

mit $j_1 < j_2 < \ldots < j_k \leq n$.

Satz 6.2.5 *Jede Matrix M läßt sich durch Anwendung geeigneter elementarer Zeilenumformungen in eine Stufenmatrix S umformen.*

Beweis. Wir wollen den Beweis durch vollständige Induktion nach der Anzahl der Spalten von M durchführen. Dazu beschreiben wir einen Algorithmus $\mathcal{E}$ bestehend aus mehreren elementaren Zeilenumformungen. Dieser wird auf eine Matrix angewendet und ergibt eine Matrix mit kleinerer Spaltenzahl, wodurch eine rekursive Anwendung möglich wird.

Algorithmus 6.2.6 (Algorithmus $\mathcal{E}$ der Umformungen von M, Eliminationsalgorithmus)

$\mathcal{E}_1$: Wenn die erste Spalte von M nur mit Nullen besetzt ist, so führen wir gar keine Umformungen durch und gehen unmittelbar zu Schritt $\mathcal{E}_5$ weiter.

$\mathcal{E}_2$: Wenn der Koeffizient in der ersten Zeile und der ersten Spalte von Null verschieden ist, so gehen wir unmittelbar zu Schritt $\mathcal{E}_4$ weiter.

$\mathcal{E}_3$: Wenn der Koeffizient in der ersten Zeile und der ersten Spalte Null ist und ein von Null verschiedener Koeffizient in der ersten Spalte und der i-ten Zeile von M steht, so vertauschen wir die i-te Zeile mit der ersten Zeile, führen also eine elementare Zeilenumformung dritter Art durch.

Wir bemerken an dieser Stelle, daß hier offenbar mehrere Wahlmöglichkeiten bestehen. Wir werden darauf zurückkommen, wenn wir später das Pivot-(Drehpunkt-)Verfahren besprechen.

Wir können also nach dieser Umformung davon ausgehen, daß in der ersten Zeile und ersten Spalte ein von Null verschiedener Koeffizient α_{11} steht. Seien jetzt die Koeffizienten der ersten Spalte

$$(\alpha_{11}, \alpha_{21}, \ldots, \alpha_{m1})^t = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{pmatrix}.$$

$\mathcal{E}_4$: In der so erhaltenen Matrix addieren wir das $(-\alpha_{11}^{-1}) \cdot \alpha_{i1}$-fache der ersten Zeile zur i-ten Zeile für alle $i = 2, \ldots, n$. Das sind elementare Zeilenumformungen zweiter Art. Dadurch erreichen wir, daß in der ersten Spalte der Vektor $(\alpha_{11}, 0, \ldots, 0)^t$ mit $\alpha_{11} \neq 0$ steht.

Damit ist ein Schritt des Algorithmus $\mathcal{E}$ vollständig beschrieben. Am Schluß haben wir durch Anwendung von $\mathcal{E}$ auf M eine Matrix

$$N = \begin{pmatrix} \alpha_{11} & \beta_{12} & \ldots & \beta_{1n} \\ 0 & \beta_{22} & \ldots & \beta_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ 0 & \beta_{m2} & \ldots & \beta_{mn} \end{pmatrix}$$

erhalten, deren erste Spalte entweder der Nullvektor oder der Vektor $(\alpha_{11}, 0, \ldots, 0)^t \neq 0$ ist.

$\mathcal{E}_5$: Ist der erste Spaltenvektor von N der Nullvektor, so betrachten wir jetzt eine Teilmatrix der Matrix $N = (\beta_{ij} | i = 1, \ldots, m, j = 1, \ldots, n)$, nämlich die Matrix

$$N' := \begin{pmatrix} \beta_{12} & \beta_{13} & \ldots & \beta_{1n} \\ \beta_{22} & \beta_{23} & \ldots & \beta_{2n} \\ \vdots & \vdots & & \vdots \\ \beta_{m2} & \beta_{m3} & \ldots & \beta_{mn} \end{pmatrix}.$$

Im anderen Fall mit $(\alpha_{11}, 0, \ldots, 0)^t \neq 0$ als erster Spalte betrachten wir

$$N' := \begin{pmatrix} \beta_{22} & \beta_{23} & \ldots & \beta_{2n} \\ \beta_{32} & \beta_{33} & \ldots & \beta_{3n} \\ \vdots & \vdots & & \vdots \\ \beta_{m2} & \beta_{m3} & \ldots & \beta_{mn} \end{pmatrix}.$$

Wenn wir eine elementare Zeilenumformung an der Matrix N' vornehmen, so können wir dieselbe elementare Zeilenumformung auch an der größeren Matrix N vornehmen, ohne in N die erste Spalte $(0, 0, \ldots, 0)^t$ bzw. $(\alpha_{11}, 0, \ldots, 0)^t$ zu verändern. Das sieht man sofort für jede einzelne der möglichen elementaren Zeilenumformungen, weil durch sie in der ersten Spalte nur die Nullen betroffen sind. In Falle des ersten Spaltenvektors $(\alpha_{11}, 0, \ldots, 0)^t$ wird auch die erste Zeile der Matrix N nicht geändert, weil für diese Zeile keine Umformungen von N' induziert werden.

Da die Matrix N' kleiner als die Matrix M ist (gemessen an der Anzahl der Spalten), kann man sie per Induktionsannahme durch endlich viele elementare Zeilenumformungen in eine Stufenmatrix umformen. Dieselben Umformungen machen dann aber auch die Matrizen M bzw. N zu Stufenmatrizen. Der Prozeß bricht natürlich ab, wenn die Matrix N' null Spalten oder null Zeilen hat.

Insgesamt haben wir damit einen Algorithmus $\mathcal{E}$ beschrieben, der mit Hilfe von elementaren Zeilenumformungen aus einer beliebigen Matrix M eine Stufenmatrix S macht.

Mit den angegebenen Umformungsverfahren kann man jetzt lineare Gleichungssysteme auf eine wesentlich einfachere Form bringen, nämlich die Form (6.3). Das führt zu dem folgenden

Satz 6.2.7 *(Das Gaußsche Eliminationsverfahren) Sei $M \cdot x = a$ ein lineares Gleichungssystem. Man löst dieses Gleichungssystem, indem man (M, a) zunächst mit elementaren Zeilenumformungen auf Stufenform (S, b) bringt. Ist in der Darstellung (6.3) des Gleichungssystems in Stufenform dann $\beta_{k+1} = 0$ (oder keine $k+1$-te Gleichung vorhanden), so ist das Gleichungssystem lösbar, sonst nicht. Man erhält alle Lösungen, indem man für alle $j \notin \{i_1, \ldots, i_k\}$ beliebige Werte für die ξ_j wählt und dann mit Hilfe von (6.3) die restlichen Werte der $\xi_{i_1}, \ldots, \xi_{i_k}$ mit Hilfe der Formeln*

$$\begin{aligned}
\xi_{i_k} &= \alpha_{ki_k}^{-1}(\beta_k - \sum_{j=i_k+1}^{n} \alpha_{kj}\xi_j) \\
&\vdots \\
\xi_{i_2} &= \alpha_{2i_2}^{-1}(\beta_2 - \sum_{j=i_2+1}^{n} \alpha_{2j}\xi_j) \\
\xi_{i_1} &= \alpha_{1i_1}^{-1}(\beta_1 - \sum_{j=i_1+1}^{n} \alpha_{1j}\xi_j)
\end{aligned} \tag{6.4}$$

(Rückwärtssubstitution) ausrechnet.

Beweis. Die Lösungsmengen der linearen Gleichungssysteme $M \cdot x = a$ und $S \cdot x = b$ stimmen überein, wenn die Matrix (S, b) aus der Matrix (M, a) durch elementare Zeilenumformungen hervorgeht, da sich nach der Bemerkung 6.2.1 die Lösungsmengen eines linearen Gleichungssystems bei elementaren Zeilenumformungen der erweiterten Gleichungsmatrix nicht ändern. Das im Satz behauptete Verhalten der Lösungen ist dann aber direkt aus der Form (6.3) durch Auflösung nach den verbleibenden ξ_i abzulesen.

Mit diesem Satz können wir sowohl eine partikuläre Lösung eines linearen inhomogenen Gleichungssystems bestimmen als auch den Lösungsraum eines linearen homogenen Gleichungssystems. Damit kann man dann nach Satz 6.1.2 die gesamte Lösungsmenge eines inhomogenen Gleichungssystems bilden.

Da sich der (Spalten-)Rang einer Matrix bei Anwendung des Gaußschen Eliminationsverfahrens nicht ändern, können wir ihn an der zugehörigen Stufenmatrix ablesen. Der Rang einer Stufenmatrix ist aber offenbar die Anzahl der Stufen.Durch Addition geeigneter Vielfacher der ersten Spalte zu den anderen Spalten können wir in den anderen Spalten an erster Stelle Nullen erhalten. Ebenso erhalten wir durch Addition geeigneter Vielfacher der Spalte der zweiten Stufe auch an der zweiten Stellen lauten Nullen. So wird jede Spalte entweder der Nullvektor oder ein Einheitsvektor. Durch Umkehren dieses Verfahrens sieht man, daß alle Spalten Linearkombinationen der Stufen sind, die selbst linear unabhängig sind. Wir erhalten also

Satz 6.2.8 *Der (Spalten-)Rang einer Matrix M ist die Anzahl der Stufen einer aus M durch das Gaußsche Eliminationsverfahren erhaltenen Stufenmatrix.*

Folgerung 6.2.9 *Eine quadratische $(n \times n)$-Matrix ist genau dann invertierbar, wenn ein zugehörige Stufenmatrix n Stufen hat.*

Beweis. Dann hat nämlich die Stufenmatrix den Rang n.

Satz 6.2.10 *Jede elementare Zeilenumformung (und damit jede Folge von elementaren Zeilenumformungen) an einer Matrix M ist Ergebnis einer Multiplikation von links mit einer regulären Matrix U, genannt Umformungsmatrix.*

Beweis. Jede elementare Zeilenumformung an einer Matrix M kann beschrieben werden durch Multiplikation mit einer Matrix U_i auf M von links. Das geht aus den folgenden Matrizenprodukten hervor.

$$\begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & \lambda & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{i1} & \dots & \alpha_{in} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} = U_1 \cdot M$$

$$= \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \lambda\alpha_{i1} & \dots & \lambda\alpha_{in} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}$$

Die zur Multiplikation von links verwendete Matrix U_1 hat Einsen in der Diagonalen bis auf den Koeffizienten $\beta_{ii} = \lambda$, ist also von der Form $E_n + (\lambda - 1)E_{ii}$, wobei allgemein E_n die Einheitsmatrix und E_{ij} eine mit Nullen und einem Koeffizienten 1 an der Stelle (i, j) besetzte Matrix bezeichnen. Die zur Multiplikation verwendete Matrix hat $E_n + (\lambda^{-1} - 1)E_{ii}$ als inverse Matrix. Sie ist daher invertierbar und heißt *Elementarmatrix erster Art.*

Weiter ist

$$\begin{pmatrix} 1 & \dots & & \dots & & \dots & 0 \\ \vdots & & & & & & \vdots \\ & & 1 & \dots & \lambda & & \\ \vdots & & & & \vdots & & \vdots \\ & & & & 1 & & \\ \vdots & & & & & & \vdots \\ 0 & \dots & & \dots & & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{i1} & \dots & \alpha_{in} \\ \vdots & & \vdots \\ \alpha_{j1} & \dots & \alpha_{jn} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} = U_2 \cdot M$$

$$= \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{i1} + \lambda\alpha_{j1} & \dots & \alpha_{in} + \lambda\alpha_{jn} \\ \vdots & & \vdots \\ \alpha_{j1} & \dots & \alpha_{jn} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}.$$

Die zur Multiplikation von links verwendete Matrix U_2 hat Einsen in der Diagonalen und einen Koeffizienten $\beta_{ij} = \lambda$ mit $i \neq j$. Alle anderen Koeffizienten sind Null. Diese Matrix, die man als $E_n + \lambda E_{ij}$ schreiben kann, hat als Inverse $E_n - \lambda E_{ij}$ und ist damit invertierbar. Sie heißt *Elementarmatrix zweiter Art.*

Schließlich ist

$$\begin{pmatrix} 1 & \dots & & \dots & & \dots & 0 \\ \vdots & & & & & & \vdots \\ & & 0 & \dots & 1 & & \\ \vdots & & \vdots & & \vdots & & \vdots \\ & & 1 & \dots & 0 & & \\ \vdots & & & & & & \vdots \\ 0 & \dots & & \dots & & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{i1} & \dots & \alpha_{in} \\ \vdots & & \vdots \\ \alpha_{j1} & \dots & \alpha_{jn} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} = U_3 \cdot M$$

$$= \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{j1} & \dots & \alpha_{jn} \\ \vdots & & \vdots \\ \alpha_{i1} & \dots & \alpha_{in} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}.$$

Die zur Multiplikation von links verwendete Matrix U_3 hat Einsen in der Diagonalen und Nullen an allen anderen Stellen mit Ausnahme von $\beta_{ii} = \beta_{jj} = 0$ und $\beta_{ij} = \beta_{ji} = 1$ für ein Paar $i \neq j$. Diese Matrix ist zu sich selbst invers und heißt *Elementarmatrix dritter Art.*

Wir haben also gesehen, daß man elementare Zeilenumformungen durch Multiplikation von links mit gewissen (invertierbaren) Elementarmatrizen erzeugen kann. Damit kann auch jede Folge von elementaren Zeilenumformungen durch Multiplikation mit einer regulären Matrix, dem Produkt der Elementarmatrizen, beschrieben werden.

Bemerkung 6.2.11 Wenn wir das Gleichungssystem $M \cdot x = b$ auf die oben beschriebene Weise gelöst haben und anschließend ein Gleichungssystem $M \cdot x = c$ lösen müssen, so müssen wir das gesamte Eliminationsverfahren neu durchrechnen. Es lohnt sich in diesen Fällen von vornherein eine etwas umfangreichere Umformung vorzunehmen. Wir ersetzen bei der Umformung in Stufenform die Matrix (M, b) durch die Matrix (M, E_m), wobei E_m die Einheitsmatrix ist. Dann formen wir diese Matrix auf Stufenform nach dem bekannten Verfahren und erhalten eine Matrix der Form (S, A). Insbesondere hat die Matrix S Stufenform. Dabei steuert der Übergang von M auf S die Wahl der elementaren Zeilenumformungen, und diese wiederum schaffen eine entsprechende Umformung von E_m auf A. Da A aus E_m durch elementare Spaltenumformungen hervorgeht und E_m regulär ist (also vom Rang m), ist auch A regulär.

Die so gewonnene Matrix A ist die im vorhergehenden Satz beschriebene Umformungsmatrix. Die elementaren Zeilenumformungen von (M, E_m) lassen sich nämlich wie folgt beschreiben:

$$U \cdot (M, E_m) = (U \cdot M, U \cdot E_m) = (S, A),$$

woraus $U = A$ folgt. Insbesondere ist $S = A \cdot M$.

Für ein beliebiges lineares Gleichungssystem der Form $M \cdot x = b$ erhält man dann die Zeilenstufenform als $S \cdot x = (A \cdot M) \cdot x = A \cdot b$. Wenn man also jetzt die Matrizen S und A berechnet hat, genügt es, für jede Wahl von b nur $A \cdot b$ zu berechnen und $S \cdot x = A \cdot b$ durch Rückwärtssubstitution zu lösen.

Übungen 6.2.12 1. Seien $M \in K_m^n, N \in K_n^m$ Matrizen mit der Eigenschaft $M \cdot N = E_m$. Zeigen Sie, das M Spaltenrang m und N Zeilenrang m hat.

2. Verwenden Sie daß Gaußsche Eliminationsverfahren, um die folgenden Gleichungssysteme zu lösen:

a)

$$\begin{aligned} 4x + 4y + 4z &= 24 \\ 2x - y + z &= -9 \\ x - 2y + 3z &= 1 \end{aligned}$$

b)

$$\begin{aligned} 2x + 2y \phantom{{}+z} &= -2 \\ y + z &= 4 \\ x \phantom{{}+2y} + z &= 1 \end{aligned}$$

c)

$$\begin{aligned} x + 2y \phantom{{}+4z} - w &= 3 \\ 2x \phantom{{}+2y} + 4z + 2w &= -6 \\ x + 2y - z \phantom{{}+w} &= 6 \\ 2x - y + z + w &= -3 \end{aligned}$$

d)

$$\begin{aligned} x + 3y - 2z - w &= 9 \\ 2x + 4y \phantom{{}-2z} + 2w &= 10 \\ -3x - 5y + 2z - w &= -15 \\ x - y - 3z + 2w &= 6 \end{aligned}$$

6.3 Inverse Matrizen, die LU-Zerlegung und die Pivot-Methode

In diesem Abschnitt sind wir an quadratischen $(n \times n)$-Matrizen M und den durch sie beschriebenen Gleichungssystemen $M \cdot x = a$ interessiert. Nach 6.2.9

ist die Matrix M genau dann regulär oder invertierbar, wenn die Stufenmatrix zu M genau n Stufen hat, d.h. wenn sie den Rang n hat.

Wir wollen zunächst einen Algorithmus zur Berechnung der Inversen einer Matrix herleiten. Dazu verändern wir den in 6.2.6 eingeführten Algorithmus $\mathcal{E}$ zu einem Algorithmus $\mathcal{J}$, und zwar fügen wir vor dem Schritt $\mathcal{E}_4$ einen weiteren Schritt ein und ergänzen den Schritt $\mathcal{E}_4$, so daß unsere neuer Algorithmus $\mathcal{J}$ aus sechs Schritten besteht:

Algorithmus 6.3.1 ($\mathcal{J}$ zum Gauß-Jordan-Verfahren)
$\mathcal{J}_1 = \mathcal{E}_1, \quad \mathcal{J}_2 = \mathcal{E}_2, \quad \mathcal{J}_3 = \mathcal{E}_3,$

$\mathcal{J}_4$: Der Koeffizient α_{11} in der ersten Zeile und ersten Spalte ist von Null verschieden. Wir multiplizieren die erste Zeile der Matrix mit α_{11}^{-1}, so daß nunmehr an der linken oberen Ecke der Matrix eine Eins steht.

$\mathcal{J}_5$: In der so erhaltenen Matrix addieren wir das $(-\alpha_{i1})$-fache der ersten Zeile zur i-ten Zeile für alle $i = 2, \ldots, n$. Das sind elementare Zeilenumformungen zweiter Art. Dadurch erreichen wir, daß in der ersten Spalte der Vektor $(1, 0, \ldots, 0)^t$ steht.

$\mathcal{J}_6$: Nach rekursiver Durchführung der Schritte $\mathcal{J}_1, \ldots, \mathcal{J}_5$ erhalten wir eine Stufenmatrix mit $k := \mathrm{rg}(M)$ Stufen, die auf den Stufen die Koeffizienten $\alpha_{ij_i} = 1$ haben. Nunmehr addieren wir für alle $i = k, \ldots, 1$ (also rückwärts) das $(-\alpha_{ij})$-fache der i-ten Zeile zur j-ten Zeile für alle $j = 1, \ldots, j_i - 1$ und erreichen dadurch, daß auch über den mit 1 besetzten Stufen jeweils Koeffizienten 0 stehen. Die Matrix hat also die Form

$$\begin{pmatrix}
0 \ldots 1\ \alpha_{1(j_1+1)} \ldots 0 & & \ldots 0 \ldots \alpha_{1n} \\
 & 1\ \alpha_{2(j_2+1)} & \ldots 0 \ldots \alpha_{2n} \\
 & & \vdots \quad\quad \vdots \\
 & & 1 \ldots \alpha_{kn} \\
 & & 0 \\
 & & \vdots \\
 & & 0
\end{pmatrix}$$

Der hierfür notwendige zusätzliche Rechenaufwand ist dabei geringfügig, insbesondere, da bei der Bestimmung der Lösung eines linearen Gleichungssystems die Rückwärtssubstitution vereinfacht wird zu einer einfachen Substitution

$$\begin{aligned}
\xi_{j_1} &= \beta_1 - \sum_{j > j_1, j \notin \{j_1, \ldots, j_k\}} \alpha_{1j} \xi_j \\
\xi_{j_2} &= \beta_2 - \sum_{j > j_2, j \notin \{j_1, \ldots, j_k\}} \alpha_{2j} \xi_j \\
&\vdots \\
\xi_{j_k} &= \beta_k - \sum_{j > j_k, j \notin \{j_1, \ldots, j_k\}} \alpha_{kj} \xi_j
\end{aligned}$$

Nach Übergang zur Stufenform mit dem Algorithmus $\mathcal{J}$ beginnen die Stufen jeweils mit 1. Weiter sind die Spaltenvektoren an den Stufen jeweils kanonische Einheitsvektoren.

Wenn wir den Algorithmus $\mathcal{J}$ nun auf eine Matrix (M, E_n) anwenden und M dabei regulär ist, dann erhalten wir (E_n, A), wobei die Stufenform S von M in besonders einfacher Form als Einheitsmatrix auftreten muß, da M den Rang n hat, also alle kanonischen Einheitsvektoren in der Stufenform auftreten. Weiter wissen wir aus 6.2.11, daß $E_n = A \cdot M$ gilt, so daß A die inverse Matrix zu M ist.

Algorithmus 6.3.2 (Algorithmus zur Bestimmung der Inversen einer Matrix) Wenn M eine reguläre Matrix ist und der Algorithmus $\mathcal{J}$ auf (M, E_n) angewendet wird, so erhält man als Ergebnis (E_n, M^{-1}).

Wir haben sogar allgemeiner den folgenden Satz bewiesen.

Satz 6.3.3 *Sei M eine (quadratische) $n \times n$-Matrix. Die Matrix M ist genau dann invertierbar, wenn sie durch elementare Zeilenumformungen in eine Einheitsmatrix übergeführt werden kann. Ist das der Fall, so erhält man die inverse Matrix zu M, indem man dieselben Zeilenumformungen, mit denen man M in die Einheitsmatrix überführt, auf die Einheitsmatrix in derselben Reihenfolge anwendet.*

Beweis. Die Umformung mit elementaren Zeilenumformungen geschieht nach 6.2.10 und 6.2.11 immer durch Multiplikation mit invertierbaren Matrizen U von links. Es gilt also $U \cdot (M, E_n) = (U \cdot M, U \cdot E_n) = (E_n, A)$ genau dann, wenn $U \cdot M = E_n$, $U \cdot E_n = A$ ist, also wenn $A = U = M^{-1}$ ist.

Wir haben gesehen, daß man elementare Zeilenumformungen durch Multiplikation von links mit gewissen (invertierbaren) Elementarmatrizen erzeugen kann. Da sich für jede invertierbare Matrix M die Einheitsmatrix E_n durch geeignete elementare Zeilenumformungen in die inverse Matrix M^{-1} überführen läßt, erhalten wir mit den entsprechenden Elementarmatrizen $F_1, ..., F_r$ die Gleichung $F_1 \cdot \ldots \cdot F_r \cdot E_n = M^{-1}$ oder $F_1 \cdot \ldots \cdot F_r = M^{-1}$. Da M die inverse Matrix von M^{-1} ist, können wir auch M in dieser Weise schreiben und haben gezeigt:

Folgerung 6.3.4 *Jede invertierbare Matrix läßt sich als Produkt von Elementarmatrizen darstellen.*

In 6.2.3 haben wir schon gesehen, daß sich der Spaltenrang einer Matrix bei elementaren Zeilenumformungen nicht ändert. Er ändert sich daher auch nicht bei Multiplikation mit einer invertierbaren Matrix A von links. Da die linearen Abbildungen $\widehat{MA} = \widehat{M}\widehat{A} : K_m \longrightarrow K_n$ und $\widehat{M} : K_m \longrightarrow K_n$ dasselbe Bild haben, ändert sich der Spaltenrang von M auch nicht bei Multiplikation mit einer invertierbaren Matrix A von rechts bzw. bei elementaren Spaltenumformungen.

Folgerung 6.3.5 *Elementare Zeilen- bzw. Spaltenumformungen einer Matrix lassen deren (Spalten-)Rang invariant. Insbesondere ist* $\mathrm{rg}(AMB) = \mathrm{rg}(M)$ *für invertierbare Matrizen* A, B.

Man ist oft daran interessiert, eine Matrix in eine besonders einfache Form umzuwandeln. Dabei muß natürlich die Methode für die Umformung festgelegt werden. Ein solche besonders einfache Form nennt man häufig auch Normalform der Matrix.

Satz 6.3.6 *(Normalformensatz)*

1. *Zu jeder Matrix M gibt es invertierbare Matrizen A und B mit*
$$M = A \begin{pmatrix} E_i & 0 \\ 0 & 0 \end{pmatrix} B.$$
Dabei ist E_i die Einheitsmatrix mit i Zeilen und Spalten und $i = \mathrm{Rang}(M)$.
2. *Jede Matrix M läßt sich durch elementare Zeilen- und Spaltenumformungen auf die Form $\begin{pmatrix} E_i & 0 \\ 0 & 0 \end{pmatrix}$ bringen.*

Beweis. 1. Wir fassen M als darstellende Matrix einer linearen Abbildung $f : V \longrightarrow W$ bezüglich fester Basen von V bzw. W auf. Wenn wir eine neue Basisfamilie für V wählen, indem wir eine Basisfamilie durch ein Komplement U von $\mathrm{Ke}(f)$ legen und mit einer Basisfamilie von $\mathrm{Ke}(f)$ zu einer Basisfamilie von V vervollständigen, dann ist das Bild der Basisfamilie von U unter f aus Dimensionsgründen eine Basisfamilie von $\mathrm{Bi}(f)$. Wenn wir diese zu einer Basisfamilie von W vervollständigen, dann hat die darstellende Matrix von f bezüglich dieser neuen Basisfamilien die Form $\begin{pmatrix} E_i & 0 \\ 0 & 0 \end{pmatrix}$ mit $i = \mathrm{Rang}(M)$. Die Basistransformationen bewirken Multiplikation von links bzw. rechts mit invertierbaren Matrizen: $M = A \begin{pmatrix} E_i & 0 \\ 0 & 0 \end{pmatrix} B$.

2. Wegen Folgerung 6.3.4 lassen sich die invertierbaren Matrizen A und B bzw. ihre Inversen als Produkte von Elementarmatrizen schreiben, die dann entsprechende Zeilen- und Spaltenumformungen bewirken.

Man kann die Aussage aus Teil 1. des Satzes auch so ausdrücken: zu jeder Matrix M gibt es eine eindeutig bestimmte Zahl i und invertierbare Matrizen A' und B', so daß gilt
$$A'MB' = \begin{pmatrix} E_i & 0 \\ 0 & 0 \end{pmatrix}.$$
Insbesondere ist die Zahl $i = \mathrm{rg}(M)$ durch den Spaltenrang von M festgelegt. Da die Aussage weder von Zeilen noch von Spalten abhängig ist, erhalten wir

Folgerung 6.3.7 *Der Zeilenrang und der Spaltenrang einer Matrix M stimmen überein.*

Schon in der Bemerkung 6.2.11 haben wir die Bedeutung und mögliche Anwendungen der Umformungsmatrix erkannt. Insbesondere haben wir gesehen, daß Umformungsmatrizen immer regulär und damit quadratisch sind. Wir wollen sie etwas systematischer einsetzen. Dazu definieren wir

Definition 6.3.8 Eine quadratische Matrix P heißt *Permutationsmatrix*, wenn sie sich als ein Produkt von Elementarmatrizen des Typs U_3 schreiben läßt. Eine quadratische Matrix L (lower) heißt *untere Dreiecksmatrix*, wenn alle Koeffizienten α_{ij} mit $i < j$ oberhalb der Hauptdiagonalen der Matrix Null sind. Eine Matrix U (upper) heißt *obere Dreiecksmatrix*, wenn alle Koeffizienten α_{ij} mit $i > j$ unterhalb der Hauptdiagonalen Null sind. Eine Matrix D heißt *Diagonalmatrix*, wenn sie sowohl eine obere als auch eine untere Dreiecksmatrix ist, d.h. wenn die einzig von Null verschiedenen Koeffizienten auf der Hauptdiagonalen liegen.

Wir beachten, daß für eine reguläre obere Dreiecksmatrix alle Koeffizienten auf der Hauptdiagonalen von Null verschieden sein müssen, sonst könnte man sie auf Stufenform bringen, bei der mindestens eine Stufe mehr als einen Schritt „einrückt", was für eine reguläre Matrix nicht möglich ist, denn dann würde die letzte Zeile der Nullvektor. Damit ist eine reguläre obere Dreiecksmatrix auch schon in Stufenform gegeben.

Lemma 6.3.9 *Das Produkt von zwei unteren Dreiecksmatrizen ist eine untere Dreiecksmatrix.*

Beweis. Wir betrachten einen Koeffizienten des Produkts $a := \sum_{j=1}^{n} \alpha_{ij}\beta_{jk}$ für $i < k$. Da $\alpha_{ij} = 0$ für $i < j$ gilt, ist $a = \sum_{j=1}^{i} \alpha_{ij}\beta_{jk}$. Da $\beta_{jk} = 0$ für $j < k$ gilt, ist $a = \sum_{j=k}^{i} \alpha_{ij}\beta_{jk}$. Wegen $j < k$ sind also keine Summanden zu addieren, und es folgt $a = 0$.

Für $\lambda \in K$ bezeichne $U_{ij}(\lambda)$ die Umformungsmatrix, die die Addition des λ-fachen der j-ten Zeile zur i-ten Zeile induziert. Nach den Überlegungen im Beweis von 6.2.10 ist $U_{ij}(\lambda)$ dann eine Einheitsmatrix mit einem zusätzlichen Eintrag λ in der i-Zeile und j-ten Spalte. Weiter bezeichnen wir mit P_{kl} die Umformungsmatrix, die die Vertauschung der k-ten und der l-ten Zeilen induziert.

Lemma 6.3.10 *Seien $i < k < l$. Dann gelten*

$$\begin{aligned}
P_{kl}U_{ij}(\lambda) &= U_{ij}(\lambda)P_{kl} \qquad \textit{für } j \notin \{i,k,l\}, \\
P_{kl}U_{ik}(\lambda) &= U_{il}(\lambda)P_{kl}, \\
P_{kl}U_{il}(\lambda) &= U_{ik}(\lambda)P_{kl}.
\end{aligned}$$

Beweis. Die letzte Gleichung folgt aus der vorletzten Gleichung, weil die Vertauschung der k-ten mit der l-ten Zeile zu sich selbst invers ist, also $P_{kl}P_{kl} = E_n$ gilt. Die erste Gleichung gilt, weil die Addition des λ-fachen der

j-ten Zeile zur i-ten Zeile nicht beeinflußt wird durch die vorherige oder anschließende Vertauschung zweier anderer Zeilen. Da die Umformungsmatrizen durch die Zeilenoperationen eindeutig bestimmt sind - sie sind ja das Ergebnis der entsprechenden Umformungen der Einheitsmatrix -, gilt die Aussage auch für die entsprechenden Umformungsmatrizen. Wenn man jedoch die l-te Zeile mit der k-ten Zeile vertauscht und dann die jetzt k-te Zeile mit λ multipliziert und zur i-ten Zeile addiert, dann ist das dasselbe, wie die Addition des λ-fachen der l-ten Zeile zur i-ten Zeile und die anschließende Vertauschung der l-ten Zeile mit der k-ten Zeile. Damit folgt auch die zweite Gleichung.

Wir werden den Algorithmus $\mathcal{E}$ zum Übergang zur Stufenmatrix nochmals genauer studieren. Zu Beginn jedes Reduktionsschrittes führen wir mit $\mathcal{E}_3$ jeweils eine Vertauschung von Zeilen oder eine Multiplikation mit einer Matrix P_{kl}, $k < l$ durch. Dann erfolgen mehrere Multiplikationen mit Matrizen der Form $U_{kj}(\lambda)$ mit jeweils verschiedenen Faktoren λ. Beim nächsten Reduktionsschritt wird der Zeilenzähler k um 1 erhöht. Die Folge der Umformungen geschrieben mit den Umformungsmatrizen (und jeweils passenden Faktoren λ) kann dann so geschrieben werden

$$U_{kk+1}(\lambda) \ldots U_{km}(\lambda) P_{kl} U_{k-1,k}(\lambda) \ldots P_{k-1,l'} \ldots M = S.$$

Da die Matrix P_{kl} an den Matrizen $U_{ij}(\lambda)$ mit $i < k$ gemäß Lemma 6.3.10 vorbeigezogen werden kann, ohne die Indizes i zu ändern, kann man dieselben Umformungen auch in der Form

$$U_{kk+1}(\lambda) \ldots U_{1m}(\lambda) P_{kl} \ldots P_{1l'} M = S$$

geschrieben werden. Das Produkt der Umformungsmatrizen $U_i(\lambda)$ ergibt eine untere Dreiecksmatrix L, das Produkt der Matrizen P_i eine Permutationsmatrix P. Also erhält man $LPM = S$. Durch Multiplikation von links mit L^{-1} ergibt sich die Gleichung $PM = L^{-1}S$. Da auch die Inversen von Umformungsmatrizen wieder Umformungsmatrizen desselben Typs sind, erhält man

Satz 6.3.11 *Zu jeder Matrix M gibt eine Stufenmatrix S, eine untere reguläre Dreiecksmatrix L und eine Permutationsmatrix P, so daß gilt*

$$PM = LS.$$

Satz 6.3.12 *(über die LU-Zerlegung von regulären Matrizen) Ist M eine reguläre Matrix, so gibt es eine reguläre obere Dreiecksmatrix U, eine reguläre untere Dreiecksmatrix L und eine Permutationsmatrix P, so daß*

$$PM = LU.$$

Beweis. Für eine reguläre Matrix ergibt sich die Stufenform als reguläre obere Dreiecksmatrix. Daher ergibt sich die Folgerung aus dem Satz.

Algorithmus 6.3.13 (Algorithmus zur LU-Zerlegung) Wir wollen einen Algorithmus zur LU-Zerlegung angeben. Dazu betrachten wir die Matrix (M, E_n, E_n, E_n), auf die wir aus dem Gaußschen Algorithmus gewonnene elementare Zeilenoperationen und zusätzliche Matrizenmultiplikationen anwenden. Bei der Durchführung dieses Algorithmus parallel zum Gaußalgorithmus sei in einem Zwischenschritt daraus die Matrix (T, L, P, E_n) entstanden mit den zusätzlichen Eigenschaften $LPM = T$, L eine untere Dreiecksmatrix und P eine Permutationsmatrix.

- Wenn der Gaußsche Algorithmus nun eine Addition eines Vielfachen einer Zeile zu einer anderen Zeile erfordert, so ist dies durch eine Umformungsmatrix U gegeben. Diese wenden wir lediglich auf die Teilmatrix (T, L) an und erhalten (UT, UL, P, E_n) mit $ULPM = UT$ und UL untere Dreiecksmatrix, d.h. wir wenden die elementare Zeilenumformung nur auf (T, L) an.

- Wenn der Gaußsche Algorithmus eine Vertauschung von zwei Zeilen erfordert, so ist dies durch eine weitere Umformungsmatrix U gegeben. Die elementare Zeilenumformung wenden wir auf die ganze Matrix (T, L, P, E_n) an und erhalten(UT, UL, UP, U). Anschließend multiplizieren wir UL von rechts mit U, da beide Matrizen explizit zur Verfügung stehen, und ersetzen U wieder durch die Einheitsmatrix, haben danach also (UT, ULU, UP, E_n) mit $ULUUPM = ULPM = UT$, wegen $UU = E_n$. Da wie in den oben durchgeführten Überlegungen ULU nur für untere Dreiecksmatrizen durchgeführt wird, deren Einträge außerhalb der Diagonalen durch die Vertauschung zwar umgestellt werden, aber nicht aus dem unteren Dreieck heraus getauscht werden (6.3.10), bleibt ULU eine reguläre untere Dreiecksmatrix.

Das Verfahren bricht ab, sobald wir eine Stufenmatrix S anstelle von T erhalten haben, also (S, L, P, E_n) mit $LPM = S$, wobei L eine reguläre untere Dreiecksmatrix, P eine Permutationsmatrix und S eine Stufenmatrix sind. Da auch L^{-1} eine reguläre untere Dreiecksmatrix ist, ist mit $PM = L^{-1}S$ das Ziel erreicht.

Wir wollen die Anwendung der LU-Zerlegung diskutieren. Wenn das lineare Gleichungssystem $M \cdot x = b$ zu lösen ist und $PM = LS$ gilt, so folgt $LSx = PMx = Pb$. Das Gleichungssystem kann gelöst werden, indem man $c := Pb$ berechnet, das Gleichungssystem $Ly = c$ durch Vorwärtssubstitution wie in 6.2.7 mit

$$\begin{aligned} \eta_1 &= \gamma_1, \\ \eta_2 &= \gamma_2 - \lambda_{21}\eta_1 \\ &\vdots \\ \eta_n &= \gamma_n - \sum_{j=1}^{n-1} \lambda_{nj}\eta_j \end{aligned}$$

wobei $L = (\lambda_{ij})$, und schließlich die Lösung y in das Gleichungssystem $Sx = y$ einsetzt und dieses durch Rückwärtssubstitution löst.

Beispiel 6.3.14 Wir schließen diesen Abschnitt mit einem Beispiel für das angegebene Verfahren. Wir wollen das lineare Gleichungssystem

$$\begin{aligned} 2y + 3z &= 2 \\ 2u + x \qquad + 3z &= 1 \\ 3u + 3x + y + 6z &= 4 \end{aligned}$$

mit Hilfe des LU-Verfahrens lösen. Dazu betrachten wir die folgenden Matrizen und Matrixumformungen:

$$\begin{pmatrix} 0 & 0 & 2 & 3 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 3 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 3 & 3 & 1 & 6 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

wird durch Vertauschen der ersten und zweiten Zeile umgewandelt in

$$\begin{pmatrix} 2 & 1 & 0 & 3 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 3 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 3 & 3 & 1 & 6 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Wir multiplizieren die Teilmatrix UL mit U von rechts und ersetzen U durch E_n:

$$\begin{pmatrix} 2 & 1 & 0 & 3 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 3 & 3 & 1 & 6 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Die nächste Umformung erfolgt durch Addition des $(-\frac{3}{2})$-fachen der ersten Zeile zur dritten Zeile in der Teilmatrix (T, L):

$$\begin{pmatrix} 2 & 1 & 0 & 3 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1.5 & 1 & 1.5 & -1.5 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Schließlich erfolgt ein Vertauschen der zweiten mit der dritten Zeile

$$\begin{pmatrix} 2 & 1 & 0 & 3 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1.5 & 1 & 1.5 & -1.5 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 2 & 3 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

und eine Multiplikation der Teilmatrix UL mit U von rechts (und Ersetzen von U durch E_n):

$$\begin{pmatrix} 2 & 1 & 0 & 3 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1.5 & 1 & 1.5 & -1.5 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 2 & 3 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Wir invertieren nun die untere Dreiecksmatrix und erhalten in der obigen Notation

$$S = \begin{pmatrix} 2 & 1 & 0 & 3 \\ 0 & 1.5 & 1 & 1.5 \\ 0 & 0 & 2 & 3 \end{pmatrix}; \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 1.5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Wir berechnen $c = Pb$ als

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

lösen das Gleichungssystem $Ly = c$ und erhalten

$$y = \begin{pmatrix} 1 \\ 2.5 \\ 2 \end{pmatrix},$$

schließlich hat das Gleichungssystem $Sx = y$ die partikuläre Lösung

$$x_0 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

und eine Basis des Lösungsraumes des homogenen Gleichungssystems ist

$$a_1 = \begin{pmatrix} 1.5 \\ 0 \\ 1.5 \\ -1 \end{pmatrix}.$$

Damit ist auch die komplette Lösung des ursprünglichen Gleichungssystems gefunden.

Bisher hatten wir bei allen Algorithmen in der Wahl der Zeile, die durch Vertauschen an die oberste Stelle kommen soll, keine Festlegung getroffen. Da der Computer beim Rechnen mit Gleitkommazahlen jedoch ungenaue Resultate erzielt, hat sich die folgende Festlegung der zu wählenden Zeile als besonders günstig erwiesen. In jedem Schritt $\mathcal{E}_3$, den man im Gauß- oder Gauß-Jordan-Verfahren durchführt:

$\mathcal{E}_3'$ vertausche man die erste Zeile mit einer Zeile, die in der ersten Spalte einen dem Betrag nach möglichst großen Koeffizienten hat, also mit der i-ten Zeile, wenn $|\alpha_{i1}| \geq |\alpha_{j1}|$ für alle $j \neq i$. Einen solchen Koeffizienten α_{i1} nennt man ein *Pivot- oder Drehpunktelement*.

Übungen 6.3.15 1. Sei $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ die Abbildung

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} xy \\ x^2 + y^2 \end{pmatrix}$$

Die Matrix der partiellen Ableitungen ist

$$Df := \begin{pmatrix} y & x \\ 2x & 2y \end{pmatrix}.$$

Entscheiden Sie, wann Df invertierbar ist.

2. Entscheiden Sie, wann

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

invertierbar ist.

3. Finden Sie Inverse für die folgenden Matrizen

$$\begin{pmatrix} 2 & 4 & 6 \\ -1 & -4 & -3 \\ 0 & 1 & -1 \end{pmatrix}, \begin{pmatrix} 2 & 2 & -4 \\ 2 & 6 & 0 \\ -3 & -3 & 5 \end{pmatrix},$$

$$\begin{pmatrix} 1 & -2 & 3 & 0 \\ 0 & 1 & -1 & 1 \\ -2 & 2 & -2 & 4 \\ 0 & 2 & -3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 & 2 \\ 2 & -1 & 1 & -1 \\ 3 & 3 & 2 & -2 \\ 1 & 2 & 1 & 0 \end{pmatrix}.$$

4. Lösen Sie die folgenden Gleichungssysteme unter Verwendung der inversen Koeffizientenmatrizen

a)

$$\begin{aligned} x - 2y + 3z \phantom{{}+ 4w} &= 4 \\ y - z + w &= -8 \\ -2x + 2y - 2z + 4w &= 12 \\ 2y - 3z + w &= -4 \end{aligned}$$

b)

$$\begin{aligned} x + y \phantom{{}+ 2z} + 2w &= 3 \\ 2x - y + z - w &= 3 \\ 3x + 3y + 2z - 2w &= 5 \\ x + 2y + z \phantom{{}- 2w} &= 3 \end{aligned}$$

5. Ein Bankkunde möchte Geld in festverzinslichen Papieren der Kategorien AAA, A und B anlegen. Die Papiere der Kategorie AAA erbringen 6% Zinsen, die der Kategorie A 7% Zinsen und die der Kategorie B 10% Zinsen. Der Kunde möchte doppelt soviel Geld in der Kategorie AAA anlegen, als in der Kategorie B. Wieviel Geld muß der Kunde in den einzelnen Kategorien anlegen, wenn

a) seine Gesamtanlage DM 50000.- beträgt und die jährliche Zinseinnahme DM 3620.- betragen soll,

b) seine Gesamtanlage DM 60000.- beträgt und die jährliche Zinseinnahme DM 4300.- betragen soll,

c) seine Gesamtanlage DM 80000.- beträgt und die jährliche Zinseinnahme DM 5800.- betragen soll?

6. Finden Sie LU-Zerlegungen von den Matrizen

$$M_1 = \begin{pmatrix} 2 & 6 & 2 \\ -3 & -8 & 0 \\ 4 & 9 & 2 \end{pmatrix}$$

$$M_2 = \begin{pmatrix} 6 & -2 & 0 \\ 9 & -1 & 1 \\ 3 & 7 & 5 \end{pmatrix}.$$

7. Lösen Sie mittels des Verfahrens der LU-Zerlegung die linearen Gleichungssysteme

$$\begin{pmatrix} 0 & 1 & 0 & 2 \\ -2 & -1 & -1 & 4 \\ 2 & 3 & 1 & 0 \\ 2 & 2 & 2 & -1 \end{pmatrix} \cdot x = b_i$$

für

$$b_1 = \begin{pmatrix} 2 \\ 3 \\ 1 \\ 0 \end{pmatrix}, b_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 2 \end{pmatrix} \quad \text{und} \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix}.$$

6.4 Ein Kapitel Codierungstheorie

Wir wollen die Erkenntnisse über lineare Abbildungen und Matrizen verwenden, um Probleme der linearen Codierung zu formulieren und zu lösen. Dazu legen wir uns ein Wörterbuch an, das die Bedeutung gewisser Objekte aus der linearen Algebra in Ausdrücke der Codierungstheorie übersetzt.

Wir werden in diesem Abschnitt als Grundkörper durchgehend den endlichen Körper $K := GF(q)$ mit q Elementen verwenden. GF ist dabei eine Abkürzung für das Wort Galois-Feld. Man kann zeigen, daß es genau dann einen Körper mit q Elementen gibt, wenn q ein Primzahlpotenz ist, d.h. wenn es eine Primzahl p und eine natürliche Zahl n mit $q = p^n$ gibt. Dieser Körper ist zudem durch die Angabe von q bis auf Isomorphie eindeutig bestimmt.

Wir kennen bisher lediglich die endlichen Körper $GF(p) = \mathbb{Z}/p\mathbb{Z}$ für Primzahlen p. Da die meisten Anwendungen jedoch nur den Körper $\mathbb{Z}/2\mathbb{Z}$ mit zwei Elementen (binäres System der Computer!) benutzen, wollen wir die Konstruktion der übrigen Körper $GF(q)$ hier nicht durchführen und verweisen den interessierten Leser auf Lehrbücher der Algebra.

Definition 6.4.1 (Wörterbuch der Codierung) Ein *Code* ist eine Menge C (von Zeichen, die geeignet sind, Informationen zu speichern und zu übermitteln). Eine *Chiffre* oder eine *Verschlüsselung* (*Codierung* oder *Chiffrierung*) ist eine Abbildung $f : C_1 \longrightarrow C_2$ eines Codes in einen anderen. Eine *Dechiffrierung* einer Chiffre $f : C_1 \longrightarrow C_2$ ist eine Abbildung $g : C_2 \longrightarrow C_1$ mit $gf = \mathrm{id}$. Die Quelle einer Chiffre $f : C_1 \longrightarrow C_2$ heißt *Klartext*, ein Element des Klartextes heißt *Nachrichtenwort*. Ein Element des Bildes einer Chiffre heißt *Codewort*. Eine Codierung heißt *lineare Codierung*, wenn C_1 und C_2 Vektorräume sind und $f : C_1 \longrightarrow C_2$ eine lineare Abbildung ist. Sinnvoll sind nur Codierungen f, die injektiv sind.

Beispiele 6.4.2 1. Sprachen im Sinne von 3.1.6, d.h. beliebige Mengen von „strings" oder Wörtern über einem beliebigen Alphabet A.

2. Das Zahlensystem, d.h. die mit den Ziffern $0, \ldots, 9$ und den Zeichen . und $-$ dargestellten Zahlen.

3. Das Morsealphabet, das mit den Zeichen . (dit) und $-$ (dah) aufgebaut wird.

4. Die q- und die z-Gruppen in der Morsesprache, das sind Gruppen von drei Buchstaben des (Buchstaben-)Alphabets, die mit q bzw. z beginnen, z.B. qth = Standort.

5. Die Barcodes zur Bezeichnung von Waren im Supermarkt.

6. Der ISBN-Code (International Standard Book Number), wie z.B. 3-519-02211-7, wobei die einzelnen Gruppen folgendes bedeuten:

3 = Erscheinungsland
519 = Verlag
02211 = fortlaufende Buchnummer
7 = Prüfnummer.

Die Prüfung auf eine korrekte Übertragung (Fehlererkennung) geschieht im Beispiel durch Überprüfung von $10 \cdot 3 + 9 \cdot 5 + 8 \cdot 1 + 7 \cdot 9 + 6 \cdot 0 + 5 \cdot 2 + 4 \cdot 2 + 3 \cdot 1 + 2 \cdot 1 + 1 \cdot 7 \equiv 0 (\mathrm{mod}\ 11)$. Die Restklassenberechnung modulo 11 kann wie in Beispiel 3.6.4 5. durch Bildung der alternierenden Quersumme vorgenommen werden.

7. Beliebiger Text der Umgangssprache kann in einen linearen Code K^n für $K = GF(q)$ übersetzt werden, indem man zunächst den Text jeweils in Gruppen von l Buchstaben und Abstände (und evtl. sonstige Zeichen) zusammenfaßt. Bei der Verwendung von $a, \ldots, z, A, \ldots, Z$,*Zwischenraum* sind also 53^l verschiedene solche Textgruppen möglich. Diesen weist man in einer beliebig festzulegenden Weise ebenso viele verschiedene Elemente in K^n zu. Damit bestimmt sich l aus $53^l \leq q^n$ als $l \leq n \cdot \frac{\ln(q)}{\ln(53)}$.

Wir werden im folgenden nur lineare Codes der Form $C = K^n$ mit $K = GF(q)$ verwenden mit der linearen Chiffrierung $f : K^k \longrightarrow K^n$. Wegen der notwendigen Dechiffrierung wird f immer als Monomorphismus vorausgesetzt.

Definition 6.4.3 Sei K^n ein linearer Code. Die *Hamming*[1]*-Metrik* auf K^n ist die Abbildung $d : K^n \times K^n \longrightarrow \mathbb{N}_0$ mit $d(x, y) :=$ Anzahl der $i \in \{1, \dots, n\}$ mit $\xi_i \neq \eta_i$. Die Auffassung ist hierbei, daß $d(x, y)$ die Anzahl der Koeffizienten von x angibt, die in y anders (falsch) angegeben werden. Der Wert $d(x, y)$ heißt *Hamming-Abstand* von x und y. Die *Hamming-Gewichtsfunktion* ist die Abbildung $\|.\| : K^n \longrightarrow \mathbb{N}_0$ mit $\|x\| := d(x, 0)$. Das ist die Anzahl der von Null verschiedenen Komponenten von x.

Definition 6.4.4 Ein Paar (M, d) heißt *metrischer Raum mit der Metrik* d, wenn M eine Menge und $d : M \times M \longrightarrow \mathbb{R}$ eine Abbildung sind mit

1. $\forall x, y \in M[d(x, y) = 0 \Longleftrightarrow x = y]$,
2. $\forall x, y \in M[d(x, y) = d(y, x)]$,
3. $\forall x, y, z \in M[d(x, z) \leq d(x, y) + d(y, z)]$ (Dreiecksungleichung).

Lemma 6.4.5 *Die Hamming-Metrik ist eine Metrik auf* K^n.

Beweis. folgt unmittelbar aus der Definition.

Wir vermerken noch eine leicht einzusehende zusätzliche Translationsinvarianz $d(x, y) = d(x+z, y+z)$, die zeigt, daß d durch das Hamming-Gewicht $\|.\|$ schon vollständig bestimmt ist, denn $d(x, y) = \|x - y\|$.

Ein weiteres bekanntes Beispiel für eine Metrik ist der reelle Vektorraum $\mathbb{R}^n$ mit der sogenannten Euklidischen Metrik (vgl. Kapitel 8)

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (\xi_i - \eta_i)^2}.$$

Lemma 6.4.6 *Für das Hamming-Gewicht* $\|.\| : K^n \longrightarrow \mathbb{N}_0$ *gelten*

1. $\|x\| = 0 \Longleftrightarrow x = 0$,
2. $\forall \lambda \neq 0[\|\lambda x\| = \|x\|]$,
3. $\|x + y\| \leq \|x\| + \|y\|$.

Beweis. leicht nachzurechnen, da $\|x\|$ die Anzahl der von Null verschiedenen Komponenten von x ist.

Definition 6.4.7 Ein Monomorphismus $f : K^k \longrightarrow K^n$ wird eine (k, n) - *lineare Codierung* genannt. Die darstellende Matrix von f heißt *erzeugende Matrix*. Eine lineare Abbildung $h : K^n \longrightarrow K^{n-k}$ heißt *Kontrollabbildung*, wenn $\mathrm{Ke}(h) = \mathrm{Bi}(f)$. Die darstellende Matrix von h heißt *Kontrollmatrix*. Die *Hamming-Norm* oder der *Hamming-Abstand* einer (k, n)-linearen Codierung f ist definiert als

$$\|f\| := \mathrm{Min}\left\{\|f(x)\| \, | x \in K^k, x \neq 0\right\}.$$

[1] Richard W. Hamming (1915–1998)

Wir bemerken, daß $x \in \mathrm{Bi}(f)$ genau dann, wenn $h(x) = 0$. Ein solche Kontrollabbildung existiert immer, wie wir z.B. in 6.1.4 gesehen haben. Weiter ist $\|f\| = \mathrm{Min}\{d(f(x), f(y)) | x, y \in K^k, x \neq y\}$, weil $d(f(x), f(y)) = d(f(x) - f(y), 0) = d(f(x - y), 0) = \|f(x - y)\|$.

Im folgenden Satz gehen wir von der allgemeinen Vorstellung aus:

$$\boxed{\text{Klartext}} \xrightarrow{\text{Codierung}} \boxed{\text{Sender}} \xrightarrow{\text{Übertragung/Störung}}$$

$$\boxed{\text{Empfänger}} \xrightarrow{\text{Decodierung}} \boxed{\text{Klartext}}$$

Es wird also ein Klartext codiert, über eine Informationsleitung zum Empfänger übermittelt, (wobei der Text aufgrund der Codierung eventuell auch abhörsicher ist,) wird auf der Übertragungsstrecke mit Störungen verschiedener Art verändert und beim Empfänger wieder decodiert. Wir wollen Methoden finden, die Fehler bei der Übertragung zu erkennen und möglichst auch zu korrigieren. Wenn also x ein codiertes ausgesandtes Wort ist und y das empfangene Wort ist, dann soll festgestellt werden, ob es tatsächlich durch die Codierung entstanden ist oder verändert wurde und ob man daraus das Wort x rekonstruieren kann. Wenn bei einer Codierung $f : K^k \longrightarrow K^n$ Fehler an höchstens r Stellen des übertragenen Wortes immer erkannt werden können, so sagen wir, daß die Codierung r-fehlerentdeckend ist. Wenn Fehler an höchstens s Stellen durch die restliche Information im übertragenen Wort korrigiert werden können, so heißt die Codierung s-fehlerkorrigierend.

Satz 6.4.8 *Sei $f : K^k \longrightarrow K^n$ eine (k, n)-lineare Codierung, sei $C := \mathrm{Bi}(f)$ und sei $y \in K^n$.*

1. *(Fehlererkennung:) Wenn es ein $x \in C$ mit $x \neq y$ gibt, so daß $d(x, y) < \|f\|$, dann ist $y \notin C$, d.h. das Wort y ist kein Codewort, also falsch.*
2. *(Fehlerkorrektur:) Wenn es ein $x \in C$ gibt mit $d(x, y) < \frac{1}{2}\|f\|$, dann gilt für alle $z \in C, z \neq x [d(x, y) < d(z, y)]$, d.h. x ist das einzige Element von C mit dem gegebenen Abstand $d(x, y)$ und somit eindeutig durch y bestimmt.*

Beweis. 1. Wenn $y \in C$ wäre, so wäre $d(x, y) \geq \|f\|$ oder $x = y$ nach Definition von $\|f\|$.

2. Für $z \in C$ und $z \neq x$ gilt $2d(x, y) < \|f\| \leq d(x, z) \leq d(x, y) + d(y, z)$, also $d(x, y) < d(z, y)$.

Wenn bei der Übertragung von $f(a)$ weniger als $\frac{1}{2}\|f\|$ Fehler aufgetreten sind und $y \in K^n$ empfangen wurde, dann ist also $x \in C$ mit $d(x, y) < \frac{1}{2}\|f\|$ das übertragene Element $f(a)$. Somit ist eine lineare Codierung immer $\|f\| - 1$-fehlererkennend und $[\frac{1}{2}(\|f\| - 1)]$-fehlerkorrigierend. Es kommt also jetzt darauf an, Codierungen f mit möglichst großer Norm $\|f\|$ zu finden. Wir betrachten einige

Beispiele 6.4.9 1. Paritäts-Prüfungs-Codes (Parity-Check-Codes): Sei $n \geq 2$ und $k = n-1$ und $f : K^k \longrightarrow K^n$ gegeben durch $f(\alpha_1, ..., \alpha_{n-1}) = (\alpha_1, ..., \alpha_n)$ mit $\alpha_n = -(\alpha_1 + \ldots + \alpha_{n-1})$. Offenbar ist f ein Monomorphismus. Weiter ist $a \in \mathrm{Bi}(f)$ genau dann, wenn $\sum_{i=1}^n \alpha_i = 0$. Wenn $q = 2$ ist, dann wird jede ungerade Anzahl von Fehlern dadurch erkannt, daß $\sum_{i=1}^n \alpha_i = 1$ gilt. Ein gerade Anzahl von Fehlern wird nicht erkannt. Für $x \in \mathrm{Bi}(f)$ und $x \neq 0$ müssen mindestens zwei Koeffizienten von Null verschieden sein, also ist $\|f\| = 2$. Damit kann zwar ein Fehler (und sogar eine ungerade Anzahl von Fehlern) erkannt werden, jedoch ergibt sich keine Möglichkeit zur Korrektur von Fehlern. Diese Codierungen werden z.B. in PCs verwendet, wenn 8-Bit Worte in 9-Bit Speichern gespeichert werden und das 9. Bit durch die Abbildung f bestimmt wird.

2. Wiederholungscode: Eine einfache Möglichkeit einer sichereren Übertragung auf einer gestörten Übertragungsstrecke ist die dreifache Übertragung jedes einzelnen Wortes. Dabei ist $n = 3k$ und $f : K^k \longrightarrow K^n$ durch $f(x) = (x, x, x)$ gegeben. Das ist wieder eine lineare Codierung. Man sieht sofort, daß $\|f\| = 3$ ist, also ist nach 6.4.8 diese Codierung 2-fehlerentdeckend und 1-fehlerkorrigierend.

Wenn die beiden Vektorräume K^k und K^n dieselbe Dimension haben, dann muß die Codierung f ein Isomorphismus sein. Dann ist $\|f\| = 1$ und eine Fehlerentdeckung oder -Korrektur offenbar nicht möglich. $\|f\|$ hängt also offenbar auch von den gegebenen Dimensionen ab.

Satz 6.4.10 *Sei* $f : K^k \longrightarrow K^n$ *eine* (k, n)-lineare Codierung *und* $h : K^n \longrightarrow K^{n-k}$ *eine Kontrollabbildung. Sei* B *die zugehörige Kontrollmatrix. Dann sind je* $\|f\| - 1$ *Spaltenvektoren von* B *linear unabhängig, und es gibt* $\|f\|$ *linear abhängige Spaltenvektoren, d.h.* $\|f\|$ *ist die Minimalzahl von linear abhängigen Spaltenvektoren von* B.

Beweis. Sei $x = \sum \lambda_i e_i \in \mathrm{Bi}(f)$ ein von Null verschiedener Vektor mit $\|x\| = \|f\|$ minimal. Dann sind genau $\|f\|$ Faktoren λ_i von Null verschieden. Wegen $0 = h(x) = B \cdot x = \sum \lambda_i b_i$ sind die $\|f\|$ Spaltenvektoren b_i linear abhängig. Sei eine Teilmenge $I \subseteq \{1, \ldots, n\}$ mit $|I| = \|f\| - 1$ gegeben und sei $\sum_{i \in I} \lambda_i b_i = 0$ mit $\lambda_i \neq 0$ für alle $i \in I$. Dann ist $h(\sum_{i \in I} \lambda_i e_i) = \sum_{i \in I} \lambda_i b_i = 0$, also ist $x = \sum_{i \in I} \lambda_i e_i \in \mathrm{Bi}(f)$. Dann sind mindestens $\|f\|$ verschiedene Skalare $\lambda_i \neq 0$, was nach Voraussetzung ausgeschlossen ist, oder es ist $x = 0$, und damit sind alle $\lambda_i = 0$. Die Menge der $\{b_i | i \in I\}$ ist also linear unabhängig.

Folgerung 6.4.11 *Für jede* (k, n)*-lineare Codierung* $f : K^k \longrightarrow K^n$ *gilt*

$$\|f\| \leq n - k + 1.$$

Beweis. Der Rang jeder Kontrollmatrix zu f ist $n - k$. Damit sind je $n - k + 1$ Vektoren der Kontrollmatrix linear abhängig. Nach dem vorhergehenden Satz ist also $\|f\| \leq n - k + 1$.

Die Kontrollmatrix kann herangezogen werden, um Fehler in der Übertragung zu erkennen und evtl. zu korrigieren. Insbesondere kann mit ihr die Anzahl der Fehler abgeschätzt werden. Wir bezeichnen die Spaltenvektoren von B mit $b_i \in K^{n-k}$, also $B = (b_1, \ldots, b_n)$.

Folgerung 6.4.12 *1. Sei $x \in \mathrm{Bi}(f)$, und seien bei der Übertragung genau t Fehler aufgetreten. Wenn $y \in K^n$ der empfangene Wert ist, dann ist die minimale Anzahl der Koeffizienten λ_i mit $By = \sum \lambda_i b_i$ höchstens t.*

2. Sei $x \in \mathrm{Bi}(f)$, seien bei der Übertragung genau t Fehler aufgetreten, und sei $t < \frac{1}{2}\|f\|$. Wenn $y \in K^n$ der empfangene Wert ist, dann gibt es t eindeutig bestimmte Koeffizienten λ_i mit $By = \sum \lambda_i b_i$. Der Übertragungsfehler ist dann $\sum \lambda_i e_i$, und es gilt $x = y - \sum \lambda_i e_i$.

Beweis. Mit 1. kann die Anzahl der aufgetretenen Fehler nach unten abgeschätzt werden. Sei $\sum \mu_i e_i$ der Übertragungsfehler, d.h. $y = x + \sum \mu_i e_i$. Dann ist $By = Bx + \sum \mu_i B e_i = \sum \mu_i b_i$. Seien genau t Fehler aufgetreten, so ist die Anzahl der Summanden t. Wenn in 2. weniger als $\frac{1}{2}\|f\|$ Fehler aufgetreten sind, dann ist die Anzahl der Summanden in der Darstellung $By = \sum \mu_i b_i$ kleiner als $\frac{1}{2}\|f\|$, also sind die verwendeten b_i linear unabhängig. Wenn $By = \sum \lambda_j b_j$ eine weitere Darstellung ist und die Anzahl der Summanden minimal, insbesondere also kleiner als $\frac{1}{2}\|f\|$, ist, dann ist $\sum \mu_i b_i - \sum \lambda_j b_j = 0$ mit weniger als $\|f\|$ Summanden. Nach 6.4.10 sind die verwendeten b_i linear unabhängig, also stimmen die λ_i mit den μ_i überein, d.h. die Darstellung $By = \sum \mu_i b_i$ mit weniger als $\frac{1}{2}\|f\|$ Summanden ist eindeutig, und der Fehler ist $y - x = \sum \lambda_i e_i$. Man beachte hier jedoch, daß der Fehler nur unter der Annahme $t < \frac{1}{2}\|f\|$ korrigiert werden konnte.

Man kann die Minimalzahl von linear abhängigen Vektoren in B und damit die Hamming-Norm von f bestimmen, daher ist es sinnvoll eine beliebige Kontrollmatrix B' zu konstruieren und dann aus ihr die Codierung $f : K^r \longrightarrow K^n$ abzuleiten. Man kann dann B unmittelbar mit einem möglichst großen Minimum an linear abhängigen Vektoren konstruieren. Dazu brauchen wir lediglich einen Monomorphismus $f : K^r \longrightarrow K^n$ mit $\mathrm{Bi}(f) = \mathrm{Ke}(\widehat{B} : K^n \longrightarrow K^{n-r})$ zu konstruieren, was wegen $r = \dim \mathrm{Ke}(\widehat{B})$ immer möglich ist.

Beispiele 6.4.13 1. Sei $K = \mathbb{Z}/11\mathbb{Z}$. Sei die Kontrollmatrix

$$B := \begin{pmatrix} 2 & 1 & 2 & 1 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 1 & 2 & 1 & 2 & 1 \end{pmatrix}.$$

Dann kann man nachrechnen, daß die Minimalzahl von linear abhängigen Spaltenvektoren 4 ist. Eine entsprechende Codierungsmatrix für die Codierung $f : K^4 \longrightarrow K^7$ gewinnt man aus der Basis des Kerns von $\widehat{B}$. Sie ist

$$A := \begin{pmatrix} 9 & 9 & 8 & 7 \\ 1 & 2 & 2 & 5 \\ 2 & 1 & 2 & 3 \\ 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}.$$

Diese Codierung hat die Hamming-Norm 4, kann also 3 Fehler erkennen und einen Fehler korrigieren.

2. Sei $K = \mathbb{Z}/2\mathbb{Z}$. Sei die Kontrollmatrix

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

Dann kann man nachrechnen, daß die Minimalzahl von linear abhängigen Spaltenvektoren 3 ist. Eine entsprechende Codierungsmatrix für die Codierung $f : K^4 \longrightarrow K^7$ ist

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

Diese Codierung hat die Hamming-Norm 3, kann also 2 Fehler erkennen und einen Fehler korrigieren.

Definition 6.4.14 Sei K ein beliebiger Körper. Die Menge der Folgen $K[[x]] := \{\varphi : \mathbb{N}_0 \longrightarrow K\} = \{(\alpha_0, \alpha_1, \ldots) | \alpha_i \in K\} = K^{\mathbb{N}_0}$ heißt (formaler) *Potenzreihenring* über K.

Lemma 6.4.15 $K[[x]]$ *ist ein Ring unter den folgenden Operationen:*

$(\varphi + \psi)(n) = \varphi(n) + \psi(n)$,
$(\varphi \cdot \psi)(n) = \sum_{i=0}^{n} \varphi(i)\psi(n-i)$.

Das Einselement ist die Folge $(1, 0, 0, \ldots)$.

Beweis. Unter der Addition liegt sogar ein Vektorraum vor nach dem Hauptbeispiel für Vektorräume 5.1.4. Die Assoziativität und Distributivität der Multiplikation ist eine einfache Rechenübung. Die Eigenschaft des Einselements folgt aus der Tatsache, daß bei der Multiplikation die Summe jeweils auf einen einzigen Summanden zusammenfällt.

Definition 6.4.16 Sei K ein beliebiger Körper. Die Menge der endlichwertigen Folgen

$$K[x] = K^{(\mathbb{N}_0)} := \{(\alpha_i) \in K^{\mathbb{N}_0} \mid \text{für nur endlich viele } i \in \mathbb{N}_0 \text{ gilt } \alpha_i \neq 0\}$$

heißt (formaler) *Polynomring* über K. Die Elemente von $K[x]$ heißen *Polynome.*

Lemma 6.4.17 *$K[x]$ ist ein Unterring von $K[[x]]$.*

Beweis. Nach Beispiel 5.1.7 ist $K[x] \subseteq K[[x]]$ ein Untervektorraum. Es bleibt nur die Abgeschlossenheit bezüglich der Multiplikation zu zeigen. Wenn also (α_i) und (β_i) in $K[x]$ gegeben sind, die beide nur noch Koeffizienten Null für Indizes $> n$ haben, dann sind in $(\alpha_i) \cdot (\beta_i) = (\sum_{j=0}^{i} \alpha_j \beta_{i-j})$ alle Koeffizienten mit Index $i > 2n$ Null, denn in der Summe sind alle Summanden Null.

Bemerkung 6.4.18 In $K[x]$ bezeichnen wir $x := (0, 1, 0, 0, 0, ...)$. Dann ist $x^2 = (0, 0, 1, 0, 0, \ldots)$. Allgemein ist $x^n = e_n$ die Folge mit einer Eins an der $n + 1$-sten Stelle und Null sonst. In 5.2.4 (4) haben wir gezeigt, daß die Menge der e_i bzw. hier die Menge der x^i eine Basis für $K[x]$ bilden. Jeder Vektor aus $K[x]$ läßt sich daher in eindeutiger Weise (mit eindeutig bestimmten Koeffizienten) als $\sum_{i=0}^{n} \alpha_i x^i = \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n$ schreiben. Die oben angegebene Multiplikation ist dann die bekannte Multiplikation von Polynomen. Die eindeutig bestimmte Zahl n mit $\alpha_n \neq 0$ und $\alpha_{n+i} = 0$ für alle $i \in \mathbb{N}$ heißt der *Grad* des Polynoms $(\alpha_i) \in K[x] \setminus \{0\}$. α_n heißt der höchste Koeffizient des Polynoms. Man sieht durch Betrachtung der höchsten Koeffizienten sofort ein, daß $K[x]$ ein nullteilerfreier Ring ist.

Lemma 6.4.19 *Die Polynome in $K[x]$ vom Grade höchstens n bilden einen Vektorraum P_n der Dimension $n + 1$.*

Beweis. Diese Polynome werden von den linear unabhängigen Polynomen $1 = x^0, x, x^2, x^3, \ldots, x^n$ erzeugt.

Satz 6.4.20 *Im Polynomring $K[x]$ gilt der Euklidische Divisionsalgorithmus: zu jedem Paar $f, g \in K[x]$ von Polynomen mit $g \neq 0$ gibt es ein eindeutig bestimmtes Paar von Polynomen $q, r \in K[x]$ (Quotient und Rest), so daß gilt*

$$f = q \cdot g + r \qquad \textit{und} \qquad \mathrm{Grad}(r) < \mathrm{Grad}(g).$$

Beweis. Wir zeigen zunächst die Eindeutigkeit der Zerlegung. Sei $f = qg + r = q'g + r'$ mit $\mathrm{Grad}(r) < \mathrm{Grad}(g)$ und $\mathrm{Grad}(r') < \mathrm{Grad}(g)$. Dann ist $(q - q')g + (r - r') = 0$. Da auch $\mathrm{Grad}(r - r') < \mathrm{Grad}(g)$ gilt, ist $(q - q')g = 0$. Insbesondere muß der höchste Koeffizient von $(q - q')g$ Null sein, was nur geht, wenn $q - q' = 0$ gilt. Dann ist aber auch $r - r' = 0$ und damit die Eindeutigkeit gezeigt.

Wenn der Grad von f kleiner ist, als der Grad von g, dann setzen wir $q = 0$ und $r = f$. Wenn $\mathrm{Grad}(f) = n + 1$ ist und der Satz für Polynome vom Grad n schon bewiesen ist, dann sei $\gamma := \alpha_{n+1}/\beta_k$ der Quotient der höchsten Koeffizienten α_{n+1} von f und β_k von g. Dann ist $f' := f - \gamma g$ ein Polynom vom Grad kleiner oder gleich n. Wir können also schreiben $f' = q'g + r$ mit $\mathrm{Grad}(r) < \mathrm{Grad}(g)$. Also ist $f = (\alpha + q')g + r$ mit $\mathrm{Grad}(r) < \mathrm{Grad}(g)$.

Bemerkung 6.4.21 1. Ein Polynom $f(x)$ in $K[x]$ vom Grad n hat höchstens n Nullstellen in K. Seien nämlich $\alpha_1, \ldots, \alpha_k$ Nullstellen von $f(x)$, dann ist $f(x) = (x - \alpha_1) \cdot (x - \alpha_2) \cdot \ldots \cdot (x - \alpha_k) \cdot g(x)$, also $k \leq n$. Nach dem Divisionsalgorithmus ist nämlich $f(x) = (x - \alpha_1) \cdot g_1(x) + \beta_1$. Wenn man für x den Wert α_1 einsetzt, dann erhält man $0 = \beta_1$. Für jede weitere Nullstelle α_i von $f(x)$ ist dann aber $0 = f(\alpha_i) = (\alpha_i - \alpha_1)g(\alpha_i)$, also sind die $\alpha_2, \ldots, \alpha_k$ Nullstellen von $g_1(x)$. Durch Induktion nach dem Grad erhält man die behauptete Aussage.

2. Von Polynomen in $K[x]$ können wir wie im reellen Fall Ableitungen bilden, hier *formale Ableitungen* genannt. Wir bilden nämlich die eindeutig bestimmte lineare Abbildung $d/dx : K[x] \longrightarrow K[x]$, indem wir auf der Basis (x^i) vorschreiben $d/dx(x^i) := ix^{i-1}$ (für $i = 0$ soll $d/dx(x^0) = 0$ gelten). Die Produktregel gilt auch hierfür, denn es ist $d/dx(x^i x^j) = (i + j)x^{i+j-1} = ix^{i-1}x^j + jx^i x^{j-1} = d/dx(x^i)x^j + x^i d/dx(x^j)$. Daraus leitet sich wegen der Linearität die Produktregel ab:

$$d/dx(fg) = d/dx(f)g + fd/dx(g).$$

Seien k und n mit $k < n$ gegeben und sei g ein Polynom vom Grad $n-k$. Dann definiert g die folgende lineare Abbildung $g : P_{k-1} \ni f \mapsto gf \in P_{n-1}$. Wir betrachten die entsprechende lineare Abbildung auf den Koordinatensystemen $\widehat{g} : K^r \longrightarrow K^n$. Wenn $g = \sum_{i=0}^{n-k} \gamma_i x^i$ ist, dann ist die darstellende Matrix von g bezüglich der Basen $1, x, x^2, \ldots, x^k$ bzw. $1, x, x^2, \ldots, x^n$ gegeben durch

$$M = \begin{pmatrix}
\gamma_0 & 0 & & \ldots & 0 \\
\gamma_1 & \gamma_0 & 0 & \ldots & 0 \\
\vdots & \ddots & \ddots & & \vdots \\
\vdots & & \ddots & \ddots & \vdots \\
\vdots & & & \gamma_1 & \gamma_0 \\
\vdots & & & & \gamma_1 \\
\gamma_{n-k} & & & & \vdots \\
0 & \gamma_{n-k} & & & \vdots \\
\vdots & & \ddots & & \vdots \\
\vdots & & & \ddots & \vdots \\
0 & \ldots & & 0 & \gamma_{n-k}
\end{pmatrix},$$

wie man sofort aus dem Polynomprodukt

$$gf = \sum_{i=0}^{n-k} \gamma_i x^i \sum_{j=0}^{k-1} \beta_j x^j = \sum_{t=0}^{n-1} (\sum_{j=0}^{k-1} \gamma_{t-j} \beta_j) x^t$$

abliest.

Definition 6.4.22 Eine Codierung der Form $\widehat{g} : K^k \longrightarrow K^n$ mit $\mathrm{Grad}(g) \leq n-k$ heißt *Polynomcode.*

Definition 6.4.23 Ein Codierung $h : K^k \longrightarrow K^n$ heißt *zyklisch*, wenn für alle $x = (\xi_1, \ldots, \xi_n) \in \mathrm{Bi}(h)$ gilt $(\xi_n, \xi_1, \ldots, \xi_{n-1}) \in \mathrm{Bi}(h)$. Also ist jede zyklische Vertauschung eines Codewortes wieder ein Codewort.

Satz 6.4.24 *Seien k und n mit $k < n$ gegeben. Sei $g \in P_{n-k}$ vom Grad $n-k$ ein Teiler von $x^n - 1 \in K[x]$. Dann ist der durch g erzeugte Polynomcode ein zyklischer Code. g heißt dann ein* Generatorpolynom *für den zyklischen Code.*

Beweis. Sei $gg' = x^n - 1$. Dann ist $x^n = gg' + 1$. Sei $gf = \alpha_0 + \alpha_1 x + \ldots + \alpha_{n-1}x^{n-1}$ Darstellung eines Codewortes $(\alpha_0, ..., \alpha_{n-1})$. Wir multiplizieren diese Gleichung mit x und erhalten $xgf = \alpha_0 x + \alpha_1 x^2 + \ldots + \alpha_{n-1}x^n = \alpha_0 x + \alpha_1 x^2 + \ldots + \alpha_{n-1}(gg' + 1) = \alpha_{n-1} + \alpha_0 x + \alpha_1 x^2 + \ldots + \alpha_{n-1}gg'$ und daraus $\alpha_{n-1} + \alpha_0 x + \alpha_1 x^2 + \ldots + \alpha_{n-2}x^{n-1} = xgf - \alpha_{n-1}gg' = g(xf - \alpha_{n-1}g')$. Da $\mathrm{Grad}(xf - \alpha_{n-1}g') \leq n - 1 - \mathrm{Grad}(g) = k - 1$ gilt, ist also auch $(\alpha_{n-1}, \alpha_0, \ldots, \alpha_{n-2})$ ein Codewort und der Polynomcode zyklisch.

Bemerkung 6.4.25 Es gilt auch die Umkehrung des Satzes. Sei $f : K^r \longrightarrow K^n$ ein zyklischer Code. Dann gibt es ein Polynom $g \in P_{n-k}$ mit g teilt $x^n - 1$, das diesen zyklischen Code erzeugt. Wir benötigen den Beweis hier nicht.

Sei $K = GF(q)$ mit $q = p^t$ und p einer Primzahl. In K gilt $p = 1 + \ldots + 1 \equiv 0 (\mathrm{mod}\ p)$. Dann ist $(x-1)^p = x^p - 1$ in $K[x]$, denn nach der binomischen Formel ist $(x-1)^p = \sum_{i=0}^{p} \binom{p}{i} (-1)^{p-i} x^i = x^p - 1$ und $\binom{p}{i} = \frac{p \cdot \ldots \cdot (p-i+1)}{1 \cdot \ldots \cdot i} \equiv 0 (\mathrm{mod}\ p)$, weil p als Primzahl in diesem Bruch nicht gekürzt werden kann. Damit ist $(x-1)^{p-k}(x-1)^k = x^p - 1$.

Mit diesen Hilfsmitteln können wir jetzt Polynome angeben, die einen zyklischen Code größtmöglicher Hamming-Norm erzeugen.

Satz 6.4.26 *Der durch $g_k := (x-1)^{p-k}$ über $K = GF(q)$ generierte zyklische Code hat die Hamming-Norm $p - k + 1$.*

Beweis. Für $k = 1$ ist $\widehat{g_1} : K^1 \longrightarrow K^p$ gegeben durch $g_1(\alpha) = \alpha \cdot (x-1)^{p-1}$. Nun ist $(x-1)(x-1)^{p-1} = (x-1)^p = x^p - 1 = (x-1)(x^{p-1} + x^{p-2} + \ldots + x + 1)$. Da $K[x]$ nullteilerfrei ist, ist $(x-1)^{p-1} = x^{p-1} + x^{p-2} + \ldots + x + 1$. Damit erhalten wir $g_1 \cdot \alpha = \alpha \cdot (x^{p-1} + x^{p-2} + \ldots + x + 1)$, also $\widehat{g_1}(\alpha) = (\alpha, \ldots, \alpha)$. Offenbar hat diese Codierung nach Definition die Hamming-Norm p.

Für $k = p$ ist $g_p = 1$, also $\widehat{g_p} : K^p \longrightarrow K^p$ die identische Abbildung mit der Hamming-Norm 1.

Wir zeigen jetzt $\|\widehat{g}_{k+1}\| < \|\widehat{g}_k\|$ für alle $0 \leq k < p$. Sei $(\alpha_0, \ldots, \alpha_{p-1}) \in \mathrm{Bi}(\widehat{g}_k)$ ein Codewort minimaler Norm. Weil wir einen zyklischen Code haben, können wir $\alpha_0 \neq 0$ annehmen. Für $f \neq 0$ gilt $\mathrm{Grad}(g_k f) = \mathrm{Grad}(g_k) + \mathrm{Grad}(f) = (p-k) + \mathrm{Grad}(f) \geq p - k$, also ist auch $\alpha_i \neq 0$ für ein $i > 0$. Dann ist das zugehörige Polynom $\alpha_0 + \alpha_1 x + \ldots + \alpha_{p-1}x^{p-1} = g_k f = (x-1)^{p-k} f$. Wir bilden die formale Ableitung und erhalten $\alpha_1 + 2\alpha_2 x + \ldots + (p-1)\alpha_{p-1}x^{p-2} = (g_k f)' = (p-k)(x-1)^{p-k-1} f + (x-1)^{p-k} f' = (x-1)^{p-(k+1)}((p-k)f + (x-1)f')$. Daher ist $(\alpha_1, 2\alpha_2, \ldots, (p-1)\alpha_{p-1}, 0) \in \mathrm{Bi}(\widehat{g}_{k+1})$ ein Element kleinerer Norm. Damit ist $\|\widehat{g}_{k+1}\| < \|\widehat{g}_k\|$ und

$$1 = \|\widehat{g}_p\| < \|\widehat{g}_{p-1}\| < \ldots < \|\widehat{g}_1\| = p,$$

woraus $\|\widehat{g}_k\| = p - k + 1$ folgt.

Übungen 6.4.27 1. Die folgende Codierung sei gegeben

$$\begin{array}{ll} f(00) = 000000, & f(10) = 101101, \\ f(01) = 010011, & f(11) = 111110. \end{array}$$

a) Zeigen Sie, daß f eine lineare Codierung ist.
b) Bestimmen Sie die erzeugende Matrix.
c) Finden Sie eine Kontrollabbildung und die zugehörige Kontrollmatrix.
d) Bestimmen Sie den Hamming-Abstand von f.
e) Entziffern Sie folgende Worte

$$100000, 101111, 010010, 010101, 011110.$$

2. Finden Sie je eine (3,6)-Codierung f mit $||f|| = n$ für $n = 1, 2, 3, 4$.
3. Sei $K = \mathbb{Z}/(2)$. Zeigen Sie:

a) $A := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ ist eine Codierungsmatrix.

b) Finden Sie eine geeignete Kontrollmatrix dazu.
c) Bestimmen Sie den Hamming-Abstand des Codes.
d) Es werden die Nachrichten

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

empfangen. Analysieren Sie!

7. Eigenwerttheorie

Die Determinante einer quadratischen Matrix oder eines Endomorphismus ist eine der wichtigsten Invarianten. Mit ihr kann man feststellen, ob eine Matrix invertierbar ist. Sie gestattet es aber auch, geometrische Eigenschaften eines Endomorphismus genauer zu studieren. Dies wird im Abschnitt über Eigenwerte geschehen.

7.1 Determinanten

Wir führen den Begriff der Determinante auf eine wenig übliche Weise ein. Die verwendete Methode führt besonders schnell zu den wichtigsten Eigenschaften.

Definition 7.1.1 Eine Abbildung $\Delta : K_n^n \longrightarrow K$ heißt eine *Determinantenfunktion*, wenn gelten

(Δ1) $\quad \Delta(B) = \Delta(A)$,

falls B aus A durch Addition einer Zeile zu einer anderen Zeile entsteht, und

(Δ2) $\quad \Delta(B) = \alpha \cdot \Delta(A)$,

falls B aus A durch Multiplikation einer Zeile mit einem Faktor α entsteht.

Satz 7.1.2 *Sei $\Delta : K_n^n \longrightarrow K$ eine Determinantenfunktion. Dann gilt*

1. *$\Delta(A) = 0$ für alle $A \in K_n^n$ mit Rang$(A) < n$,*
2. *$\Delta = 0$, falls $\Delta(E_n) = 0$.*

Beweis. 1. Für Elementarmatrizen zweiter Art von der Form $F_{ij} = E_n + E_{ij}$ gilt nach 6.2.10

$$\Delta(F_{ij}A) = \Delta(A). \tag{7.1}$$

Für Elementarmatrizen erster Art von der Form $F_i(\alpha) = E_n + (\alpha - 1)E_{ii}$ gilt

$$\Delta(F_i(\alpha)A) = \alpha\Delta(A). \tag{7.2}$$

Daraus folgt für Elementarmatrizen der Form $F_{ij}(\alpha) = E_n + \alpha E_{ij} = F_j(\alpha^{-1})F_{ij}F_j(\alpha)$

$$\Delta(F_{ij}(\alpha)A) = \Delta(A) \tag{7.1a}$$

wobei

$$F_{ij}(\alpha) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_i + \alpha \cdot a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}.$$

Elementarmatrizen dritter Art der Form P_{ij} können geschrieben werden als $P_{ij} = F_i(-1)F_{ij}F_j(-1)F_{ji}F_i(-1)F_{ij}$, d.h. die Vertauschung zweier Zeilen kann mit elementaren Zeilenoperationen erster und zweiter Art dargestellt werden, insbesondere gilt

$$\Delta(P_{ij}A) = -\Delta(A). \tag{7.3}$$

Also gibt es zu jeder invertierbaren Matrix B einen von Null verschiedenen Faktor b mit

$$\Delta(B \cdot A) = b\Delta(A)$$

für alle Matrizen A, denn die Faktoren, die man bei einer Zerlegung von B in Elementarmatrizen erhält, hängen nur von B und nicht von A ab. Ist $\mathrm{Rang}(A) < n$, dann hat die Stufenform $B \cdot A = S$ von A als letzte Zeile den Nullvektor. Daher gilt $\Delta(A) = b^{-1}\Delta(S) = 0$ wegen $(\Delta 2)$.

2. Ist $\mathrm{Rang}(A) = n$ und $\Delta(E_n) = 0$, so ist $\Delta(A) = \Delta(A \cdot E_n) = a \cdot \Delta(E_n) = 0$, also $\Delta = 0$.

Wir vergleichen Determinantenfunktionen und erhalten dabei den Begriff der Determinante.

Folgerung 7.1.3 *Seien $\Delta_1, \Delta_2 : K_n^n \longrightarrow K$ Determinantenfunktionen. Dann gilt*

$$\Delta_1(E_n)\Delta_2(A) = \Delta_1(A)\Delta_2(E_n).$$

Beweis. $\Delta(A) := \Delta_1(E_n)\Delta_2(A) - \Delta_1(A)\Delta_2(E_n)$ ist ebenfalls eine Determinantenfunktion, da sie $(\Delta 1)$ und $(\Delta 2)$ erfüllt. Weiter ist $\Delta(E_n) = 0$. Also ist $\Delta = 0$ und damit die Behauptung bewiesen.

Folgerung 7.1.4 *Sei $\Delta : K_n^n \longrightarrow K$ eine Determinantenfunktion mit $\Delta(E_n) = 1$. Dann gilt $\Delta(A \cdot B) = \Delta(A)\Delta(B)$ für alle $A, B \in K_n^n$.*

Beweis. $\Delta_1(A) := \Delta(A \cdot B)$ ist eine Determinantenfunktion wegen (7.1) und (7.2). Also folgt mit $\Delta_2 = \Delta$ aus 7.1.3
$\Delta(A \cdot B) = \Delta_1(A) \cdot \Delta_2(E_n) = \Delta_1(E_n) \cdot \Delta_2(A) = \Delta(B) \cdot \Delta(A)$.

Definition 7.1.5 Eine Abbildung $\Delta : K_n^n \longrightarrow K$ heißt *(zeilen-) multilinear*, wenn Δ aufgefaßt als Abbildung auf dem n-Tupel der Zeilenvektoren in jedem Argument (in jeder Zeile) linear ist, d.h. wenn

$$\lambda \cdot \Delta \begin{pmatrix} a_1 \\ \vdots \\ x \\ \vdots \\ a_n \end{pmatrix} + \mu \cdot \Delta \begin{pmatrix} a_1 \\ \vdots \\ y \\ \vdots \\ a_n \end{pmatrix} = \Delta \begin{pmatrix} a_1 \\ \vdots \\ \lambda x + \mu y \\ \vdots \\ a_n \end{pmatrix} \text{ gilt.}$$

Satz 7.1.6 *Zu jedem n gilt es eine eindeutig bestimmte Determinantenfunktion* $\det : K_n^n \longrightarrow K$ *mit* $\det(E_n) = 1$*. Diese Determinantenfunktion ist multilinear. Ist* $\Delta : K_n^n \longrightarrow K$ *eine Determinantenfunktion, so gilt für alle* $A \in K_n^n$

$$\Delta(A) = \det(A) \cdot \Delta(E_n).$$

Beweis. Die Eindeutigkeit: Seien Δ_1 und Δ_2 Determinantenfunktion mit $\Delta_1(E_n) = 1 = \Delta_2(E_n)$. Nach 7.1.3 folgt dann $\Delta_1(A) = \Delta_1(A)\Delta_2(E_n) = \Delta_1(E_n)\Delta_2(A) = \Delta_2(A)$. Für den Existenzbeweis verwenden wir vollständige Induktion nach n. Für $n = 1$ ist $\det : K_1^1 \ni (\alpha) \mapsto \alpha \in K$ offenbar eine Determinantenfunktion mit $\det(E_1) = 1$. Sie ist außerdem (multi-)linear. Sei die Existenz für $n - 1$ bewiesen, und sei

$$A = \begin{pmatrix} \alpha_{11} & b_1 \\ \vdots & \vdots \\ \alpha_{n1} & b_n \end{pmatrix} \in K_n^n. \text{ Sei } B_k = \begin{pmatrix} b_1 \\ \vdots \\ \widehat{b_k} \\ \vdots \\ b_n \end{pmatrix} \in K_{n-1}^{n-1},$$

wobei die Zeile b_k fortzulassen ist. Das soll mit $\widehat{b_k}$ angedeutet werden. Nach Induktionsannahme ist $\det(B_k)$ definiert und multilinear in den Zeilen b_i. Wir definieren

$$\det(A) := \sum_{k=1}^{n} (-1)^{k+1} \alpha_{k1} \det(B_k). \tag{7.4}$$

Die Abbildung det ist multilinear in den Zeilen. Wir zeigen dieses für die erste Zeile:

$$\det \begin{pmatrix} \lambda\alpha_{11} + \mu\alpha'_{11} & \lambda b_1 + \mu b'_1 \\ \vdots & \vdots \\ \alpha_{n1} & b_n \end{pmatrix}$$

$$= (-1)^{1+1}(\lambda\alpha_{11} + \mu\alpha'_{11}) \det(B_1) =$$

$$
\begin{aligned}
&+\sum_{k=2}^{n}(-1)^{k+1}\alpha_{k1}(\lambda \det(B_k) + \mu \det(B'_k)) \\
&= \lambda\sum_{k=1}^{n}(-1)^{k+1}\alpha_{k1}\det(B_k) + \mu\sum_{k=1}^{n}(-1)^{k+1}\alpha_{k1}\det(B'_k) \\
&= \lambda \cdot \det\begin{pmatrix} \alpha_{11} & b_1 \\ \vdots & \vdots \\ \alpha_{n1} & b_n \end{pmatrix} + \mu \det\begin{pmatrix} \alpha'_{11} & b'_1 \\ \vdots & \vdots \\ \alpha_{n1} & b_n \end{pmatrix}.
\end{aligned}
$$

Insbesondere ist $(\Delta 2)$ erfüllt.

Wir betrachten wieder nur den Spezialfall $\lambda = \mu = 1$ und $(\alpha'_{11}b'_1) = (\alpha_{21}b_2)$, so ist

$$
\det\begin{pmatrix} \alpha_{11}+\alpha_{21} & b_1+b_2 \\ \vdots & \vdots \\ \alpha_{n1} & b_n \end{pmatrix} = \det\begin{pmatrix} \alpha_{11} & b_1 \\ \vdots & \vdots \\ \alpha_{n1} & b_n \end{pmatrix} + \det\begin{pmatrix} \alpha_{21} & b_2 \\ \vdots & \\ \alpha_{n1} & b_1 \end{pmatrix}
$$

wobei in der letzten Matrix die erste und zweite Zeile gleich sind. In (7.4) fallen damit der erste und zweite Summand fort und bei $B_3, \ldots, B_n$ sind jeweils erste und zweite Zeile gleich. Da $\det(B_k)$ eine Determinantenfunktion ist, ist $\det(B_k) = 0$ für $k = 3, \ldots, n$. Man kann nämlich eine Nullzeile erhalten.

Schließlich ist $\det(E_n) = 1$, weil $\det(E_{n-1}) = 1$. Endlich ist $\Delta(A) = \Delta(A)\det(E_n) = \Delta(E_n)\det(A)$ nach 7.1.3.

Definition 7.1.7 Die Determinantenfunktion $\det : K_n^n \longrightarrow K$ heißt *Determinante.*

Die Rechenregeln für Determinantenfunktionen ergeben jetzt die wichtigsten Eigenschaften der Determinante.

Folgerung 7.1.8 *1. Die Determinante einer Matrix ist genau dann Null, wenn die Zeilen bzw. Spalten der Matrix linear abhängig sind.*

2. Die Determinante einer Matrix ändert sich nicht, wenn man zu einer Zeile eine Linearkombination der anderen Zeilen addiert.

3. Die Determinante einer Matrix ändert ihr Vorzeichen, wenn man zwei Spalten vertauscht.

4. $\det(A \cdot B) = \det(A) \cdot \det(B)$.

5. $\det(\lambda \cdot A) = \lambda^n \cdot \det(A)$.

6. $\det(A^t) = \det(A)$.

7. Es gilt $\det\begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \alpha_{nn} \end{pmatrix} = \alpha_{11} \cdot \ldots \cdot \alpha_{nn}$,

d.h. die Determinante einer oberen Dreiecksmatrix ist das Produkt aller Elemente auf der Diagonalen.

Beweis. 2. Zunächst ist wegen ($\varDelta 1$) die Addition einer Zeile zu einer anderen möglich. Ein Vielfaches der $j-$ten Zeile kann ebenfalls aus $i-$ten Zeile addiert werden, indem man zunächst die $j-$te Zeile mit $\alpha \neq 0$ multipliziert, dann addiert, dann die j-te Zeile mit α^{-1} multipliziert. Wegen ($\varDelta 2$) ändert sich die Determinante nicht. Der Prozeß kann für mehrere Zeilen wiederholt werden.

1. Wenn die Zeilen linear abhängig sind, ist der Rang $< n$ und nach 7.1.2 1. die Determinante Null. Ist Rang $A = n$, also A regulär, so ist nach 7.1.4 $\det(A) \cdot \det(A^{-1}) = \det(A \cdot A^{-1}) = \det(E_n) = 1$, also $\det(A) \neq 0$.

3. folgt aus (7.3)

4. ist 7.1.4

5. folgt aus ($\varDelta 2$), für jede Zeile einmal angewendet.

6. Wegen $(A^t)^{-1} = (A^{-1})^t$ ist A genau dann invertierbar, wenn A^t invertierbar ist. Nur in diesem Fall ist die Behauptung zu zeigen. A ist Produkt von Elementarmatrizen $A = F_1 \cdot \ldots \cdot F_k$. Dann ist $A^t = F_k^t \cdot \ldots \cdot F_1^t$. Da für Elementarmatrizen F_i gilt $\det(F_i^t) = \det(F_i)$, genauer für Elementarmatrizen

$$\begin{array}{rl} F_i \text{ erster Art} & \det(F_i) = \alpha, \\ F_i \text{ zweiter Art} & \det(F_i) = 1, \\ F_i \text{ dritter Art} & \det(F_i) = -1 \end{array}$$

(wie in (7.1), (7.2), (7.3) mit $A = E_n$), ist also $\det(A^t) = \det(A)$ nach Teil 4.

7. Wenn eines der $\alpha_{ii} = 0$, dann ist Rang $A < n$, also $\det(A) = 0 = \alpha_{11} \cdot \ldots \cdot \alpha_{nn}$. Wenn alle $\alpha_{ii} \neq 0$ sind, dann können Vielfache der unteren Zeilen zu den oberen Zeilen so addiert werden, daß eine Diagonalmatrix mit $\alpha_{11}, \ldots, \alpha_{nn}$ in der Diagonalen entsteht. Die Determinante ändert sich nicht. Nach ($\varDelta 2$) auf jede Zeile angewendet folgt $\det(A) = \alpha_{11} \cdot \ldots \cdot \alpha_{nn} \cdot \det(E_n) = \alpha_{11} \cdot \ldots \cdot \alpha_{nn}$.

Beispiele 7.1.9 1. $\det(\alpha) = \alpha$

2. $\det \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \alpha\delta - \gamma\beta$

3.

$$\det \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix} = \alpha_1 \det \begin{pmatrix} \beta_2 & \gamma_2 \\ \beta_3 & \gamma_3 \end{pmatrix}$$

$$-\alpha_2 \det \begin{pmatrix} \beta_1 & \gamma_1 \\ \beta_3 & \gamma_3 \end{pmatrix} + \alpha_3 \det \begin{pmatrix} \beta_1 & \gamma_1 \\ \beta_2 & \gamma_2 \end{pmatrix}$$

$$= \alpha_1\beta_2\gamma_3 - \alpha_1\beta_3 v\gamma_2 - \alpha_2\beta_1 v\gamma_3 + \alpha_2\beta_3\gamma_1 + \alpha_3\beta_1\gamma_2 - \alpha_3\beta_2\gamma_1$$

$$= \begin{array}{ccccccccc} \alpha_1 & & \beta_1 & & \gamma_1 & & \alpha_1 & & \beta_1 \\ & \diagdown & & \times & & \times & & \diagup & \\ \alpha_2 & & \beta_2 & & \gamma_2 & & \alpha_2 & & \beta_2 \\ & \diagup & & \times & & \times & & \diagdown & \\ \alpha_3 & & \beta_3 & & \gamma_3 & & \alpha_3 & & \beta_3 \end{array} \qquad \text{(Regel von Sarrus)}$$

Die Berechnung von Determinanten ist kompliziert. Wir geben nur eine Methode dazu an. Eine algorithmisch schnellere Methode erhält man, wenn man bei elementaren Zeilenumformungen in geschickter Weise die Determinanten der verwendeten Elementarmatrizen sammelt, bis durch Zeilenumformungen die Einheitsmatrix erreicht ist.

Definition 7.1.10 Sei $A \in K_n^n$. Mit A_{ij} werde die aus A durch Streichung der i-ten Zeile und j-ten Spalte entstehende Matrix in K_{n-1}^{n-1} bezeichnet. A_{ij} heißt dann auch *Streichungsmatrix*, $\det(A_{ij})$ *Streichungsdeterminante.*

Satz 7.1.11 *(Entwicklungssatz nach der j-ten Spalte) Sei $A \in K_n^n$ gegeben. Dann ist*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \alpha_{ij} \det(A_{ij}).$$

Beweis. Wir vertauschen in A die j-te Spalte schrittweise mit den vorhergehenden Spalten, bis sie an erster Stellte steht. Diese neue Matrix A' hat $\det(A') = (-1)^{j-i} \det(A)$ und damit $\det(A') = \sum_{i=1}^{n} (-1)^{i+1} \alpha_{ij} \det(A_{ij})$ als Determinante. Daraus folgt die Behauptung.

Bemerkung 7.1.12 1. Wegen $\det(A^t) = \det(A)$ folgt auch ein entsprechender Entwicklungssatz nach der i-ten Zeile:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ij} \det(A_{ij}).$$

2. Zur praktischen Berechnung kann man also jeweils eine Entwicklung nach einer besonders geeigneten (mit möglichst vielen Nullen besetzten) Zeile oder Spalte durchführen, z.B.

$$\det \begin{pmatrix} 0 & 1 & 2 \\ 3 & 7 & 1 \\ 0 & 1 & 0 \end{pmatrix} = (-1)^4 \cdot 0 \cdot \det \begin{pmatrix} 1 & 2 \\ 7 & 1 \end{pmatrix} + (-1)^5 \cdot 1 \cdot \det \begin{pmatrix} 0 & 2 \\ 3 & 1 \end{pmatrix}$$

$$+(-1)^6 \cdot 0 \cdot \det \begin{pmatrix} 0 & 1 \\ 3 & 7 \end{pmatrix} = 6$$

durch Entwicklung nach der dritten Zeile.

3. Für $A \in K_n^n$ ist $\det(A) = \sum_\sigma (\pm 1) \alpha_{1\sigma(1)} \cdot \alpha_{2\sigma(2)} \cdot \ldots \cdot \alpha_{n\sigma(n)}$ mit geeigneten Vorzeichen, wobei σ durch alle Permutationen der Zahlen $\{1, \ldots, n\}$ läuft, also $n!$ Summanden definiert. Aus jeder Zeile und jeder Spalte kommt in jedem der Produkte genau ein Koeffizient vor. Das sieht man durch vollständige Induktion und Auswertung der Entwicklung nach einer Zeile. Denn in (7.4) kommen in $\det(A_{ij})$ aus jeder Zeile von A außer der i-ten und jeder Spalte von A außer der j-ten in den Summanden jeweils genau ein Faktor vor, und man erhält $(n-1)!$ Summanden. Also ergibt (7.4) insgesamt $n \cdot (n-1)! = n!$ Summanden.

Die Determinante einer Matrix kann zur Berechnung der inversen Matrix verwendet werden. Dazu definieren wir

Definition und Folgerung 7.1.13 *Sei* $A \in K_n^n$ *und* $B = (\beta_{ij})$ *mit* $\beta_{ij} = (-1)^{i+j}\det(A_{ji})$ *definiert.* B *heißt* Komplementärmatrix *zu* A. *Damit ist*

$$A \cdot B = \det(A)E_n.$$

Ist A *regulär, so ist* $A^{-1} = \dfrac{1}{\det(A)}B$.

Beweis. Es ist $\sum_k \alpha_{ik}\beta_{kj} = \sum_k (-1)^{j+k}\alpha_{ik}\det(A_{jk}) = \det(A)\cdot\delta_{ij}$, denn für $i = j$ ist dies der Entwicklungssatz nach der j-ten Zeile. Ist jedoch $i \neq j$, so können wir diesen Ausdruck ebenfalls nach dem Entwicklungssatz nach der j-ten Zeile als Determinante einer Matrix A' auffassen, die aus A durch Ersetzen der j-ten Zeile durch die i-te Zeile entstanden ist. Diese Matrix ist singulär, also ist $\det(A') = 0$.

Satz 7.1.14 *(Cramersche[1] Regel) Sei* $A = (a_1, \dots, a_n) \in K_n^n$ *eine reguläre Matrix mit den Spaltenvektoren* a_i. *Dann hat das lineare Gleichungssystem* $A \cdot x = b$ *die Lösung*

$$\xi_i = \frac{1}{\det(A)}\det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n).$$

Beweis. (Die eindeutig bestimmte) Lösung ist $x = A^{-1} \cdot b = \frac{1}{\det(A)}Bb$. Wenn $b = (\beta_i)$ ist und $B = (\beta_{ij})$ die Komplementärmatrix zu A ist, dann ist $\det(A)\xi_i = \sum_{j=1}^m \beta_{ij}\beta_j = \sum_{j=1}^n (-1)^{i+j}\beta_j \det(A_{ji}) = \det(a_1, ..., a_{i-1}, b, a_{i+1}, ..., a_n)$, woraus die Behauptung folgt.

Bisher haben wir lediglich Determinanten von Matrizen betrachtet. Jetzt wenden wir uns Determinanten für Endomorphismen von endlichdimensionalen Vektorräumen zu.

Satz 7.1.15 *Sei* V *ein endlichdimensionaler Vektorraum und* $f : V \longrightarrow V$ *ein Endomorphismus. Sei* B *eine Basisfamilie von* V. *Dann ist die Determinante der darstellenden Matrix von* f *bezüglich* B *unabhängig von der Wahl von* B.

Beweis. Seien B und B' Basisfamilien von V. Sei S die Basistransformation von B nach B' und M die darstellende Matrix von f bzgl. B. Nach 5.5.20 ist dann SMS^{-1} die darstellende Matrix von f bzgl. B'. Weiter ist

$$\det(SMS^{-1}) = \det(S)\det(M)\det(S^{-1}) = \det(SS^{-1})\det(M) = \det(M).$$

[1] Gabriel Cramer (1704–1752)

Definition 7.1.16 Für einen Endomorphismus $f : V \longrightarrow V$ mit darstellender Matrix M definieren wir die Determinante $\det(f) := \det(M)$. Sie ist nach 7.1.15 unabhängig von der Wahl der Basisfamilie.

Übungen 7.1.17 1. Zeigen Sie, daß $\begin{pmatrix} a \\ b \end{pmatrix}$ und $\begin{pmatrix} c \\ d \end{pmatrix}$ genau dann linear unabhängig sind, wenn $ad - bc \neq 0$ ist.
2. Berechnen Sie elementargeometrisch (mit Schulkenntnissen) den Flächeninhalt des Parallelogramms mit den Eckpunkten

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} a+c \\ b+d \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}.$$

Vergleichen Sie Ihr Ergebnis mit der Determinante von $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$.
3. Seien $x_1, x_2, x_3 \in K$ Elemente eines Körpers K. Zeigen Sie:

$$\det \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{pmatrix} = (x_3 - x_2)(x_3 - x_1)(x_2 - x_1).$$

4. Entscheiden Sie, ob die folgende Aussage richtig ist (ja/nein).
Sind alle Streichungsdeterminanten einer quadratischen Matrix Null, so ist auch die Determinante dieser Matrix Null.

7.2 Eigenwerte und Eigenvektoren

Eigenwerte und Eigenvektoren haben breite Anwendungsbereiche, u.a. bei Differentialgleichungen, in der Technik und der Physik. Wir beschränken uns hier auf endlichdimensionale Vektorräume und ihre Endomorphismen.

Sei im folgenden $V \neq 0$ ein Vektorraum und $f : V \longrightarrow V$ ein Endomorphismus.

Definition 7.2.1 Ein Skalar $\lambda \in K$ heißt ein *Eigenwert* von f, wenn es ein $v \in V, v \neq 0$ gibt mit

$$f(v) = \lambda v.$$

Ein Vektor $v \in V, v \neq 0$ mit $f(v) = \lambda v$ heißt ein *Eigenvektor* zum Eigenwert λ. Die Menge $V_\lambda := \{v \in V | f(v) = \lambda v\}$ heißt *Eigenraum* zum Eigenwert λ. Die Menge $\{\lambda \in K | \lambda \text{ Eigenwert von } f\}$ heißt *Spektrum* von f.

Bemerkung 7.2.2 V_λ ist ein Untervektorraum von V. Die Eigenvektoren zu λ sind genau die von Null verschiedenen Vektoren in V_λ.

Beispiele 7.2.3 1. Sei $f : V \longrightarrow V$ nicht injektiv. Dann ist 0 ein Eigenwert von f und $\mathrm{Ke}(f) = V_0 \neq 0$. Es gibt nämlich ein $v \in V$ mit $v \neq 0$ und $f(v) = 0$, also $f(v) = 0 \cdot v$. Weiter ist $v \in \mathrm{Ke}(f) \Longleftrightarrow f(v) = 0 \cdot v \Longleftrightarrow v \in V_0$.

2. Für $\mathrm{id} : V \longrightarrow V$ ist 1 der einzige Eigenwert, und es ist $V = V_1$. Es ist nämlich $\mathrm{id}(v) = v = 1 \cdot v$ für alle $v \in V$.

3. Die lineare Abbildung $\widehat{M} : K_2 \longrightarrow K_2$ mit $M = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ hat die Eigenwerte 1 und 2 und die Eigenräume $V_1 = \left\{ \begin{pmatrix} \alpha \\ 0 \end{pmatrix} \right\}$ und $V_2 = \left\{ \begin{pmatrix} 0 \\ \beta \end{pmatrix} \right\}$.

Satz 7.2.4 *Seien* $\{\lambda_i | i \in I\}$ *paarweise verschiedene Eigenwerte von* f. *Dann gilt* $\sum_{i \in I} V_{\lambda_i} = \oplus_{i \in I} V_{\lambda_i}$.

Beweis. Wir müssen zeigen, daß aus $\sum v_i = 0$ mit $v_i \in V_{\lambda_i}$ folgt $v_i = 0$ für alle $i \in I$. Wenn das nicht der Fall ist, dann gibt es eine Summe $\sum_{i=1}^n v_i = 0$ kürzester Länge $n > 0$. Offenbar sind in einer solchen Summe alle $v_i \neq 0$. Daraus folgt $n \geq 2$. Es ist $\sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n f(v_i) = f(\sum_{i=1}^n v_i) = f(0) = 0$ ebenso wie $\sum_{i=1}^n \lambda_1 v_i = \lambda_1 \sum_{i=1}^n v_i = 0$. Die Differenz ist $\sum_{i=2}^n (\lambda_i - \lambda_1) v_i = 0$, eine Summe kürzerer Länge mit von Null verschiedenen Summanden $(\lambda_i - \lambda_1) v_i$, ein Widerspruch.

Definition 7.2.5 Eigenwerte, Eigenvektoren, Eigenräume und Spektrum einer quadratischen(!) Matrix M sind die der linearen Abbildung $\widehat{M} : K_n \longrightarrow K_n$.

Übungen 7.2.6 1. Zeigen Sie: Eine nilpotente Matrix (eine Matrix M mit $M^n = 0$) hat nur den Eigenwert 0.

2. Bestimmen Sie die Eigenwerte und die Eigenvektoren der folgenden Matrix:

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

3. Habe $f : V \longrightarrow V$ einen Eigenwert λ. Zeigen Sie:
   a) f^n hat einen Eigenwert λ^n.
   b) Es gilt $V_\lambda(f) \subseteq V_{\lambda^n}(f^n)$.
   c) Finden Sie ein $f : V \longrightarrow V$ und einen Eigenwert λ von f, so daß gilt $V_\lambda(f) \subsetneqq V_{\lambda^n}(f^n)$.

7.3 Das charakteristische Polynom

Zur Bestimmung von Eigenwerten und damit auch von Eigenräumen verwendet man häufig das charakteristische Polynom. Damit wird der Zusammenhang zwischen Eigenwerten und Determinanten hergestellt.

Definition 7.3.1 Sei V endlichdimensional mit Basisfamilie B, und sei $f : V \longrightarrow V$ ein Endomorphismus mit der darstellenden Matrix M.

Für $\lambda \in K$ heißt

$$\det(f - \lambda \cdot \mathrm{id}) = 0$$

die *charakteristische Gleichung* von f.

$$\chi_M(x) := \det(M - x \cdot E_n)$$

heißt das *charakteristische Polynom* von M.

Lemma 7.3.2 *$\chi_M(x)$ ist ein Polynom von Grad n. Es gilt für $M = (\alpha_{ij})$*

$$\chi_M(x) = (-1)^n \cdot \big(x^n - \sum_{i=1}^{n} \alpha_{ii} x^{n-1} + \ldots \pm \det(M)\big)$$

Beweis. Nach Bemerkung 7.1.12 3. ist

$$\det(M) = \sum_{\sigma} (\pm 1) \alpha_{1\sigma(1)} \cdot \ldots \cdot \alpha_{n\sigma(n)}.$$

Für $\det(M - xE_n)$ sind die Faktoren in der Diagonalen von der Form $\alpha_{ii} - x$. Die höchste Potenz von x ergibt sich, wenn alle Faktoren im Produkt auf der Diagonalen liegen, also für $\sigma = \mathrm{id}$, mit $(\alpha_{11} - x) \cdot \ldots \cdot (\alpha_{nn} - x) = (-1)^n \cdot x^n + (-1)^{n-1} \sum \alpha_{ii} x^{n-1} +$ Terme vom Grad $\leq n-2$. Wenn im Produkt ein Faktor nicht auf der Diagonalen liegt, so muß auch noch ein zweiter Faktor außerhalb der Diagonalen auftreten. Der Polynomgrad des Produkts ist dann $\leq n-2$. Um den konstanten Koeffizienten des Polynoms zu erhalten, setze man $x = 0$ ein. Dann ist der konstante Koeffizient gleich $\det(M - 0 \cdot E_n) = \det(M)$.

Satz 7.3.3 *Sei* $\dim V = n < \infty, f \in \mathrm{Hom}_K(V, V)$*. Dann gilt für* $\lambda \in K$*:*
λ *ist Eigenwert von* $f \Longleftrightarrow \det(f - \lambda \mathrm{id}) = 0$.

Beweis. λ Eigenwert von $f \Longleftrightarrow \exists v \neq 0 [f(v) = \lambda v] \Longleftrightarrow \exists v \neq 0 [(f - \lambda \mathrm{id})(v) = 0] \Longleftrightarrow f - \lambda \mathrm{id}$ nicht bijektiv $\Longleftrightarrow \det(f - \lambda \mathrm{id}) = 0$.

Bemerkung 7.3.4 1. $V_\lambda = \mathrm{Ke}(f - \lambda \mathrm{id})$.

2. Es gibt höchstens n Werte λ mit $\det(f - \lambda \mathrm{id}) = 0$, weil das charakteristische Polynom $\det(M - xE_n)$ höchstens n Nullstellen hat und weil $\det(f - \lambda \mathrm{id}) = \det(M - \lambda E_n)$, wobei $M - \lambda E_n$ eine darstellende Matrix für $f - \lambda \mathrm{id}$ ist.

3. $f : V \longrightarrow V$ hat höchstens n verschiedene Eigenwerte.

Definition 7.3.5 1. Eine Nullstelle λ eines Polynoms $f(x)$ hat die *Vielfachheit* k, wenn $f(x)$ durch $(x - \lambda)^k$ (ohne Rest) teilbar ist, aber nicht durch $(x - \lambda)^{k+1}$.

2. Ein Eigenwert λ der Matrix M bzw. des Endomorphismus f hat die *algebraische Vielfachheit* $k = \mu(\chi_f, \lambda)$, wenn er als Nullstelle des charakteristischen Polynoms die Vielfachheit k hat.

3. Ein Eigenwert λ der Matrix M bzw. des Endomorphismus f hat die *geometrische Vielfachheit* k, wenn $k = \dim V_\lambda$.

Satz 7.3.6 *Sei* $\dim V = n < \infty$ *und* $f \in \mathrm{Hom}_K(V,V)$. *Sei* $\lambda \in K$ *ein Eigenwert von* f. *Dann ist* $\mu(\chi_f, \lambda) \geq \dim V_\lambda$.

Beweis. Sei $(v_1, \ldots, v_k)$ eine Basisfamilie von V_λ und $(v_1, \ldots, v_n)$ eine Fortsetzung zu einer Basisfamilie von V. Dann ist bezüglich dieser Basisfamilie die darstellende Matrix von f von der Form

$$M = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & * \\ 0 & & \lambda & \\ \hline & 0 & & M^1 \end{array}\right) \quad \text{und}$$

$\det(M - xE_n) = \chi_f = (\lambda - x)^k \cdot \det(M^1 - xE_{n-k})$ durch Entwicklungen nach der 1. bis r. Spalte.

Übungen 7.3.7 1. Bestimmen Sie die reellen bzw. komplexen Eigenwerte von

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

2. Bestimmen Sie die Eigenwerte und Eigenräume von

$$\begin{pmatrix} 2 & 3 & -1 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \quad \text{und von} \quad \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

3. Entscheiden Sie, ob die folgende Aussage richtig ist (ja/nein).
Die Summe der geometrischen Vielfachheiten eines Endomorphismus eines endlichdimensionalen Vektorraumes ist gerade die Dimension dieses Vektorraumes.
4. Sei

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

Zeigen Sie $\chi_A = x^2 - \mathrm{Spur}(A) + \det A$, wobei für eine $n \times n$-Matrix $M = (\alpha_{ij})$ die Spur definiert ist als $\mathrm{Spur}(M) := \sum_{i=1}^n \alpha_{ii}$.
5. Zeigen Sie für den in der vorangehenden Aufgabe definierten Begriff der Spur:
   a) Die Abbildung $\mathrm{Spur} : K_n^n \longrightarrow K$ ist linear.
   b) Es gilt $\mathrm{Spur}(AB) = \mathrm{Spur}(BA)$ für alle $A, B \in K_n^n$.
   c) Es gilt im allgemeinen nicht $\mathrm{Spur}(ABC) = \mathrm{Spur}(BAC)$ für $A, B, C \in K_n^n$.
   d) Sei V ein endlichdimensionaler Vektorraum und $f : V \longrightarrow V$ eine lineare Abbildung. Sei B eine Basisfamilie von V und M die darstellende Matrix von f bezüglich B. Zeigen Sie, daß $\mathrm{Spur}(M)$ nicht von der Wahl von B abhängt (so daß also $\mathrm{Spur}(f) := \mathrm{Spur}(M)$ wohldefiniert ist).

7.4 Diagonalisierbare Matrizen und Endomorphismen

Gewisse quadratische Matrizen bzw. Endomorphismen kann man durch geschickte Wahl einer Basis auf eine besonders einfache Form, die Diagonalform, bringen. Wir wollen diese Endomorphismen charakterisieren und ihre Eigenschaften studieren.

Definition 7.4.1 1. $f \in \mathrm{Hom}_K(V, V)$ heißt *diagonalisierbar*, wenn es eine Basisfamilie B von V so gibt, daß die darstellende Matrix von f bzgl. B eine Diagonalmatrix ist.

2. Eine Matrix $M \in K_n^n$ heißt *diagonalisierbar*, wenn es eine reguläre Matrix S so gibt, daß SMS^{-1} eine Diagonalmatrix ist.

Satz 7.4.2 *Für $f \in \mathrm{Hom}_k(V, V)$ sind äquivalent*

1. *f ist diagonalisierbar,*
2. *es gibt in V eine Basis aus Eigenvektoren von f,*
3. *a) $\chi_f(x)$ ist ein Produkt von Linearfaktoren $(\lambda_i - x)$,*
   b) für alle Eigenwerte λ_i von f gilt $\mu(\chi_f, \lambda_i) = \dim V_{\lambda_i}$,
4. *ist $\{\lambda_1, \ldots, \lambda_k\}$ das Spektrum von f, so ist*

$$V = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}.$$

Beweis. 1. $\iff$ 2. : weil die Basisfamilie, bezüglich der f eine darstellende Matrix in Diagonalform hat, eine Basisfamilie aus Eigenvektoren ist: $f(v_i) = \lambda_i v_i$.

1. $\Longrightarrow$ 4. : $\oplus V_{\lambda_i} \subseteq V$ gilt immer. Da $\oplus V_{\lambda_i}$ eine Basis von V enthält, ist $V = \oplus V_{\lambda_i}$.

4. $\Longrightarrow$ 3. : Wegen $\chi_f(x) = (x - \lambda_1)^{i_1} \cdot \ldots \cdot (x - \lambda_k)^{i_k} \cdot g(x)$ ist $n = \sum_{i=1}^k \dim V_{\lambda_i} \leq \sum_{i=1}^k \mu(\chi_f, \lambda_i) \leq n$, also gilt Gleichheit, insbesondere $\mu(\chi_f, \lambda_i) = \dim V_{\lambda_i}$ und $g(x) = \alpha \in K$.

3. $\Longrightarrow$ 2. : $n = \sum_{i=1}^k \mu(\chi_f, \lambda_i) = \sum_{i=1}^k \dim V_{\lambda_i}$ impliziert wegen 7.3.6 die Gleichung $V = \oplus_{i=1}^k V_{\lambda_i}$. Insbesondere hat V eine Basis(-familie) von Eigenvektoren.

Beispiele 7.4.3 1. $\begin{pmatrix} 0 & 2 & 4 \\ 1 & 1 & 0 \\ -2 & 2 & 5 \end{pmatrix}$ ist diagonalisierbar wegen

$$\begin{pmatrix} 1 & 0 & 2 \\ 1 & 2 & 1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 3 & -2 & -4 \\ -1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 4 \\ 1 & 1 & 0 \\ -2 & 2 & 5 \end{pmatrix}.$$

Es ist $\chi_M(x) = \det \begin{pmatrix} -x & 2 & 4 \\ 1 & 1-x & 0 \\ -2 & 2 & 5-x \end{pmatrix} = (-x)(1-x)(5-x) + 8 - 2(5 - x) + 8(1-x) = -x^3 + 6x^2 - 11x + 6 = -(x-1)(x-2)(x-3)$. Da $\mu(\chi_M(x), 1) =$

$\mu(\chi_M(x), 2) = \mu(\chi_M(x), 3) = 1$, ist nach 7.3.6 $\dim V_1 = \dim V_2 = \dim V_3 = 1$, also M diagonalisierbar mit $SMS^{-1} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Die Matrizen S bzw. S^{-1} sind Transformationsmatrizen. Wenn b_1, b_2, b_3 eine Basis von K_3 aus Eigenvektoren ist, dann transformiert S^{-1} die Basis e_1, e_2, e_3 in b_1, b_2, b_3, und es gilt

$$S^{-1} = (b_1, b_2, b_3).$$

Die Eigenvektoren erhält man aus dem linearen Gleichungssystem $(M - \lambda_i E_n) \cdot x = 0$.

2. $M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist nicht diagonalisierbar, weil

$$\chi_M(x) = \det \begin{pmatrix} 1-x & 1 \\ 0 & 1-x \end{pmatrix} = (1-x)^2,$$

also 1 einziger Eigenwert der algebraischen Vielfachheit 2 ist, und weil $(M - \lambda E_2)x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x = 0$ einen 1-dimensionalen Lösungsraum $E_1 = K \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ hat.

3. $M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ ist über $\mathbb{R}$ nicht diagonalisierbar, weil $\chi_M(x) = \det \begin{pmatrix} -x & -1 \\ 1 & -x \end{pmatrix} = x^2 + 1$ keine Nullstellen hat, M also keine Eigenwerte bzw. Eigenvektoren hat. M ist eine Drehung um 90°!

Übungen 7.4.4 1. Ist

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

diagonalisierbar?

2. Finden Sie eine diagonalisierbare Matrix, die dasselbe charakteristische Polynom wie $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ hat.
3. Zeigen Sie: $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ ist diagonalisierbar.
4. Zeigen Sie: $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ ist nicht diagonalisierbar.
5. Entscheiden Sie, ob die folgende Aussage richtig ist (ja/nein).
   Eine Matrix ist genau dann diagonalisierbar, wenn ihr charakteristisches Polynom in Linearfaktoren zerfällt.
6. a) Sei M diagonalisierbar als $D = SMS^{-1}$. Zeigen Sie: $D^n = SM^nS^{-1}$.
   b) Berechnen Sie mit Hilfe des Teils 1. M^{10} für $M = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$.

7.5 Potenzmethode zur Bestimmung dominanter Eigenwerte (R. v. Mises)

Wir haben gesehen, daß Eigenwerte eine besondere Bedeutung haben und haben auch schon Methoden kennengelernt, sie zu berechnen. Es gibt ein numerisches Verfahren, gewisse Eigenwerte näherungsweise zu bestimmen, das wir zum Abschluß dieses Kapitels besprechen wollen.

Definition 7.5.1 Sei $f : K_n \longrightarrow K_n$ ein diagonalisierbarer Endomorphismus mit den Eigenwerten $\lambda_1, \ldots \lambda_n$, wobei Eigenwerte entsprechend ihrer Vielfachheit gezählt werden. Sei $K = \mathbb{R}$ oder $\mathbb{C}$ und sei $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$. Dann heißt λ_1 ein *dominanter Eigenwert.*

Satz 7.5.2 *Sei λ_1 ein dominanter Eigenwert von $f : K_n \longrightarrow K_n$. Sei $y \in K_n \setminus \{0\}$ und sei $y^{(m)} := f(y^{(n-1)}) = f^m(y)$, $y^{(0)} := y$ und sei $y^{(m)} = (y_1^{(m)}, \ldots, y_n^{(m)})^t \in K_n$. Dann konvergieren die Folgen $\big(\frac{1}{\lambda_1^m} \cdot y_i^{(m)} | m \in \mathbb{N}_0\big)$. Insbesondere ist $\lim_{m\longrightarrow\infty} \frac{y_i^{(m)}}{y_i^{(m-1)}} = \lambda_1$ für alle i, für die $\lim_{m\longrightarrow\infty} \frac{1}{\lambda_1^m} y_i^{(m)}$ $\neq 0$ gilt.*

Beweis. Seien $x_1, \ldots, x_n$ eine Basis aus Eigenvektoren von f zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$. Sei $y = \sum \eta_i x_i$, und sei $x_i = (\xi_{ik}|k)$. Dann folgt $(|\lambda_1| > |\lambda_2| ... \Longrightarrow \lambda_1 \neq 0)$

$$f(y) = y^{(1)} = \sum \eta_i f(x_i) = \sum \eta_i \lambda_i x_i = \lambda_1 \cdot \sum \eta_i (\frac{\lambda_i}{\lambda_1}) x_i \text{ und}$$

$$\frac{1}{\lambda_1^m} \cdot y^{(m)} = \frac{1}{\lambda_1^m} f^m(y) = \sum \eta_i (\frac{\lambda_i}{\lambda_1})^m x_i.$$

Für genügend große m wird

$$|\frac{1}{\lambda_1^{m+r}} y_k^{(m+r)} - \frac{1}{\lambda_1^m} y_k^{(m)}| = |\sum_{i=2}^n \eta_i (\frac{\lambda_i}{\lambda_1})^m \Big((\frac{\lambda_i}{\lambda_1})^r - 1 \Big) \xi_{ik}| \leq$$

$$\sum_{i=2}^n |\eta_i| \cdot |(\frac{\lambda_i}{\lambda_1})^r - 1| \cdot |\xi_{ik}| \cdot |\frac{\lambda_i}{\lambda_1}|^m < \epsilon$$

da aus $|\lambda_1| > |\lambda_i|$ folgt $|\frac{\lambda_i}{\lambda_1}| < 1$.

Wenn $\lim_{m\longrightarrow\infty} \frac{1}{\lambda_1^m} y_i^{(m)} \neq 0$, dann ist für genügend große m

$$\epsilon > |\frac{\frac{1}{\lambda_1^{m+1}} y_i^{(m+1)}}{\frac{1}{\lambda_1^m} y_i^{(m)}} - 1| = |\frac{1}{\lambda_1} \cdot \frac{y_i^{(m+1)}}{y_i^{(m)}} - 1| = |\frac{1}{\lambda_1}| |\frac{y_i^{(m+1)}}{y_i^{(m)}} - \lambda_1|$$

$$\text{also } \lim_{m\longrightarrow\infty} \left(\frac{y_i^{(m+1)}}{y_i^{(m)}} \right) = \lambda_1.$$

Beispiel 7.5.3 Für $f : K_n \longrightarrow K_n$ verwenden wir $\widehat{M} : \mathbb{R}_3 \longrightarrow \mathbb{R}_3$ mit $M = \begin{pmatrix} 1 & 2 & -1 \\ 7 & 6 & -1 \\ -4 & -4 & 1 \end{pmatrix}$ und $y = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Dann ist

M			$y^{(0)}$	$y^{(1)}$	$y^{(2)}$	$y^{(3)}$	$y^{(4)}$	$\approx \lambda_1^{12}$	$\approx \lambda_1^{23}$	$\approx \lambda_1^{34}$
1	2	−1	1	2	33	282	2553	16.5	8.54	9.05
7	6	−1	1	12	93	852	7653	7.75	9.16	8.98
−4	−4	1	1	−7	−63	−567	−5103	9	9	9

Folgerung 7.5.4 *Wenn* $\lim_{m\longrightarrow\infty} \frac{1}{\lambda_1^m} y^{(m)} = z \neq 0$*, dann ist* z *Eigenvektor zu* λ_1.

Beweis. $f(z) = f\bigl(\lim_{m\longrightarrow\infty} \frac{1}{\lambda_1^m} y^{(m)}\bigr) = \lim_{m\longrightarrow\infty} \frac{1}{\lambda_1^m} f(y^{(m)}) = \lim_{m\longrightarrow\infty} \lambda_1 \cdot \frac{1}{\lambda_1^{(m+1)}} y^{(m+1)} = \lambda_1 \cdot z$. (Dabei ist f als lineare Abbildung stetig.)

Übungen 7.5.5 1. Stellen Sie fest, ob die folgenden Matrizen dominante Eigenwerte besitzen und bestimmen Sie diese.

$$\begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & 4 \\ 1 & -1 \end{pmatrix}, \quad \begin{pmatrix} 4 & 0 & 0 \\ 2 & -5 & 0 \\ 1 & 3 & 6 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

2. Verwenden Sie die Potenzmethode, um einen dominanten Eigenwert und einen zugehörigen Eigenvektor der folgenden Matrix zu bestimmen:

$$\begin{pmatrix} 4 & 3 & 6 \\ 4 & 13 & 20 \\ -2 & -6 & -9 \end{pmatrix}.$$

8. Euklidische Vektorräume

Wir führen in diesem Kapitel weitere geometrische Eigenschaften für einem Vektorraum ein, insbesondere Längen von Vektoren und Winkel zwischen Vektoren. Damit gelingt es nun, viele elementargeometrische Aussagen zu beweisen. Wesentliches Hilfsmittel hierfür wird das zusätzliches Strukturdatum des Skalarprodukts.

Wir setzen in diesem Kapitel voraus, daß alle Vektorräume über dem Körper $\mathbb{R}$ definiert sind, also reelle Vektorräume sind.

8.1 Skalarprodukte

Der zentrale neue Begriff ist der des Skalarprodukts, eines Produkts zwischen Vektoren, das reelle Werte annimmt.

Definition 8.1.1 Sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung $\sigma : V \times V \longrightarrow \mathbb{R}$ heißt eine *Bilinearform* auf V, wenn $\forall y \in V[V \ni x \mapsto \sigma(x,y) \in \mathbb{R}]$ Homomorphismus und $\forall x \in V[V \ni y \mapsto \sigma(x,y) \in \mathbb{R}]$ Homomorphismus. Eine Bilinearform σ heißt

nichtausgeartet, wenn $\forall x \in V, x \neq 0[\sigma(x,V) \neq 0$ und $\sigma(V,x) \neq 0]$;
symmetrisch, wenn $\forall x, y \in V[\sigma(x,y) = \sigma(y,x)]$;
positiv definit, wenn $\forall x \in V, x \neq 0[\sigma(x,x) > 0]$.

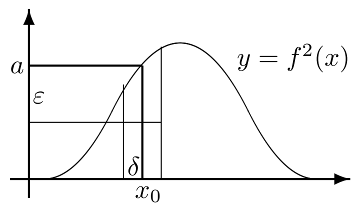
Eine positiv definite, symmetrische Bilinearform heißt ein *Skalarprodukt* auf V. Ein reeller Vektorraum V zusammen mit einem Skalarprodukt σ heißt *Euklidischer*[1] *Vektorraum*. Wir schreiben dann auch $\langle x, y\rangle := \sigma(x,y)$.

Beispiele 8.1.2 1. In $V = \mathbb{R}_n$ ist $\sigma(x,y) := x^t y = \sum_{i=1}^n \xi_i \eta_i$ ein Skalarprodukt. Sicher ist nämlich $\sigma(x,y) = x^t y$ bilinear, $\sigma(x,y) = \sum \xi_i \eta_i$ symmetrisch, und es ist $\sigma(x,x) = \sum \xi_i^2 > 0$ für $x \neq 0$. Dieses Skalarprodukt heißt *kanonisches Skalarprodukt* des $\mathbb{R}_n$.

2. Das nächste Beispiel sieht zunächst recht exotisch aus. Es ist aber Grundlage für große Teile der Analysis, insbesondere der Funktionalanalysis und der Theorie der Differentialgleichungen, auf die wir in diesem Buch nicht weiter eingehen können.

[1] Euklid ca. 300 v.Chr.

Sei $V = \{f\colon [0,1] \longrightarrow \mathbb{R} \mid f \text{ stetig}\}$. Dann ist $\sigma(f,g) := \int_0^1 f(x)g(x)dx$ ein Skalarprodukt. Die Bilinearität und Symmetrie in f und g sind trivial. Ist $f \neq 0$, so ist $\int_0^1 f(x)f(x)dx > 0$. Es gibt nämlich ein $x_0 \in [0,1]$ mit $f(x_0) = a \neq 0$. Da f stetig ist, ist auch f^2 stetig und $f^2(x_0) = a^2 > 0$. Dann gibt es zu $\varepsilon := \frac{1}{2}a^2$ ein $\delta > 0$ mit $|f^2(x) - f^2(x_0)| < \varepsilon$ für alle x mit $|x - x_0| < \delta$. Also ist $f^2(x) > \frac{1}{2}a^2$ für alle x mit $|x - x_0| < \delta$. Daraus folgt $\int_0^1 f^2(x)dx \geq \frac{1}{2}a^2 \cdot \delta > 0$.



Für die Einführung der geometrischen Begriffe in Euklidischen Vektorräumen ist die folgende Cauchy-Schwarzsche Ungleichung von besonderer Bedeutung.

Satz 8.1.3 *(Cauchy[2]-Schwarz[3]sche Ungleichung) Sei (V,σ) ein Euklidischer Vektorraum. Dann gilt für alle $x, y \in V$*

$$\langle x, y\rangle^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle. \tag{8.1}$$

Weiterhin gilt: x, y linear abhängig $\Longleftrightarrow \langle x, y\rangle^2 = \langle x, x\rangle \cdot \langle y, y\rangle$.

Beweis. Für $y = 0$ ist die Aussage des Satzes klar. Sei also $y \neq 0$. Zunächst gilt

$$0 \leq \langle x - \alpha y, x - \alpha y\rangle = \langle x, x\rangle - 2\alpha\langle x, y\rangle + \alpha^2\langle y, y\rangle.$$

Wir setzen $\alpha := \langle x, y\rangle / \langle y, y\rangle$ und multiplizieren mit $\langle y, y\rangle$. Das ergibt

$$\begin{aligned} 0 &\leq \langle x, x\rangle\langle y, y\rangle - 2\langle x, y\rangle\langle x, y\rangle + \langle x, y\rangle\langle x, y\rangle \\ &= \langle x, x\rangle\langle y, y\rangle - \langle x, y\rangle^2. \end{aligned}$$

Sei nun $x = \beta y$. Dann ist $\langle x, y\rangle\langle x, y\rangle = \beta^2\langle y, y\rangle^2 = \langle x, x\rangle\langle y, y\rangle$, also gilt Gleichheit in (8.1). Gelte schließlich Gleichheit in (8.1). Wenn $x = 0$ oder $y = 0$, dann sind x und y linear abhängig. Sei also $x \neq 0$ und $y \neq 0$. Dann ist $0 < \langle x, x\rangle\langle y, y\rangle = \langle x, y\rangle^2$, also ist $\langle x, y\rangle \neq 0$. Wir zeigen jetzt, daß $x = \dfrac{\langle x, x\rangle}{\langle x, y\rangle} y$ gilt. Dann sind nämlich x und y linear abhängig. Es ist

[2] Baron August-Louis Cauchy (1789–1857)
[3] Hermann Amandus Schwarz (1843–1921)

$$\begin{aligned}\langle x - \frac{\langle x,x\rangle}{\langle x,y\rangle}y, x - \frac{\langle x,x\rangle}{\langle x,y\rangle}y\rangle &= \langle x,x\rangle - 2\frac{\langle x,x\rangle}{\langle x,y\rangle}\langle x,y\rangle + \frac{\langle x,x\rangle\langle x,x\rangle}{\langle x,y\rangle\langle x,y\rangle}\langle y,y\rangle \\ &= \langle x,x\rangle - 2\frac{\langle x,x\rangle\langle x,y\rangle}{\langle x,y\rangle} + \frac{\langle x,y\rangle\langle x,y\rangle}{\langle x,y\rangle\langle x,y\rangle}\langle x,x\rangle \\ &= \langle x,x\rangle - 2\langle x,x\rangle + \langle x,x\rangle = 0.\end{aligned}$$

Da σ positiv definit ist, folgt $x - \frac{\langle x,x\rangle}{\langle x,y\rangle}y = 0$ und daraus die Behauptung.

Jetzt haben wir alle Hilfsmittel bereit, um Winkel und Längen einzuführen.

Definition 8.1.4 1. Sei $(V, \langle,\rangle)$ ein Euklidischer Vektorraum. Zwei Vektoren $x, y \in V$ heißen *orthogonal* oder *senkrecht* $(x \perp y)$, wenn $\langle x,y\rangle = 0$.

2. Wegen $\langle x,y\rangle^2 \le \langle x,x\rangle\langle y,y\rangle$ gilt

$$-1 \le \frac{\langle x,y\rangle}{\sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle}} \le 1$$

(für $x, y \neq 0$) also existiert genau ein $\varphi \in [0,\pi]$ mit

$$\cos\varphi = \frac{\langle x,y\rangle}{\sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle}}.$$

φ heißt der *Winkel* zwischen x und y. Wir schreiben $\Theta(x,y) := \varphi$.

Bemerkung 8.1.5 1. $x \perp y \Longleftrightarrow \Theta(x,y) = \frac{\pi}{2}$.

2. x, y linear abhängig $\Longleftrightarrow \Theta(x,y) \in \{0,\pi\}$.

Übungen 8.1.6 1. Stellen Sie fest, welche der folgenden Bildungen ein Skalarprodukt definieren und welche nicht. Geben Sie ggf. an, welche Axiome verletzt sind. Für $x = (\xi_1,\xi_2,\xi_3), y = (\eta_1,\eta_2,\eta_3) \in \mathbb{R}^3$ seien definiert:

a) $\langle x,y\rangle := \xi_1\eta_1 + \xi_3\eta_3$,
b) $\langle x,y\rangle := \xi_1^2\eta_1^2 + \xi_2^2\eta_2^2 + \xi_3^2\eta_3^2$,
c) $\langle x,y\rangle := 4\xi_1\eta_1 + \xi_2\eta_2 + 3\xi_3\eta_3$,
d) $\langle x,y\rangle := \xi_1\eta_1 - \xi_2\eta_2 + \xi_3\eta_3$,
e) $\langle x,y\rangle := \xi_1\eta_2 + \xi_2\eta_1 + \xi_3\eta_3$.

2. Sei $V = M_2$ der Vektorraum der reellen 2×2-Matrizen. Stellen Sie fest, ob durch

$$\langle \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} u & x \\ y & z \end{pmatrix} \rangle = au + bx + by - cx + cy + dz$$

ein Skalarprodukt definiert wird.

3. Benutzen Sie das Skalarprodukt

$$\langle f,g\rangle = \int_0^1 f(x)g(x)dx,$$

um den Wert von $\langle f,g\rangle$ auszurechnen für

a) $f = \cos(2\pi x)$, $g = \sin(2\pi x)$.
b) $f = x$, $g = e^x$.
c) $f = \tan \frac{\pi}{4} x$, $g = 1$.

4. Sei P_2 der Vektorraum der reellen Polynome vom Grad ≤ 2.
   a) Zeigen Sie, daß P_2 mit
   $$\langle a_2x^2 + a_1x + a_0, b_2x^2 + b_1x + b_0 \rangle := a_0b_0 + a_1b_1 + a_2b_2$$
   ein Euklidischer Vektorraum ist.
   b) Bestimmen Sie $||p||$ für $p = -1 + 2x + x^2$.
   c) Bestimmen Sie den Winkel zwischen $p = -1 + 5x + 2x^2$ und $q = 2 + 4x - 9x^2$.
   d) Bestimmen Sie den Winkel zwischen $p = x - x^2$ und $q = 7 + 3x + 3x^2$.
5. Sei P_2 der Vektorraum der reellen Polynome vom Grad ≤ 2 mit dem Skalarprodukt
   $$\langle p, q \rangle = \int_0^1 p(x)q(x)dx.$$
   a) Bestimmen Sie $||p||$ für $p = -1 + 2x + x^2$.
   b) Bestimmen Sie den Winkel zwischen $p = -1 + 5x + 2x^2$ und $q = 2 + 4x - 9x^2$.
   c) Bestimmen Sie den Winkel zwischen $p = x - x^2$ und $q = 7 + 3x + 3x^2$.
6. Zeigen Sie mit Hilfe der Cauchy-Schwarzschen Ungleichung, daß für alle $a, b \in \mathbb{R}$ und alle Winkel φ gilt:
   $$(a\cos(\varphi) + b\sin(\varphi))^2 \leq a^2 + b^2.$$
7. Zeigen Sie mit Hilfe der Cauchy-Schwarzschen Ungleichung, daß für alle positive reellen Zahlen $a_1, \ldots, a_n$ gilt
   $$n^2 \leq (a_1 + \ldots + a_n) \left(\frac{1}{a_1} + \ldots + \frac{1}{a_n} \right).$$
8. Zeigen Sie mit Hilfe der Cauchy-Schwarzschen Ungleichung, daß für alle reellen Zahlen $a_1, \ldots, a_n$ gilt
   $$(a_1 + \ldots + a_n)^2 \leq n^2(a_1^2 + \ldots + a_n^2).$$
9. Die Vektoren $(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)$ werden als Kanten des Einheitswürfels im $\mathbb{R}^n$ aufgefaßt (jede Kante tritt mehrfach auf, wie oft?). Zeigen Sie, daß der Winkel φ zwischen der Diagonalen $(1, 1, \ldots, 1)$ und den Kanten die Gleichung
   $$\cos(\varphi) = \frac{1}{\sqrt{n}}$$
   erfüllen.

8.2 Normierte Vektorräume

Die Bildung der Länge oder Norm eines Vektors in einem Euklidischen Vektorraum erfüllt die Gesetze einer Norm. Es gibt viele Beispiele von normierten Vektorräumen, auch von solchen, die nicht aus einem Euklidischen Vektorraum entstehen.

Definition 8.2.1 Sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung

$$\|.\| : V \ni x \mapsto \|x\| \in \mathbb{R}$$

heißt *Norm*, wenn

1. $\|\alpha x\| = |\alpha|\|x\|$,
2. $\|x + y\| \leq \|x\| + \|y\|$
3. $\|x\| = 0 \Longrightarrow x = 0$.

$(V, \|.\|)$ heißt *normierter Vektorraum.* $x \in V$ heißt *normiert*, wenn $\|x\| = 1$.

Bemerkung 8.2.2 Ist $(V, \|.\|)$ ein normierter Vektorraum, so ist $\|x\| \geq 0$ für alle $x \in V$. Es ist nämlich $0 = |0|\|x\| = \|0 \cdot x\| = \|x + (-x)\| \leq \|x\| + \|-x\| = \|x\| + |-1|\|x\| = 2\|x\|$.

Satz 8.2.3 *Sei* $(V, \langle,\rangle)$ *ein Euklidischer Vektorraum. Dann ist* $\|x\| := \sqrt{\langle x, x\rangle}$ *eine Norm auf* V*.* $\|x\|$ *heißt* Norm *oder* Länge *von* x*.*

Beweis. 1. $\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha^2}\sqrt{\langle x, x\rangle} = |\alpha|\|x\|$.

2. $\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \leq \langle x, x\rangle + 2\sqrt{\langle x, y\rangle^2} + \langle y, y\rangle \leq \langle x, x\rangle + 2\sqrt{\langle x, x\rangle\langle y, y\rangle} + \langle y, y\rangle = (\|x\| + \|y\|)^2 \Longrightarrow \|x+y\| \leq \|x\| + \|y\|$.

3. $\|x\| = 0 \Longrightarrow \sqrt{\langle x, x\rangle} = 0 \Longrightarrow \langle x, x\rangle = 0 \Longrightarrow x = 0$.

Mit den Gesetzen der Norm und des Skalarprodukts können wir jetzt einige der wichtigsten geometrischen Sätze beweisen.

Satz 8.2.4 *(Pythagoras[4])*

1. $\forall x, y \in V[x \perp y \Longrightarrow \|x + y\|^2 = \|x\|^2 + \|y\|^2]$.
2. $\forall x, y \in V[\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle]$.

Beweis. 1. $x \perp y \Longrightarrow 2\langle x, y\rangle = 0$.

2. $\|x + y\|^2 = \langle x + y, x + y\rangle = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$.

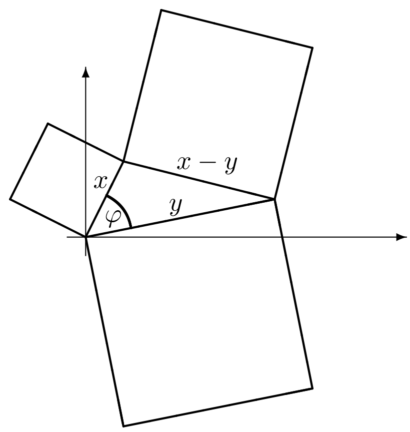
Lemma 8.2.5 $\forall x, y \in V \setminus \{0\}[\langle x, y\rangle = \|x\|\|y\| \cos(\Theta(x, y))]$.

Beweis. $\|x\|\|y\| \cos(\Theta(x, y)) = \|x\|\|y\| \cdot \frac{\langle x,y\rangle}{\|x\|\|y\|} = \langle x, y\rangle$.

[4] Pythagoras (580–500)

Satz 8.2.6 *(Cosinus Satz)*

$$\forall x, y \in V \setminus \{0\}[\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\| \cos(\Theta(x, y)).$$



Beweis. $\|x-y\|^2 = \|x\|^2+\|y\|^2-2\langle x, y\rangle = \|x\|^2+\|y\|^2-2\|x\|\|y\| \cos(\Theta(x, y))$.

Wir haben gezeigt, daß Euklidische Vektorräume auch normierte Vektorräume sind. Jetzt zeigen wir, daß normierte Vektorräume auch metrische Räume sind. Schließlich sind dann metrische Räume auch topologische Räume. Jede dieser Folgerungen ist eine echte Folgerung, also nicht umkehrbar.

Lemma 8.2.7 *Sei $(V, \|.\|)$ ein normierter Vektorraum. Dann ist durch $d : V \times V \ni (x, y) \mapsto \|x - y\| \in \mathbb{R}$ eine Metrik auf V gegeben. (vgl. 6.4.4)*

Beweis. 1. $d(x, y) = \|x - y\| = 0 \Longleftrightarrow x - y = 0 \Longleftrightarrow x = y$.
2. $d(x, y) = \|x - y\| = |-1|\|y - x\| = d(y, x)$.
3. $d(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$.

Bemerkung 8.2.8 Sei (X, d) ein metrischer Raum (vgl. 4.2.4 und 6.4.4). Dann gilt $\forall x \neq y[d(x, y) > 0]$.

Beweis. $2d(x, y) = d(x, y) + d(y, x) \geq d(x, x) = 0 \Longrightarrow d(x, y) \geq 0$. $x \neq y \Longrightarrow d(x, y) > 0$.

Wir führen nur gewisse offene Mengen in einem metrischen Raum ein, definieren jedoch nicht, was genau ein topologischer Raum ist, da wir diesen Begriff später nicht weiter brauchen.

Definition 8.2.9 Sei (M, d) ein metrischer Raum. Sei $r \in \mathbb{R}^+ \setminus \{0\}$ und $a \in M$. Dann heißt $K(a;r) := \{b \in M | d(a,b) \leq r\}$ *abgeschlossene Kugel* um a mit dem *Radius* r. $K_\circ(a;r) := \{b \in M | d(a,b) < r\}$ heißt *offene Kugel.* $U \subseteq M$ heißt *offen,* wenn es zu jedem $a \in U$ ein $r \in \mathbb{R}^+ \setminus \{0\}$ gibt mit $K_\circ(a;r) \subseteq U$.

Bemerkung 8.2.10 Jeder metrische Raum ist ein topologischer Raum.

Definition 8.2.11 Sei $(V, \langle , \rangle)$ ein endlichdimensionaler Euklidischer Vektorraum. Eine Basis $b_1, \ldots, b_n$ von V heißt *Orthonormalbasis* von V, wenn gilt $\langle b_i, b_j \rangle = \delta_{ij}$.

Alle Vektoren in einer Orthonormalbasis haben also die Länge 1 und je zwei verschiedene Basisvektoren sind senkrecht zueinander.

Satz 8.2.12 *(Gram[5]- Schmidtsches[6] Orthonormalisierungsverfahren)*
Sei $(V, \langle ., . \rangle)$ ein endlichdimensionaler Euklidischer Vektorraum. Dann existiert eine Orthonormalbasis von V.

Beweis. Der Beweis zeigt ein Verfahren zur Konstruktion einer Orthonormalbasis auf. Das Verfahren selbst ist für viele Rechnungen sehr wichtig.

Sei $(c_1, \ldots, c_n)$ eine Basis von V. Wir bilden induktiv

$$\begin{aligned} b_1 &:= \tfrac{1}{\|c_1\|} c_1; \\ d_{r+1} &:= c_{r+1} - \textstyle\sum_{i=1}^{r} \langle c_{r+1}, b_i \rangle b_i; \\ b_{r+1} &:= \tfrac{1}{\|d_{r+1}\|} d_{r+1}. \end{aligned}$$

Der mittlere Ausdruck $d_{r+1} := c_{r+1} - \sum_{i=1}^{r} \langle c_{r+1}, b_i \rangle b_i$ stellt die Projektion des Basisvektors d_{r+1} auf den von den $b_1, \ldots, b_r$ aufgespannten Untervektorraum dar, wie wir im folgenden sehen werden.

Mit der Konstruktion wird $(b_1, \ldots, b_n)$ eine Orthonormalbasis. Zunächst sind die b_i normiert, denn für $x \neq 0$ ist $\|\frac{1}{\|x\|} x\| = \frac{1}{\|x\|} \|x\| = 1$. Weiter ist $\langle b_j, d_{r+1} \rangle = \langle b_j, c_{r+1} \rangle - \sum_{i=1}^{r} \langle b_i, c_{r+1} \rangle \langle b_j, b_i \rangle = \langle b_j, c_{r+1} \rangle - \langle b_j, c_{r+1} \rangle = 0$, also auch $\langle b_j, b_{r+1} \rangle = 0$ für $j < r+1$. Wegen $c_{r+1} \notin \langle b_1, \ldots, b_r \rangle = \langle c_1, \ldots, c_r \rangle$ (Ind. Ann.) ist $d_{r+1} \neq 0$. Außerdem ist $b_{r+1} \in \langle c_{r+1}, b_1, \ldots, b_r \rangle$ und $c_{r+1} \in \langle b_1, \ldots, b_{r+1} \rangle$, also ist $\langle b_1, \ldots, b_{r+1} \rangle = \langle c_1, \ldots, c_{r+1} \rangle$.

Satz 8.2.13 *Sei V ein Euklidischer Vektorraum der Dimension $n < \infty$. Sei $B = (b_1, \ldots, b_n)$ eine Orthonormalbasis von V und $x \in V$. Dann gelten*

$$x = \sum_{i=1}^{n} \langle x, b_i \rangle b_i = \sum_{i=1}^{n} \|x\| \cos(\Theta(x, b_i)) b_i \qquad \textit{und}$$

$$\|x\|^2 = \sum_{i=1}^{n} \langle x, b_i \rangle^2.$$

[5] J8rgen Amandus Gram (1850–1916)
[6] Erhardt Schmidt (1876-1959)

Beweis. $\langle \sum_i \langle x, b_i\rangle b_i, b_j\rangle = \sum \langle x, b_i\rangle\langle b_i, b_j\rangle = \langle x, b_j\rangle$ für alle b_j, also gilt $\langle \sum \langle x, b_i\rangle b_i, y\rangle = \langle x, y\rangle$ für alle $y \in V$ oder $\langle \sum_i \langle x, b_i\rangle b_i - x, y\rangle = 0$. Insbesondere ist $\langle \sum_i \langle x, b_i\rangle b_i - x, \sum_i \langle x, b_i\rangle b_i - x\rangle = 0$, also $\sum \langle x, b_i\rangle b_i = x$. Dann folgt
$\|x\|^2 = \langle x, x\rangle = \langle \sum_i \langle x, b_i\rangle b_i, \sum_j \langle x, b_j\rangle b_j\rangle = \sum_i \langle x, b_i\rangle^2$.

Satz 8.2.14 *Sei V ein Euklidischer Vektorraum und sei $U \subseteq V$ ein endlichdimensionaler Untervektorraum. Dann gilt $V = U \oplus U^\perp$, wobei $U^\perp := \{v \in V | \forall u \in U[\langle u, v\rangle = 0]\}$ das* orthogonale Komplement *von U ist.*

Beweis. U erbt das Skalarprodukt von V. Sei $b_1, \ldots, b_n$ eine Orthonormalbasis von U. Sei $f : V \longrightarrow U \subseteq V$ definiert durch $f(v) := \sum_{i=1}^n \langle v, b_i\rangle b_i$. Dann gilt für alle $u \in U$

$$f(u) = \sum \langle u, b_i\rangle b_i = u.$$

Insbesondere gilt $\mathrm{Bi}(f) = U$. Weiter ist $f(v) = 0 \iff \sum \langle v, b_i\rangle b_i = 0 \iff \forall i = 1, \ldots, n[\langle v, b_i\rangle = 0] \iff \forall u \in U[\langle v, u\rangle = 0] \iff v \in U^\perp$. Also ist $U^\perp = \mathrm{Ke}(f)$ ein Untervektorraum. Weiter ist $f^2(v) = f(f(v)) = f(v)$, weil $f(v) \in U$, also $f^2 = f \Longrightarrow V = \mathrm{Ke}(f) \oplus \mathrm{Bi}(f)$ und $\mathrm{Ke}(f) = U^\perp$ und $\mathrm{Bi}(f) = U \Longrightarrow V = U \oplus U^\perp$ (vgl. 5.4.20).

Bemerkung 8.2.15 Die oben konstruierte Abbildung $f : V \longrightarrow V$ heißt *Projektion auf den Untervektorraum U*. Wir werden später (8.5.4) zeigen, daß f nicht von der Wahl der Orthonormalbasis $b_1, \ldots, b_n$ von U abhängt.

Folgerung 8.2.16 *Sei V ein endlichdimensionaler Euklidischer Vektorraum. Ist $U \subseteq V$ ein Untervektorraum, so gelten*

1. $V = U \oplus U^\perp$ *und* $\dim V = \dim U + \dim U^\perp$.
2. $U^{\perp\perp} = U$.

Beweis. 1. ist trivial

2. Wegen $\langle u, u'\rangle = 0$ für alle $u \in U$ und $u' \in U^\perp$ ist $U \subseteq U^{\perp\perp}$. Aus Dimensionsgründen folgt $U = U^{\perp\perp}$

Definition 8.2.17 Sei V ein Euklidischer Vektorraum und seien $U_1, \ldots, U_n$ Untervektorräume von V. V ist *orthogonale Summe* der U_i, wenn

1. $V = U_1 + \ldots + U_n$,
2. $\forall i \neq j[U_i \perp U_j]($ d.h. $\forall x \in U_i, y \in U_j[\langle x, y\rangle = 0])$.

Bemerkung 8.2.18 Eine orthogonale Summe ist eine direkte Summe.

Beweis. Sei $\sum_i u_i = 0$. Dann ist $\langle u_j, u_j\rangle = \langle \sum_i u_i, u_j\rangle = 0$, also $u_j = 0$.

Satz 8.2.19 *Jeder endlichdimensionale Euklidische Vektorraum ist eine orthogonale Summe von eindimensionalen Untervektorräumen.*

Beweis. Sei $b_1, \ldots, b_n$ eine Orthonormalbasis. Dann ist $V = \mathbb{R}b_1 \perp \ldots \perp \mathbb{R}b_n$.

Übungen 8.2.20 1. Zeigen Sie, daß in einem Euklidischen Vektorraum V gilt

$$\langle x, y\rangle = \frac{1}{4}\left(||x+y||^2 - ||x-y||^2\right).$$

2. Zeigen Sie, daß in einem Euklidischen Vektorraum V gilt

$$||x+y||^2 + ||x-y||^2 = 2||x||^2 + 2||y||^2.$$

3. Zeigen Sie: wenn in einem normierten Vektorraum V

$$||x+y||^2 - ||x-y||^2$$

linear in x und in y ist, dann ist V ein Euklidischer Vektorraum und $||x||$ die durch ein Skalarprodukt definierte Norm.
4. Betrachten Sie $\mathbb{R}^4$ als Euklidischen Vektorraum mit dem kanonischen Skalarprodukt. Seien $u, x, y, z \in \mathbb{R}^4$ gegeben mit

$$u = (1, 0, 0, 1),\ x = (-1, 2, 0, 1),$$
$$y = (-1, -1, 2, 1),\ z = (2, 2, 3, -2).$$

   a) Zeigen Sie, daß je zwei der gegebenen Vektoren orthogonal zueinander sind und normieren Sie die gegebenen Vektoren zu einer Orthonormalbasis.
   b) Drücken Sie den Vektor $(1, 1, 1, 1)$ als Linearkombination den gefundenen Orthonormalbasis aus.
5. Betrachten Sie $\mathbb{R}^2$ als Euklidischen Vektorraum mit dem kanonischen Skalarprodukt. Benutzen Sie das Gram-Schmidtsche Orthonormalisierungsverfahren, um aus jeder der Mengen

$$\{(1, -3), (2, 2)\} \qquad \text{bzw.} \qquad \{(1, 0), (3, -5)\}$$

eine Orthonormalbasis zu machen.
6. Betrachten Sie $\mathbb{R}^4$ als Euklidischen Vektorraum mit dem kanonischen Skalarprodukt. Benutzen Sie das Gram-Schmidtsche Orthonormalisierungsverfahren, um aus der Menge

$$\{(0, 2, 1, 0), (1, -1, 0, 0), (1, 2, 0, -1), (1, 0, 0, 1)\}$$

eine Orthonormalbasis zu machen.
7. Betrachten Sie $\mathbb{R}^3$ als Euklidischen Vektorraum mit dem kanonischen Skalarprodukt. Sei $U \subseteq \mathbb{R}^3$ der Untervektorraum, der von den Vektoren $(0, 1, 2)$ und $(-1, 0, 1)$ erzeugt wird. Benutzen Sie das Gram-Schmidtsche Orthonormalisierungsverfahren, um eine Orthonormalbasis von U zu finden.
8. Sei P_2 der Vektorraum der reellen Polynome vom Grad ≤ 2 mit dem Skalarprodukt

$$\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)dx.$$

a) Benutzen Sie das Gram-Schmidtsche Orthonormalisierungsverfahren, um aus der Menge $\{1, x, x^2\}$ eine Orthonormalbasis zu machen. (Sie erhalten die ersten drei normalisierten Legendre[7] Polynome.)
b) Drücken Sie die folgenden Polynome als Linearkombinationen der gefundenen Legendre Polynome aus:

$$1 + x + 2x^2, 3 - 4x^2, 5 + 2x.$$

9. Zeigen Sie, daß in einem Euklidischen Vektorraum $||x|| = ||y||$ genau dann gilt, wenn $x + y$ und $x - y$ zueinander orthogonal sind.

8.3 Die Hessesche Normalform

Eine schöne Anwendung des Skalarprodukts in einem Euklidischen Vektorraum ist die Hessesche[8] Normalform, mit der ein affiner Unterraum beschrieben werden kann.

Definition und Lemma 8.3.1 *Sei V ein endlichdimensionaler Euklidischer Vektorraum. Sei $a + U$ ein affiner Unterraum von V, und $b'_1, \ldots, b'_k$ eine Orthonormalbasis von $U^\perp$. Dann gilt*

$$\begin{aligned} a + U &= \{x \in V | \forall z \in U^\perp : \langle z, x - a\rangle = 0\} \\ &= \{x \in V | \forall i = 1, \ldots, k : \langle b'_i, x - a\rangle = 0\}. \end{aligned}$$

Die Gleichungen

$$\langle u_i, x - a\rangle = 0$$

heißen Hessesche Normalform *von $a + U$.*

Beweis. $\langle z, x\rangle = \langle z, a\rangle \Longleftrightarrow \langle z, x - a\rangle = 0 \Longleftrightarrow x - a \in U^{\perp\perp} = U \Longleftrightarrow x \in a + U$.

Lemma 8.3.2 *Sei $b \in V$. Dann gibt es genau ein $x_\circ \in a + U$ mit $b - x_\circ \in U^\perp$.*

Beweis. $b - a \in U \oplus U^\perp \Longrightarrow b - a = u + u^\perp \Longrightarrow b - (a + u) = u^\perp$. Mit $x_\circ := a + u$ gilt $b - x_\circ = u^\perp \in U^\perp$. Sei $b - x \in U^\perp, b - x' \in U^\perp, x, x' \in a + U \Longrightarrow x - x' \in U^\perp \cap U = 0 \Longrightarrow x = x'$.

Aus der Beschreibung eines affinen Unterraumes erhalten wir jetzt leicht den Abstand von beliebigen Punkten zu ihm.

Satz 8.3.3 *Sei $b - x_\circ \in U^\perp, x_\circ \in a + U$. Dann gilt*

$$\|b - x_\circ\| = Min\{\|b - x\| | x \in a + U\}$$

$\|b - x_\circ\|$ heißt Abstand *von b nach $a + U$, $x_\circ$ heißt* Fußpunkt des Lotes *von b auf $a + U$.*

[7] Adrien Marie Legendre (1752-1833)
[8] Ludwig Otto Hesse (1811-1874)

Beweis. $\|b-x\|^2 = \|(b-x_\circ)+(x_\circ-x)\|^2 = \|b-x_\circ\|^2+\|x_\circ-x\|^2 \geq \|b-x_\circ\|^2$ wegen $b-x_\circ \in U^\perp, x_\circ - x \in U$.

Folgerung 8.3.4 $\|b-x_\circ\| = \|\sum_{i=1}^k \langle b-a, u'_i\rangle u'_i\| = \sqrt{\sum \langle u'_i, b-a\rangle^2}$ *ist der Abstand von* b *nach* $a+U$.

Beweis. 8.3.1 $\Longrightarrow b - x_\circ \in U^\perp$ und $b-a = u + (b-x_\circ) \Longrightarrow \langle b-a, u'_i\rangle = \langle b-x_\circ, u'_i\rangle \Longrightarrow \|b-x_\circ\| = \|\sum \langle b-x_\circ, u'_i\rangle u'_i\| = \|\sum\langle b-a, u'_i\rangle u'_i\|$.

Der Fall eines affinen Unterraumes der Kodimension 1 ist der bekannteste Fall der Hesseschen Normalform. Er ergibt sich leicht aus den bisherigen Überlegungen.

Bemerkung 8.3.5 Ist $\dim U = n-1$, so ist $\dim U^\perp = 1$. Ist u' eine Orthonormalbasis von $U^\perp$, so liegen zwei Punkte b, b' genau dann „auf derselben Seite" von $a+U$, wenn $\langle u', b-a\rangle$ und $\langle u', b'-a\rangle$ dasselbe Vorzeichen haben. $\langle x-a, u'\rangle = 0$ ist dann die Hessesche Normalform von $a+U$ und $|\langle b-a, u'\rangle|$ der Abstand von b zu $a+U$.

Übungen 8.3.6 1. Betrachten Sie $\mathbb{R}^3$ als Euklidischen Vektorraum mit dem kanonischen Skalarprodukt. Sei $U \subseteq \mathbb{R}^3$ der Untervektorraum, der von den Vektoren $(\frac{3}{5}, 0, \frac{4}{5})$ und $(0,1,0)$ erzeugt wird. Drücken Sie den Vektor $x = (1,2,3)$ als Summe von zwei Vektoren $x = y+z$ mit $y \in U$ und $z \in U^\perp$ aus.

2. Zeigen Sie, daß die Lösungsmenge der Gleichung $5x-3y+z = 2$ eine affine Ebene $A = a+U$ im Euklidischen Vektorraum $\mathbb{R}^3$ ist und bestimmen Sie die Hessesche Normalform dafür. Finden Sie den Fußpunkt von $x = (1,-2,4)$ in A und seinen Abstand zu A.
3. Im Euklidischen $\mathbb{R}^3$ sei die Gerade G durch
$$x = 2t, y = -t, z = 4t \quad (t \in \mathbb{R})$$
gegeben. Finden Sie den Fußpunkt von $(2,1,1)$ in G und seinen Abstand zu G.
4. Im $\mathbb{R}^3$ seien die Geraden $G_1 = a_1 + U_1$ und $G_2 = a_2 + U_2$ gegeben, wobei
$$a_1 = (1,0,1), \quad a_2 = (2,1,0),$$
$$U_1 = \langle(1,1,0)\rangle, U_2 = \langle(1,0,2)\rangle.$$
Finden Sie den Abstand zwischen G_1 und G_2 (d.h. das Minimum der Abstände $||p-q||$ für $p \in G_1$ und $q \in G_2$).
5. Sei V der Euklidische Vektorraum der stetigen Funktionen $f : [a,b] \longrightarrow \mathbb{R}$ mit dem Skalarprodukt
$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$
Sei $U \subseteq V$ ein endlichdimensionaler Untervektorraum und $p : V \longrightarrow U$ die orthogonale Projektion. Zeigen Sie, daß für $f \in V$ gilt

$$\mathrm{Min}_{g \in U} \int_a^b (f(x) - g(x))^2 dx = \int_a^b (f(x) - p(f(x)))^2 dx,$$

d.h. $p(f)$ ist die beste quadratische Approximation von f in U.

8.4 Isometrien

Wie bei allen bisher studierten Strukturen gibt es auch für die Struktur Euklidischer Vektorräume strukturerhaltende Abbildungen, die Isometrien genannt werden. Sie erhalten insbesondere Winkel und Längen.

Definition 8.4.1 Seien $(V, \langle , \rangle), (W, \langle , \rangle)$ Vektorräume mit Skalarprodukt. Eine Abbildung $f : V \longrightarrow W$ heißt *Isometrie (orthogonale Abbildung),* wenn

$$\forall x, y \in V : \qquad \langle f(x), f(y) \rangle = \langle x, y \rangle.$$

$\mathcal{O}(V, \langle , \rangle) = \{f \in \mathrm{GL}(V) | f \text{ Isometrie }\}$ ist eine Untergruppe von $\mathrm{GL}(V)$, genannt *orthogonale Gruppe.*

Satz 8.4.2 *Eine Isometrie $f : V \longrightarrow W$ ist ein Monomorphismus von Vektorräumen.*

Beweis. 1. Es ist $\langle f(\alpha x+\beta y)-\alpha f(x)-\beta f(y), f(\alpha x+\beta y)-\alpha f(x)-\beta f(y)\rangle = \langle f(\alpha x + \beta y), f(\alpha x + \beta y)\rangle + \alpha^2\langle f(x), f(x)\rangle + \beta^2\langle f(y), f(y)\rangle - 2\alpha\langle f(\alpha x + \beta y), f(x)\rangle - 2\beta\langle f(\alpha x + \beta y), f(y)\rangle - 2\alpha\beta\langle f(x), f(y)\rangle = \langle \alpha x + \beta y, \alpha x + \beta y\rangle + \alpha^2\langle x, x\rangle + \beta^2\langle y, y\rangle - 2\alpha\langle \alpha x+\beta y, x\rangle - 2\beta\langle \alpha x+\beta y, y\rangle - 2\alpha\beta\langle x, y\rangle = \langle \alpha x+\beta y - \alpha x - \beta y, \alpha x + \beta y - \alpha x - \beta y\rangle = 0$. Daraus folgt $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$. Also ist f ein Homomorphismus.

2. $f(x) = 0 \Longrightarrow 0 = \langle f(x), f(x)\rangle = \langle x, x\rangle \Longrightarrow x = 0$.

Satz 8.4.3 *Sei $f \in \mathrm{End}(V)$ und V ein Euklidischer Vektorraum. Gelte $\forall v \in V[\|f(v)\| = \|v\|]$. Dann ist f eine Isometrie.*

Beweis. Aus $\langle x, y\rangle = 1/2(\langle x + y, x + y\rangle - \langle x, x\rangle - \langle y, y\rangle) = 1/2(\|x + y\|^2 - \|x\|^2 - \|y\|^2)$, genannt Polarisierung, folgt

$$\begin{aligned}\langle f(x), f(y)\rangle &= \tfrac{1}{2}(\|f(x) + f(y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2)\\ &= \tfrac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2)\\ &= \langle x, y\rangle.\end{aligned}$$

Übungen 8.4.4 1. Drehen Sie den $\mathbb{R}^3$ um die x-Achse um den Winkel $\pi/3$ und bestimmen Sie den Bildvektor des Vektors $(1, -2, 4)$.

2. Im $\mathbb{R}^4$ mit den Koordinaten u, x, y, z werde zunächst eine Drehung um die $y - z$-Ebene mit dem Winkel $\pi/4$ und sodann eine Drehung um die $u - x$-Ebene um den Winkel $\pi/4$ vorgenommen. Wie sieht die gesamte Drehmatrix aus?
3. Im $\mathbb{R}^4$ mit den Koordinaten u, x, y, z werde zunächst eine Drehung um die $y - z$-Ebene mit dem Winkel $\pi/3$ und sodann eine Drehung um die $x - y$-Ebene um den Winkel $\pi/4$ vorgenommen. Wie sieht die gesamte Drehmatrix aus?

8.5 Orthogonale Matrizen

Die zu Isometrien gehörigen Matrizen sind die orthogonalen Matrizen als darstellende Matrizen bzgl. einer Orthonormalbasis. Wir wollen sie in diesem Abschnitt studieren.

Definition 8.5.1 Eine Matrix $M \in \mathrm{GL}(n, \mathbb{R})$ heißt *orthogonal*, wenn für die transponierte Matrix M^t gilt $M^t = M^{-1}$.

Bemerkung 8.5.2 Wenn M orthogonal ist, so folgt $MM^t = MM^{-1} = E \Longrightarrow 1 = \det E = \det(MM^t) = (\det M)^2 \Longrightarrow |\det(M)| = 1$. Eine orthogonale Matrix M heißt *eigentlich orthogonal*, wenn $\det M = 1$ gilt.

$\mathrm{O}(n) := \{M \in \mathrm{GL}(n, \mathbb{R}) | M^t = M^{-1}\}$ heißt *orthogonale Gruppe*.

$\mathrm{SO}(n) := \{M \in \mathrm{O}(n) | \det M = 1\}$ heißt *spezielle orthogonale Gruppe*.

Satz 8.5.3 *Sei V ein endlichdimensionaler Euklidischer Vektorraum mit Orthonormalbasis B. Ein Homomorphismus $f \in \mathrm{End}(V)$ ist genau dann eine Isometrie (orthogonale Abbildung), wenn die darstellende Matrix M von f bzgl. B eine orthogonale Matrix ist.*

Beweis. Seien $(\zeta_1, \ldots, \zeta_n)^t = \zeta$ bzw. $M \cdot \zeta$ die Koordinatenvektoren von $x \in V$ bzw. $f(x)$ bezüglich der Basis B. Sei $\eta = (\eta_1, \ldots, \eta_n)^t$ Koordinatenvektor von $y \in V$. Dann gilt

$$\langle x, y \rangle = \langle \sum_i \zeta_i b_i, \sum_j \eta_j b_j \rangle = \sum_{i,j} \zeta_i \eta_j \langle b_i, b_j \rangle = \zeta^t \eta = \langle \zeta, \eta \rangle$$

und damit

$$\langle x, y \rangle = \langle f(x), f(y) \rangle \Longleftrightarrow \zeta^t \eta = (M\zeta)^t M\eta = \zeta^t M^t M \eta.$$

Diese äquivalenten Aussagen gelten für alle $x, y \in V$ genau dann, wenn $M^t M = E_n$ ist. Man setze für die x bzw y Elemente aus B ein.

Bemerkung 8.5.4 Eine Transformationsmatrix für eine Basistransformation zwischen Orthonormalbasen ist also insbesondere eine orthogonale Matrix.

Wir kommen zurück auf die Projektion auf einen Untervektorraum, wie sie in (8.2.14) und (8.2.15) definiert wurde. Wenn $b_1, \ldots, b_n$ und $c_1, \ldots, c_n$ Orthonormalbasen sind und $c_i = \sum \alpha_{ij} b_j$ ist, dann ist

$$\sum_i \langle v, c_i \rangle c_i = \sum_{i,j,k} \langle v, \alpha_{ij} b_j \rangle \alpha_{ik} b_k = \sum_{j,k} \sum_i \alpha_{ij} \alpha_{ik} \langle v, b_j \rangle b_k = \sum_j \langle v, b_j \rangle b_k.$$

Also ist die Projektion $f : V \longrightarrow V$ von der Wahl der Orthonormalbasis von U unabhängig.

Satz 8.5.5 *Folgende Aussagen über $M \in \mathbb{R}_n^n$ sind äquivalent.*

1. *M ist orthogonal.*
2. *Die Spaltenvektoren von M bilden eine Orthonormalbasis für $\mathbb{R}_n$.*
3. *Die Zeilenvektoren von M bilden eine Orthonormalbasis für $\mathbb{R}^n$.*

Beweis. Wenn M orthogonal ist, dann ist auch die Transponierte M^t orthogonal. Deshalb gehen (2) $\Longleftrightarrow$ (3) ineinander durch Transponieren über. Weiter ist $M^tM = E_n \Longleftrightarrow \forall i,j[m_i^t m_j = \delta_{ij}] \Longleftrightarrow$ Spalten von M sind Orthonormalbasis

Eine orthogonale Matrix läßt sich durch Koordinatentransformation in eine besonders einfache Form überführen. Wir nennen diese Form Normalform einer orthogonalen Abbildung. Der Beweis ist etwas schwieriger. Wir verweisen den Leser hierzu auf spezielle Lehrbücher und verzichten hier auf eine Darstellung des Beweises.

Satz 8.5.6 (Normalform von orthogonalen Abbildungen)
Sei V ein endlichdimensionaler Euklidischer Vektorraum und $f \in \mathrm{End}(V)$ orthogonal. Dann gibt es eine Orthonormalbasis B, bezüglich der die darstellende Matrix M von f die Form hat

$$M = \begin{pmatrix} 1 & & & & & & & \cdots & 0 \\ & \ddots & & & & & & & \vdots \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & A_1 & & \\ \vdots & & & & & & & \ddots & \\ 0 & \cdots & & & & & & & A_r \end{pmatrix}$$

mit $A_i = \begin{pmatrix} \cos(\varphi_i) & -\sin(\varphi_i) \\ \sin(\varphi_i) & \cos(\varphi_i) \end{pmatrix}$

Ohne Beweis.

Folgerung 8.5.7 *Sei M eine orthogonale Matrix. Dann gibt es eine orthogonale Matrix T, so daß TMT^{-1} die Form hat*

$$TMT^{-1} = \begin{pmatrix} 1 & & & & & & & \cdots & 0 \\ & \ddots & & & & & & & \vdots \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & A_1 & & \\ \vdots & & & & & & & \ddots & \\ 0 & \cdots & & & & & & & A_r \end{pmatrix}$$

Beweis. M stellt bzgl. der Orthonormalbasis $e_1, \ldots, e_n$ eine Isometrie f dar. Eine Basistransformation von (e_i) zu einer anderen Orthonormalbasis B wird wegen 8.5.3 und 5.5.18 durch eine orthogonale Matrix T gegeben. Mit der Transformationsformel aus 5.5.20 folgt die Behauptung.

Ein interessante Anwendung der Normalform ist

Folgerung 8.5.8 *Eine Kugel im $\mathbb{R}_3$ werde um ihren Mittelpunkt (mit einer eigentlich orthogonalen Matrix) bewegt. Dann gibt es zwei Punkte auf der Oberfläche, die fest bleiben.*

Beweis. Die Matrix hat wegen $\det M = 1$ die Form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$$

und stellt daher eine Drehung um den (1−dimensionalen) Eigenraum zu 1 dar.

Satz 8.5.9 *Sei $M \in \mathbb{R}_n^n$. Dann gelten:*

1. *$M^t = M$ (M symmetrisch) $\Longrightarrow$ alle komplexen Eigenwerte von M sind reell.*
2. *$M^t = -M$ (M antisymmetrisch) $\Longrightarrow$ alle komplexen Eigenwerte von M sind rein imaginär.*
3. *$M^t = M^{-1}$ (M orthogonal) $\Longrightarrow$ alle komplexen Eigenwerte von M haben den Betrag 1.*

Beweis. 1. und 2.: Operiere M auf $\mathbb{C}_n$. Sei $\epsilon = \pm 1$ und gelte $M^t = \epsilon M$. Sei λ ein Eigenwert von M mit Eigenvektor z. Dann gilt $\lambda \overline{z}^t z = \overline{z}^t \lambda z = \overline{z}^t M z = \epsilon \overline{z}^t \overline{M}^t z = \epsilon (\overline{M z})^t z = \epsilon \overline{\lambda z}^t = \epsilon \overline{\lambda}\overline{z}^t z$, also $\lambda = \epsilon \overline{\lambda}$, wobei $\overline{z} \in \mathbb{C}_n$ der konjugiert komplexe Vektor zu z ist und $\overline{M} = M$ gilt, wenn $M \in \mathbb{R}_n^n$. Ist $\epsilon = 1$, so ist λ wegen $\lambda = \overline{\lambda}$ reell. Ist $\epsilon = -1$, so ist λ wegen $\lambda = -\overline{\lambda}$ rein imaginär.

3.: Aus $\overline{z}^t z = \overline{z}^t E_n z = \overline{z}^t M^t M z = \overline{z}^t \overline{M}^t M z = (\overline{M z})^t M z = (\overline{\lambda z})^t \lambda z = \overline{\lambda}\lambda \overline{z}^t z$ folgt $\overline{\lambda}\lambda = 1$, also hat λ den Betrag 1.

Um beliebige Euklidische Vektorräume zu gewinnen, braucht man ein Skalarprodukt. Im $\mathbb{R}_n$ ist ein Skalarprodukt immer gegeben in der Form $\langle b, c \rangle = b^t S c$ mit einer Matrix S, die positiv definit ist.

Definition 8.5.10 Ein Matrix $S \in \mathbb{R}_n^n$ heißt *positiv definit*, wenn $\mathbb{R}_n \times \mathbb{R}_n \ni (x, y) \mapsto x^t S y \in \mathbb{R}$ eine positiv definite symmetrische Bilinearform ist.

Satz 8.5.11 *Sei $S \in \mathbb{R}_n^n$. S ist genau dann positiv definit, wenn es eine invertierbare Matrix M gibt mit $S = M^t M$.*

Beweis. Sei S positiv definit. Dann gibt es eine Orthonormalbasis B für $\mathbb{R}_n$ mit dem durch S definierten Skalarprodukt. Wegen $b_i^t S b_j = \delta_{ij}$ ist $(b_1, \ldots, b_n)^t S (b_1, \ldots, b_n) = E_n$ und mit $M = (b_1, \ldots, b_n)^{-1}$ folgt $S = M^t M$. Wenn $S = M^t M$ gilt, dann ist $(b_1, \ldots, b_n)^t S (b_1, \ldots, b_n) = E_n$, wenn man $(b_1, \ldots, b_n) = M^{-1}$ definiert. Also ist $b_i^t S b_j = \delta_{ij}$. Damit wird die durch S definierte Bilinearform ein Skalarprodukt und S positiv definit.

Übungen 8.5.12 1. Vervollständigen Sie die Matrix

$$\begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & * \\ 1 & 0 & * \\ 0 & \frac{1}{\sqrt{2}} & * \end{pmatrix}$$

zu einer orthogonalen Matrix.

2. Vervollständigen Sie die Matrix

$$\begin{pmatrix} -\frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ * & * & * \end{pmatrix}$$

zu einer orthogonalen Matrix.

3. Sei

$$M = \begin{pmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Zeigen Sie, daß M invertierbar ist und bestimmen Sie die gemäß Satz 8.5.11 zugehörige positiv definite Matrix S. Bestimmen Sie den Winkel zwischen den Vektoren $(1, 0, 0)$ und $(0, 1, 0)$ bzgl. des durch S definierten Skalarprodukts auf $\mathbb{R}^3$.

4. Zeigen Sie, daß

$$S = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 5 & 3 \\ 2 & 3 & 3 \end{pmatrix}$$

eine positiv definite Matrix ist.

8.6 Adjungierte Abbildungen

Das Skalarprodukt eines endlichdimensionalen Euklidischen Vektorraumes V induziert eine Vielzahl von interessanten Isomorphismen. Zunächst betrachten wir einen kanonischen Isomorphismus zwischen V und und dem dualen Vektorraum $V^* = \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R})$.

Lemma 8.6.1 *Sei σ eine nichtausgeartete Bilinearform auf V. Dann ist $\sigma^* : V \ni x \mapsto \sigma(x, -) \in V^* = \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R})$ ein Monomorphismus.*

Beweis. Da σ bilinear ist, ist $\sigma(x,-)$ im zweiten Argument linear, also ein Element von $\mathrm{Hom}(V,\mathbb{R}) = V^*$. Weiter ist $\sigma^*(\alpha x+\beta y,-)(z) = \sigma(\alpha x+\beta y, z) = \alpha\sigma(x,z) + \beta\sigma(y,z) = \alpha\sigma^*(x)(z) + \beta\sigma^*(y)(z) = (\alpha\sigma^*(x) + \beta\sigma^*(y))(z)$, also gilt $\sigma^*(\alpha x+\beta y) = \alpha\sigma^*(x)+\beta\sigma^*(y)$. σ^* ist also ein Homomorphismus. Wenn $\sigma^*(x) = 0$ ist, dann ist $\sigma^*(x)(z) = \sigma(x,z) = 0$ für alle $z \in V$. Weil σ nicht ausgeartet ist, folgt $x = 0$. Damit ist σ^* ein Monomorphismus.

Folgerung 8.6.2 *Sei σ eine nichtausgeartete Bilinearform auf einem endlichdimensionalen Vektorraum V. Dann ist $\sigma^* : V \longrightarrow V^*$ ein Isomorphismus.*

Beweis. Nach 5.4.16 ist $\dim V = \dim V^*$, also ist σ^* nach 5.4.12 ein Isomorphismus.

Eine weitere wichtige Konstruktion ist die von adjungierten Endomorphismen. Diese Konstruktion führt zu einem Automorphismus von $\mathrm{End}(V)$.

Satz 8.6.3 *Seien σ und V wie in 8.6.2. Zu jedem $f \in \mathrm{End}(V)$ gibt es genau ein $f^{\mathrm{ad}} \in \mathrm{End}(V)$ mit*

$$\forall x, y \in V[\sigma(x, f(y)) = \sigma(f^{\mathrm{ad}}(x), y)].$$

f^{ad} heißt die zu f adjungierte Abbildung.

Beweis. Es ist $\sigma(x, f(-)) \in V^*$. Also gibt es nach 8.6.2 genau ein Element $f^{\mathrm{ad}}(x) \in V$ mit $\sigma^*(f^{\mathrm{ad}}(x)) = \sigma(f^{\mathrm{ad}}(x), -) = \sigma(x, f(-))$. $f^{\mathrm{ad}}(x)$ hängt linear von x ab, denn $\sigma(f^{\mathrm{ad}}(\alpha x + \beta y), z) = \sigma(\alpha x + \beta y, f(z)) = \alpha\sigma(x, f(z)) + \beta\sigma(y, f(z)) = \alpha\sigma(f^{\mathrm{ad}}(x), z) + \beta\sigma(f^{\mathrm{ad}}(y), z) = \sigma(\alpha f^{\mathrm{ad}}(x) + \beta f^{\mathrm{ad}}(y), z)$ für alle $z \in V$. Also ist $\sigma^*(f^{\mathrm{ad}}(\alpha x+\beta y)) = \sigma^*(\alpha f^{\mathrm{ad}}(x)+\beta f^{\mathrm{ad}}(y))$, also $f^{\mathrm{ad}}(\alpha x+\beta y) = \alpha f^{\mathrm{ad}}(x) + \beta f^{\mathrm{ad}}(y)$. Offenbar ist f^{ad} eindeutig bestimmt.

Folgerung 8.6.4 *Sei σ eine symmetrische, nichtausgeartete Bilinearform auf dem endlichdimensionalen Vektorraum V. Dann ist die Abbildung*

$$\mathrm{End}(V) \ni f \mapsto f^{\mathrm{ad}} \in \mathrm{End}(V)$$

ein involutorischer Antiautomorphismus des Endomorphismenringes (versuchen Sie mit dieser Aussage Ihre Freunde zu beeindrucken), d.h.

1. $(f + g)^{\mathrm{ad}} = f^{\mathrm{ad}} + g^{\mathrm{ad}}$,
2. $(f \cdot g)^{\mathrm{ad}} = g^{\mathrm{ad}} \cdot f^{\mathrm{ad}}$,
3. $(\mathrm{id})^{\mathrm{ad}} = \mathrm{id}$,
4. $(f^{\mathrm{ad}})^{\mathrm{ad}} = f$.

Beweis. 1. $\sigma((f + g)^{\mathrm{ad}}(x), y) = \sigma(x, (f + g)(y)) = \sigma(x, f(y)) + \sigma(x, g(y)) = \sigma(f^{\mathrm{ad}}(x), y) + \sigma(g^{\mathrm{ad}}(x), y) = \sigma((f^{\mathrm{ad}} + g^{\mathrm{ad}})(x), y)$ für alle $y \in V$ impliziert $\sigma^*((f + g)^{\mathrm{ad}}(x)) = \sigma^*((f^{\mathrm{ad}} + g^{\mathrm{ad}})(x))$, also folgt 1.

2. $\sigma((f \cdot g)^{\mathrm{ad}}(x), y) = \sigma(x, f(g(y))) = \sigma(f^{\mathrm{ad}}(x), g(y)) = \sigma(g^{\mathrm{ad}}(f^{\mathrm{ad}}(x)), y)$ $\Longrightarrow$ 2.

3. trivial
4. $\sigma((f^{\mathrm{ad}})^{\mathrm{ad}}(x),y) = \sigma(x,f^{\mathrm{ad}}(y)) = \sigma(f^{\mathrm{ad}}(y),x) = \sigma(y,f(x)) = \sigma(f(x),y)$, also $(f^{\mathrm{ad}})^{\mathrm{ad}} = f$.

Mit der Konstruktion von adjungierten Endomorphismen ergibt sich die interessante Frage, unter welchen Umständen f und f^{ad} übereinstimmen. Solche Endomorphismen nennen wir selbstadjungiert. Sie haben sehr weitreichende Bedeutung in der Physik, der Geometrie und in der Analysis.

Definition 8.6.5 $f \in \mathrm{End}(V)$ heißt *selbstadjungiert*, wenn $f^{\mathrm{ad}} = f$ gilt.

Lemma 8.6.6 $\mathrm{Bil}(V) := \{\sigma : V \times V \longrightarrow \mathbb{R} | \sigma$ *Bilinearform* $\}$ *ist ein Vektorraum.*

Beweis. $\mathrm{Bil}(V) \subseteq \mathrm{Abb}(V \times V, \mathbb{R})$ ist ein Untervektorraum, wie man leicht sieht.

Satz 8.6.7 *Sei σ eine nichtausgeartete symmetrische Bilinearform auf V und* $\dim V < \infty$. *Dann ist*

$$\Psi_\sigma : \mathrm{End}(V) \ni f \mapsto \sigma(-, f(-)) \in \mathrm{Bil}(V)$$

ein Isomorphismus.

Beweis. Offenbar ist $\sigma(-, f(-))$ eine Bilinearform. Weiter ist wegen $\Psi_\sigma(\alpha f + \beta g) = \sigma(-, (\alpha f + \beta g)(-)) = \alpha\sigma(-, f(-)) + \beta\sigma(-, g(-)) = \alpha\Psi_\sigma(f) + \beta\Psi_\sigma(g)$ die Abbildung Ψ_σ ein Homomorphismus. Wenn $\Psi_\sigma(f) = 0$ ist, so ist $\forall x, y \in V[\sigma(x, f(y)) = 0]$. Also gilt $\forall y \in V[f(y) = 0]$ und damit $f = 0$. Daher ist Ψ_σ injektiv. Sei schließlich $\tau \in \mathrm{Bil}(V)$. Dann ist $\tau(x, -) \in V^*$, also gibt es genau ein $f(x) \in V$ mit $\tau(x, -) = \sigma(f(x), -)$. Es ist $f \in \mathrm{End}(V)$, denn $\sigma(f(\alpha x + \beta y), -) = \tau(\alpha x + \beta y, -) = \alpha\tau(x, -) + \beta\tau(y, -) = \alpha\sigma(f(x), -) \in \beta\sigma(f(y), -) = \sigma(\alpha f(x) + \beta f(y), -)$. Wir bilden $\Psi_\sigma(f^{\mathrm{ad}}) = \sigma(-, f^{\mathrm{ad}}-) = \sigma(f-, -) = \tau(-, -)$ und haben damit Ψ_σ surjektiv gezeigt.

Bemerkung 8.6.8 Auf jedem V mit $\dim V < \infty$ gibt es eine nichtausgeartete symmetrische Bilinearform. Insbesondere ist ein Skalarprodukt nichtausgeartet. Sei nämlich $b_1, \ldots, b_n$ eine Basis von V. Dann ist $\sigma(\sum \alpha_i b_i, \sum \beta_j b_j) := \sum \alpha_i \beta_j \sigma(b_i, b_j) = \sum \alpha_i \beta_j \delta_{ij}$ nicht ausgeartet und symmetrisch, denn $\sigma(x, V) = 0$ impliziert $\sigma(\sum \alpha_i b_i, b_j) = 0 = \alpha_j$ für alle j.

Wir übertragen unsere bisherigen Betrachtungen nunmehr auf Matrizen.

Satz 8.6.9 *Sei* $\dim V = n < \infty, B$ *eine Basis und* $\sigma(b_i, b_j) = \delta_{ij}$ *eine nichtausgeartete symmetrische Bilinearform. Unter den Isomorphismen*

$$\mathbb{R}^n_n \xrightarrow{\Phi} \mathrm{End}(V) \xrightarrow{\Phi_\sigma} \mathrm{Bil}(V)$$

mit $\Phi(M)$ der von M bzgl. B dargestellten Homomorphismus gilt:

1. *äquivalent sind für* $M \in \mathbb{R}_n^n$
   a) $M^t = M$ *(*M *ist* symmetrisch*),*
   b) $\Phi(M)$ *ist selbstadjungiert,*
   c) $\Psi_\sigma\Phi(M)$ *ist symmetrisch;*
2. *äquivalent sind für* $M \in \mathbb{R}_n^n$ *:*
   a) M *regulär,*
   b) $\Phi(M)$ *Automorphismus,*
   c) $\Psi_\sigma\Phi(M)$ *nicht ausgeartet.*

Beweis. 1. Sei $f := \Phi(M)$, also $f(b_j) = \sum_{i=1}^n \alpha_{ij} b_i$ mit $M = (\alpha_{ij})$. Es gilt

$$\begin{aligned} M^t = M &\iff \forall i,j[\alpha_{ij} = \alpha_{ji}] \\ &\iff \forall i,j[\sigma(\textstyle\sum_k \alpha_{ki} b_k, b_j) = \alpha_{ji} = \sigma(b_i, \sum_k \alpha_{kj} b_k)] \\ &\iff \forall i,j[\sigma(f(b_i), b_j) = \sigma(b_i, f(b_j))] \\ &\iff \sigma(f(-),-) = \sigma(-,f(-)) \\ &\iff f^{\mathrm{ad}} = f \\ &\iff \forall x,y[\sigma(f(x),y) = \sigma(x,f(y))] \\ &\iff \forall x,y[\sigma(y,f(x)) = \sigma(x,f(y))] \\ &\iff \Psi_\sigma(f) \text{ symmetrisch.} \end{aligned}$$

2. M regulär $\iff f$ bijektiv

$$\begin{aligned} &\iff \forall y[f(y) = 0 \Longrightarrow y = 0] \\ &\iff \forall y[\Psi_\sigma(f)(V,y) = \sigma(V,f(y)) = 0 \Longrightarrow y = 0] \\ &\qquad \text{(1. Hälfte von } \Psi_\sigma(f) \text{ nichtausgeartet)} \\ &\iff f^{\mathrm{ad}} \text{ bijektiv (8.6.4)} \\ &\iff \forall y[\Psi_\sigma(f)(y,V) = \sigma(y,f(V)) = \sigma(f^{\mathrm{ad}}(y),V) = 0 \Longrightarrow \\ &\qquad y = 0] \text{(2. Hälfte von } \Psi_\sigma(f) \text{ nichtausgeartet).} \end{aligned}$$

Folgerung 8.6.10 *In* $\mathbb{R}_n$ *ist* $\tau : \mathbb{R}_n \times \mathbb{R}_n \ni (x,y) \mapsto x^t M y \in \mathbb{R}$ *eine Bilinearform und jede Bilinearform auf* $\mathbb{R}_n$ *ist von dieser Gestalt.* τ *ist genau dann symmetrisch (nichtausgeartet), wenn* M *symmetrisch (regulär) ist.*

Beweis. Für die Basis $B = (e_1, \ldots, e_n)$ und σ wie in 8.6.8 gilt

$$\Psi_\sigma\Phi(M)(e_i,e_j) = \sigma(e_i \sum_k \alpha_{kj} e_k) = \alpha_{ij} = e_i^t M e_j = \tau(e_i,e_j).$$

Übungen 8.6.11 1. Sei $\langle x,y\rangle$ das kanonische Skalarprodukt auf $\mathbb{R}^n$ und sei A eine $n \times n$-Matrix. Zeigen Sie

$$\langle x, Ay\rangle = \langle A^t x, y\rangle.$$

2. a) Zeigen Sie, daß auf R_2 durch $x^t S y$ ein Skalarprodukt definiert ist mit $S = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$.
   b) Bestimmen Sie bezüglich dieses Skalarprodukts die adjungierte Matrix zu $M = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}$.

8.7 Die Hauptachsentransformation

Wir haben schon darauf hingewiesen, daß selbstadjungierte Endomorphismen eine Vielzahl von wichtigen Anwendungen haben. Daher ist es von besonderem Interesse, ihre darstellenden Matrizen zu studieren und sie durch geeignete Wahl der Basis in eine besonders einfache Form zu bringen. Hier begegnen wir noch einmal Eigenwerten und Eigenvektoren. Die einfachste Form der darstellenden Matrizen stellt sich als Diagonalform heraus.

Satz 8.7.1 *(Hauptachsentransformation von selbstadjungierten Endomorphismen) Sei V ein endlichdimensionaler Euklidischer Vektorraum und sei $f \in \mathrm{End}(V)$ selbstadjungiert. Dann gibt es eine Orthonormalbasis B aus Eigenvektoren von f und f ist diagonalisierbar.*

Beweis. durch vollständige Induktion nach $n = \dim V$.

Für $n = 0$ ist nichts zu zeigen.

Gelte der Satz für $n-1$ und sei $\dim V = n$. Das charakteristische Polynom $\chi_f(x)$ hat (nach dem Fundamentalsatz der Algebra) n komplexe Nullstellen, die alle reell sind (8.5.8 1.). Sei $\lambda \in \mathbb{R}$ ein Eigenwert zu f mit einem Eigenvektor b_n mit $\|b_n\| = 1$. Sei $U = (\mathbb{R}b_n)^\perp$. Dann ist $\dim U = n-1$ ein Euklidischer Vektorraum. Weiter ist $f(U) \subseteq U$, denn für $u \in U$ gilt $\langle f(u), b_n\rangle = \langle u, f(b_n)\rangle = \langle u, \lambda b_n\rangle = \lambda\langle u, b_n\rangle = 0$. Es ist $f|_U$ selbstadjungiert, weil $\langle u_1, f(u_2)\rangle = \langle f(u_1), u_2\rangle$ gilt. Sei $b_1, \ldots, b_{n-1}$ eine Orthonormalbasis von U aus Eigenvektoren von $f|_U$. Dann ist $b_1, \ldots, b_n$ eine Orthonormalbasis von V aus Eigenvektoren von f.

Satz 8.7.2 *(Hauptachsentransformation für symmetrische Matrizen) Sei $S \in \mathbb{R}_n^n$ eine symmetrische Matrix. Dann gibt es eine orthogonale Matrix T, so daß TST^{-1} Diagonalform hat. Die Einträge in der Diagonalen von*

$$T^t S T = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

sind die Eigenwerte von S und die $t_1, \ldots, t_n$ mit $T = (t_1, \ldots, t_n)$ die zugehörigen Eigenvektoren.

Beweis. $\widehat{S} : \mathbb{R}_n \longrightarrow \mathbb{R}_n$ ist selbstadjungiert (8.6.9), folglich existiert eine Orthonormalbasis $t_1, \ldots, t_n$ aus Eigenvektoren: $St_j = \lambda_j t_j = t_j \cdot \lambda_j$ oder

$$ST = T \cdot \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

oder $T^{-1}ST = D$ oder $T^t ST = D$, weil T orthogonal ist (wobei D die Diagonalmatrix ist).

Der Zusammenhang zwischen symmetrischen Matrizen und Bilinearformen bringt uns zu der Frage, unter welchen Umständen die Bilinearformen Skalarprodukte sind. Diese Frage können wir zum Abschluß leicht beantworten.

Folgerung 8.7.3 *Eine symmetrische Matrix ist genau dann positiv definit, wenn alle Eigenwerte positiv sind.*

Beweis. Sei S positiv definit. Sei $b_1, \ldots, b_n$ wie in 8.7.2. Dann ist $\lambda_i = b_i^t S b_i = \|b_i\|_S^2 > 0$. Seien alle $\lambda_i > 0$, so ist für $x = \sum \zeta_i b_i$, das Skalarprodukt bzgl. S gegeben durch

$$\begin{aligned}\langle x, x\rangle_S &= (\sum \zeta_i b_i)^t S(\sum \zeta_j b_j) = \sum \zeta_i \zeta_j b_i^t S b_j \\ &= \sum \zeta_i \zeta_j \lambda_j b_i^t b_j = \sum \zeta_i^2 \lambda_i > 0.\end{aligned}$$

Übungen 8.7.4 1. Finden Sie die Dimensionen der Eigenräume der folgenden symmetrischen Matrizen:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & -4 & 2 \\ -4 & 1 & -2 \\ 2 & -2 & -2 \end{pmatrix}, \quad \begin{pmatrix} -\frac{5}{3} & -\frac{4}{3} & 0 & \frac{1}{3} \\ -\frac{4}{3} & \frac{10}{3} & 0 & -\frac{4}{3} \\ 0 & 0 & -2 & 0 \\ \frac{1}{3} & -\frac{4}{3} & 0 & -\frac{5}{3} \end{pmatrix}$$

2. Transformieren Sie die folgenden Matrizen auf Hauptachsen und geben Sie die Transformationsmatrix an:

$$\begin{pmatrix} 5 & 3\sqrt{3} \\ 3\sqrt{3} & -1 \end{pmatrix}, \quad \begin{pmatrix} -2 & 0 & -36 \\ 0 & -3 & 0 \\ -36 & 0 & -23 \end{pmatrix},$$

$$\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 5 & 2 & 0 & 0 \\ 2 & -2 & 0 & 0 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 2 & 5 \end{pmatrix}.$$